Attentional Suppression of Highly Salient Color Singletons

MANUSCRIPT IN PRESS AT *JOURNAL OF EXPERIMENTAL PSYCHOLOGY: HUMAN PERFORMANCE AND PERCEPTION*

Brad T. Stilwell and Nicholas Gaspelin
State University of New York at Binghamton

Abstract

A longstanding debate in visual attention research has been whether physically salient objects have an automatic power to capture attention. Recent evidence has supported a hybrid model. According to the signal suppression hypothesis, salient items automatically attract attention, but can be proactively suppressed via top-down control to prevent attentional capture. Although much recent evidence has suggested that salient items can be suppressed, many of these studies used color singletons with relatively low salience. It is therefore unknown whether highly salient color singletons can also be suppressed. The current study adapted the probe technique to assess capture by color singletons at large set sizes (10 or 30 items). In four experiments, we observed no evidence that highly salient color singletons captured attention and instead observed evidence that they were suppressed below baseline levels of processing. We did, however, find strong evidence of floor effects in probe report at high set sizes, which can be mitigated by limiting the number of items that are simultaneously probed. Altogether, the results support the signal suppression hypothesis.

Public Significance Statement

An important question is whether physically salient objects, such brightly colored objects or flashing lights, have an automatic ability to attract visual attention. The current study demonstrated that visual distraction by highly salient objects can be prevented via a learned suppressive process.

Keywords

visual attention
attentional capture
suppression
salience

Brad T. Stilwell and Nicholas Gaspelin, Department of Psychology, State University of New York at Binghamton.

This project was supported by the National Science Foundation Grant BCS-2045624 to N.G. All data and stimulus presentation programs in this article are available on the Open Science Framework at https://osf.io/bwg2s.

Correspondence concerning this article should be directed to Brad T. Stilwell, Department of Psychology, State University of New York at Binghamton, PO Box 6000, Binghamton, NY 13902-6000, email: stilwell@binghamton.edu, phone: (607) 777-4304.

ATTENTIONAL SUPPRESSION OF HIGHLY SALIENT COLOR SINGLETONS

Our visual system is tasked with the monumental responsibility of determining which sensory information should be attended and which should be ignored. A longstanding debate has been whether physically salient stimuli, such as brightly colored objects or flashing lights, can automatically attract attention even when they are irrelevant to our goals (for a review, Luck et al., 2021). Although this question may seem relatively straightforward to answer, research on *attentional capture* and the conditions under which it arises have been heatedly debated. Recent evidence has suggested a potential resolution: salient distractors do attract attention, but they can be proactively suppressed to prevent attentional capture (Gaspelin & Luck, 2018c, 2019). Although this potential resolution is promising, an important disagreement remains regarding the boundary conditions of proactive inhibition. Namely, it has recently been questioned whether top-down suppression can overpower stimuli with high bottom-up salience (Wang & Theeuwes, 2020; see also Theeuwes, 2004). More specifically, it has been proposed that previous evidence of proactive suppression of attentional capture may have only occurred because the eliciting stimuli had low physical salience. This question will be the focus of the current study.

THE ATTENTIONAL CAPTURE DEBATE

Researchers have long debated whether physically salient objects can automatically capture visual attention. Initial studies seemed to indicate that certain classes of physically salient stimuli have an automatic power to attract visual attention (Theeuwes, 1991, 1992; Yantis & Jonides, 1984). For example, Theeuwes (1992) had participants perform an *additional singleton paradigm* in which they searched for a target shape (e.g., a diamond) amongst a homogenous set of shapes (e.g., circles) and reported the tilt of a line inside the target. Crucially, on some trials, one distractor item was uniquely colored from the others. This *color singleton* never appeared at the target location, and thus should have been ignored by the participants. However, mean response times (RTs) were slower on singleton-present trials than singleton-absent trials (Theeuwes, 1991, 1992, 2004). This *singleton presence cost* was taken to suggest that visual attention was automatically misdirected to the color singleton, slowing detection of the target stimulus when it was present. This finding and others led to the formulation of *stimulus-driven accounts* of attentional capture which propose that attention is involuntarily attracted to physically salient stimuli, even when they are entirely task-irrelevant.

Subsequent studies, however, provided evidence against such accounts by indicating that salient stimuli do not always capture attention. According to *goal-driven accounts*, when an observer is looking for a target object, the observer will establish an *attentional set* for the features of the target (e.g., redness) and only salient items matching this attentional set will capture attention. Importantly, goal-driven accounts provided evidence that initial evidence of capture in the additional singleton paradigm may have been due to an inadvertent attentional set for the singleton. In Theeuwes (1992), the target was a shape singleton and this may have encouraged participants to search for any uniquely featured “pop out” item. This attentional set for any feature singleton (*singleton detection mode*) may have led to capture by the color singleton. As evidence of this, later studies adapted the additional singleton paradigm to discourage singleton detection mode by having the target appear amongst heterogeneously shaped distractors (Bacon & Egeth, 1994). This *feature search mode* completely eliminated singleton presence costs, suggesting that the singleton no longer captured attention when it was made task-irrelevant. This study and many others led to the conclusion that physically salient distractors do not automatically capture attention (Egeth et al., 2010; Folk et al., 2002; Lamy & Egeth, 2003; Leber & Egeth, 2006b; Lien et al., 2008, 2010).

THE SIGNAL SUPPRESSION HYPOTHESIS

The debate between these opposing accounts of attentional capture has persisted for several decades. More recent research has suggested a potential resolution based upon an inhibitory mechanism of visual attention. According to the *signal suppression hypothesis* (Gaspelin & Luck, 2018c, 2019), all salient stimuli generate a bottom-up salience signal that automatically attracts visual attention. However, salient stimuli can be proactively suppressed to prevent attentional capture. This model has garnered much recent support with converging evidence from studies of psychophysics (Chang & Egeth, 2019; Gaspelin et al., 2015; Vatterott et al., 2018; Won et al., 2019; Won & Geng, 2020), eye movements (Gaspelin et al., 2017, 2019; Gaspelin & Luck, 2018b), event-related potentials (Feldmann-Wüstefeld et al., 2020; Gaspar et al., 2016; Gaspar & McDonald, 2014; Gaspelin & Luck, 2018a; Sawaki & Luck, 2010; Weaver et al., 2017), and single-unit recordings in monkeys (Cosman et al., 2018). There is now an emerging consensus that participants can learn to suppress salient items based upon their simple features (Gaspelin et al., 2019; Gaspelin & Luck, 2018b; Stilwell et al., 2019) and locations (Maxwell et al., 2020; Ruthruff & Gaspelin, 2018; Wang & Theeuwes, 2018a, 2018b).

One important line of evidence for the signal suppression hypothesis has come from studies using the *capture-probe paradigm* (Gaspelin et al., 2015; Gaspelin & Luck, 2018a, 2018b). In a key study, Gaspelin et al. (2015) had participants perform a task that consisted of two types of randomly intermixed trials. On *search trials*, participants searched for a target shape (e.g., green diamond) amongst heterogeneous distractors and made a speeded button press to indicate the location of a dot inside (left or right). On half of trials, a color singleton appeared at a nontarget location. Importantly, on *probe trials*, a search array briefly appeared with a letter superimposed inside of each shape for 100 ms and the letters were then immediately replaced by a pattern mask (#) for 500 ms. Afterwards, participants were asked to report as many letters as possible. Participants were less likely to report probe letters at the singleton distractor location than at the average nonsingleton distractor location. This *probe suppression effect* was taken to suggest that the color singleton did not capture attention and was instead suppressed below baseline levels of processing (see also Chang & Egeth, 2019; Feldmann-Wüstefeld et al., 2020). These results have been taken as support for the signal suppression hypothesis, which proposes that salient color singletons are proactively suppressed to prevent attentional capture.

CAN HIGHLY SALIENT SINGLETONS BE SUPPRESSED?

The results of Gaspelin et al. (2015), however, have recently been questioned by stimulus-driven accounts. Specifically, it has been suggested that the color singletons were not highly salient in this study and, as a result, the singleton could not overpower suppressive mechanisms (Wang & Theeuwes, 2020; see commentary by Theeuwes in Luck et al., 2021). The singleton distractor appeared in search displays that had either a set size of 4 or 6 items. This may have yielded a weak "pop out" by the color singleton, which contrasted with only a few homogenously colored items. If the color singleton were made more salient, perhaps participants would have been unable to be suppress it.

To support this claim, Wang and Theeuwes (2020) had participants perform a capture-probe paradigm similar to Gaspelin et al. (2015), but the search display was manipulated to include set sizes of 4, 6, and 10 (see Figure 1). The logic was that the color singleton should be more salient in set size 10 displays than set size 4 displays because there are more homogeneously colored items to contrast with the singleton. At set size 4, a probe suppression effect was observed, consistent with Gaspelin et al. (2015). At set size 10, however, the singleton no longer produced a probe suppression effect. Instead, participants were slightly more likely to recall the letter at the singleton distractor (18.1%) than the average nonsingleton distractor (16.2%): a 1.9% probe capture effect. At first glance, these results seem to provide support to the claim that highly salient color singletons can overpower top-down suppression, consistent with stimulus-driven accounts of attentional capture.

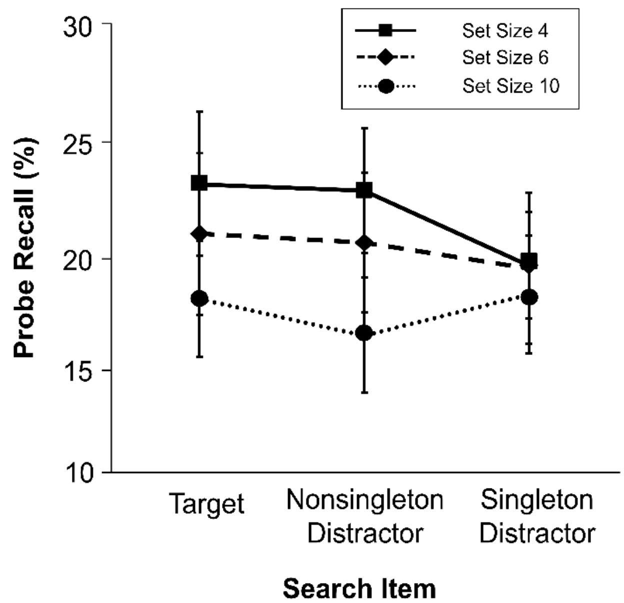
THE FLOOR EFFECT PROBLEM

There is an important caveat to the methodology of Wang and Theeuwes (2020). As the set size of the search displays were increased, the number of probe letters were also increased. Increasing the number of probe letters will reduce overall recall accuracy and potentially lead to floor effects that make suppression effects difficult to observe. In fact, this issue was directly mentioned by Gaspelin et al. (2015) when introducing the capture-probe paradigm: "…the magnitude of [the suppression] effect was limited by the already-low accuracy for probes at nonsingleton distractor locations. Our goal in Experiment 3 was to investigate whether the singleton-suppression effect can be increased by bringing overall accuracy away from floor, which we achieved by decreasing the set size from 6 to 4." (p. 7)

The *floor effect problem* occurs because there are limits on the number of probe letters that can be encoded in this paradigm. On the typical probe trial, approximately 2 letters are reported, and 1.5 of these letters are actually present in the probe array (Gaspelin et al., 2015). As a consequence, probe recall accuracy will be higher at a set size of 4 (1.5 of 4 letters = 37.5%) than at a set size of 10 (1.5 of 10 letters = 15%). Assuming that suppression reflects a reduction in probe recall accuracy at the location of the irrelevant singleton, the magnitude of the suppression effect would be expected to become smaller as the set size increases (all else being equal). For example, imagine that suppression reflects a 25% decrease in the probability of encoding the probe at the salient item relative to the average nonsalient distractor. The maximum observable suppression effect will be lower at set size 10 (15% × 25% = 3.8%) than set size 4 (37.5% × 25% = 9.4%). Thus, the magnitude of the suppression effect will be reduced at high set sizes even if the underlying base rate of suppression remains constant. This is demonstrated via a simulation in the online supplemental materials.

Consistent with a potential floor effect, the set size manipulation of Wang and Theeuwes (2020) decreased overall probe recall accuracy across search items (Figure 1). This compression of probe recall accuracy across search items is remarkably consistent with a capacity limitation on the number of probe letters that can be stored for later recall. As the set size was increased, this limitation will reduce overall probe recall accuracy for the nonsingleton distractor. This will, in turn, reduce the magnitude of the observable suppression effect. Assuming some level of measurement error, this could have made probe

Figure 1. *Probe Recall Accuracy from Wang and Theeuwes (2020).*



Note. As set size was increased, the magnitude of the probe suppression effect was reduced. This was taken as evidence of stimulus-driven capture. However, the reduction of the probe suppression effect across set sizes could also be due to a floor effect in probe report. Error bars represent 95% confidence intervals.

suppression effects difficult to detect.

It is important to highlight that performance does not have to be at chance levels to begin to suffer from floor-like compression. Chance performance can be estimated as the number of letters reported divided the number of letters possible. For example, in Wang and Theeuwes (2020), approximately 2.8 letters were reported on the average trial. Thus, the chance performance would have been 10.8% (2.8 ÷ 26 letters). However, suppression effects will suffer from compression due to capacity limitations far before chance performance is reached. The rate of decrease in the maximum observable suppression effect will be influenced by several factors, including the number of letters reported and the accuracy at storing the target letter.

THE CURRENT STUDY

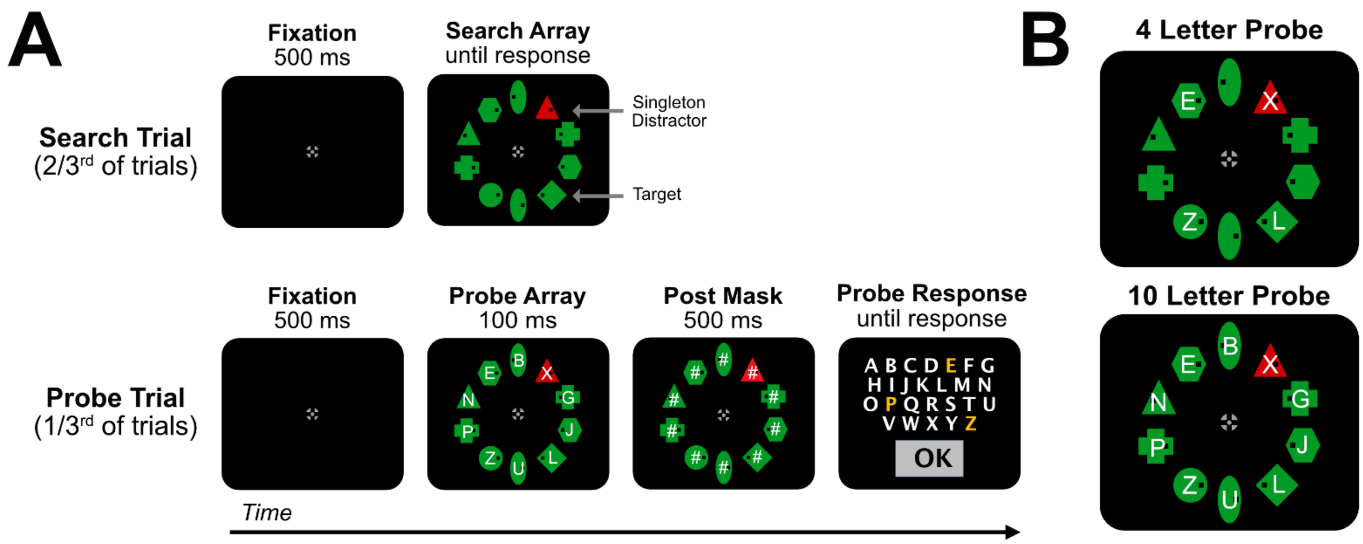
As reviewed, the capture-probe paradigm is commonly used to assess suppression of attentional capture by a salient distractor. However, the capture-probe paradigm is not well-suited for assessing suppression at high set sizes, which can yield floor effects on probe report making suppression of salient distractors below baseline levels difficult to observe. The current study will therefore introduce a new approach to the capture-probe paradigm that overcomes the floor effect problem. Participants will search displays with large set sizes (10 or 30 items) which should make the color singleton highly salient. Probe set size will be varied separately from overall search array set size, eliminating the confound of Wang and Theeuwes (2020). That is, we will use two different probe set sizes: four-letter arrays (as in Gaspelin et al., 2015) and ten-letter arrays (as in Wang & Theeuwes, 2020). Importantly, the four-letter array condition should increase overall probe recall accuracy and eliminate any potential floor effects in probe report.

EXPERIMENT 1

As depicted in Figure 2A, participants performed a capture-probe paradigm with search displays of a set size of 10. On *search trials*, participants searched for a target defined by shape and color (e.g., green diamond) and reported the location of a small dot (left vs. right) as quickly as possible. On half of trials, a color singleton distractor appeared at a nontarget location. On the other half of trials, there was no color singleton distractor. Importantly, on *probe trials,* the search array appeared briefly with letters superimposed on each search item for 100 ms. Next, the letters were replaced with a pattern mask for 500 ms. Participants then attempted to recall as many letters as possible. Crucially, we manipulated the number of letters in the probe display. On ten-letter probe trials, letter probes were superimposed on all 10 shapes (as in Wang & Theeuwes, 2020). On four-letter probe trials, however, letter probes were only superimposed onto 4 of the 10 shapes: the target, the singleton distractor (if present), and nonsingleton distractors. This should improve overall probe accuracy and eliminate any floor effects.

The signal suppression hypothesis and the stimulus-driven accounts make opposing predictions. Stimulus-driven accounts predict that the singleton should capture attention because it is now highly salient and cannot be suppressed (Wang & Theeuwes, 2020). This should yield higher recall for probe letters at the singleton distractor than the average nonsingleton distractor (a *probe capture effect*). This probe capture effect should occur at both probe set sizes because the singleton is equally salient in each condition. According to the signal suppression hypothesis, however, the singleton distractor should be proactively suppressed, resulting in lower recall for probe letters at the singleton distractor compared to the average nonsingleton distractor (a *probe suppression effect*). Additionally, probe suppression effects should be particularly large in the four-letter probe condition, which reduces any potential for floor effects.

Figure 2. *Stimuli and Procedure from Experiment 1.*



Note. (A) The trial procedure for the capture-probe paradigm. On search trials, participants searched an array of ten shapes for the target (e.g., diamond) and made a speeded button press indicating the location of a black dot inside (left vs. right). On probe trials, white letters were briefly superimposed on the search items and were then replaced with pattern masks (#). Participants attempted to report as many letters as possible. (B) The probe displays could have of two potential set sizes (4-letter vs. 10-letter), which was varied randomly by trial.

Method

Participants

A sample of 24 participants from State University of New York at Binghamton volunteered for course credit (21 women and 3 men; mean age = 18.8 years). The sample size was determined a priori based upon a power analysis of the probe suppression effect from Gaspelin et al. (2015, Exp. 4) which used similar methods and stimuli to the current experiment. Based upon the observed effect size (d_z = 0.94), 23 participants would be needed to obtain 99% power. All participants had normal color vision as indicated by an Ishihara test and had normal or corrected-to-normal visual acuity. All experimental protocols were approved by a university ethics board.

Apparatus

Stimuli were presented using PsychToolbox (Brainard, 1997) on an Asus VG245H LCD monitor (1920 x 1080 resolution) with a black background, placed at a viewing distance of 100 cm. The timing delay of the stimulus presentation system was measured using a photosensor and this value (12 ms) has been subtracted from all latency analyses in this paper. Viewing distance was maintained via a desk-mounted chinrest.

Stimuli & Procedure

As depicted in Figure 2A, each search display contained ten shapes arranged in a notional circle at an eccentricity of 3.1° from the center of the screen. The fixation cross was presented in the center of the screen and consisted of a gray (30.0 cd/m², x = .306, y = .320) circle (0.49° diameter) cross-sectioned by two black rectangles (each 0.49° by 0.12°), with a gray dot in the center (0.12° diameter). Each search display contained a diamond (1.3° by 1.3°), a circle (1.3° diameter), two triangles (1.3° by 1.3°), two hexagons (1.3° by 1.3°), two ovals (1.6° by 1.0°), and two crosses (1.3° by 1.3°). The shapes were either red (30.0 cd/m^2, x = .627, y = .330) or green (30.0 cd/m^2, x = .292, y = .631) which were photometrically isoluminant. The target shape (circle or diamond) and target color (green or red) were held constant for the experimental session, and all four potential combinations were counterbalanced across participants. On *singleton-absent trials* (half of trials), all ten shapes were the target color. On *singleton-present trials* (other half of trials), a randomly selected distractor item was rendered in a unique color, creating a singleton distractor. Each shape contained a small black dot (0.2° x 0.2°) on either the left or right side, selected at random. The location of the target shape and location of singleton distractor were selected at random with the exception that that the singleton distractor could not appear at the target location.

The distractor shapes were selected to eliminate any possible use of singleton detection mode. Namely, if only linear shapes were used as distractors, the circle target might "pop out" at high set sizes because it would be a curvilinear singleton and this could potentially induce singleton detection mode (Bacon & Egeth, 1994). Therefore, ellipse distractors were included to make the circle targets less salient and encourage use of feature search mode.

On *search trials* (two-thirds of trials), a fixation cross appeared for 500 ms. Next, the search array appeared, and participants attempted to locate the target shape as quickly as possible. They responded to the location of the dot inside the target shape (left or right) by clicking the left- or right-button on a computer mouse, respectively. Participants received immediate performance feedback after search trials. If the participant made an error, a 200 Hz tone sounded for 300 ms. If a participant took longer than 3000 ms to respond, a screen with the feedback "Too Slow" immediately appeared along with a 200 Hz tone for 300 ms. Participants were instructed that the singleton distractor would never be the target and should be ignored.

On *probe trials* (one-third of trials), the fixation screen appeared for 500 ms and was followed by the probe array for 100 ms. The probe array consisted of white capital letters in Arial typeface (0.7° by 0.7°) superimposed on the search item. Letters were randomly chosen from the full English alphabet without replacement. Following the probe array, all ten locations were post-masked with a white "#" for 500 ms. Then, a response prompt appeared that consisted of all twenty-six letters of the alphabet in white. Participants were allowed to click between zero and the number of letters presented which depended on the probe size condition. When a letter was selected, it turned yellow. Participants were allowed to unselect letters by clicking on the letter again, turning it back to white. Final probe answers were submitted by clicking a gray "OK" button located at the bottom of the screen. Probe responses were untimed and participants did not receive any feedback about the accuracy of probe reports.

Crucially, we manipulated the number of probe letters that presented randomly on each trial (see Figure 2B). In the four-letter probe condition, four letters appeared at random locations with the constraint that one letter always appeared on the target, one letter appeared on the singleton distractor (if present), and the remaining letters appeared at two randomly selected nonsingleton distractors. On singleton-absent trials, one letter appeared on the target and three letters appeared on randomly selected nonsingleton distractor shapes. This approach ensured an ample number of trials in which the singleton distractor and target letter were probed. On ten-letter probe trials, a letter appeared on every shape. In both conditions, participants were allowed to report between zero to the number of probe letters on the screen (e.g., zero to ten letters in the ten-letter probe condition). This manipulation is important because high probe set sizes (e.g., ten-letter probes) will reduce overall probe accuracy, creating a floor effect which makes suppression of the singleton distractor difficult to observe.

Participants first performed a practice block of 60 trials on the search task alone. This allowed participants to familiarize themselves with the search task generally. Then, participants performed a practice block of 60 trials of combined search and probe trials. Finally, the participants performed 8 regular blocks of 60 trials (480 trials total) with probe trials and search trials randomly intermixed. This yielded 160 probe trials, of which 80 were singleton-present trials and 80 were singleton-absent trials. Thus, there were 40 singleton-present trials using a four-letter probe array and 40 trials using a ten-letter probe array. Participants received block-by-block feedback on mean response time (RT) and accuracy.

Data Analysis

To avoid issues of sphericity, all reported *p* values in this manuscript were Greenhouse-Geisser corrected for analysis of variances (ANOVAs) with more than two levels of a given factor. When appropriate, we adjusted the partial-eta squared in our reports of effect size for positive bias (Mordkoff, 2019). For within-subject t-tests, Cohen's d_z was used to measure effect size (Lakens, 2013).

Results

Trials with an RT less than 200 ms or greater than 2500 ms (1.0% of trials) were excluded from all search-trial analyses. Additionally, trials with an incorrect response (1.7%) were excluded from search-trial RT analyses.

Search Trials

If attention was captured by the color singleton, mean RTs should be slower on singleton-present trials than singleton-absent trials. As depicted in Figure 3A, no singleton presence cost was observed. If anything, mean RTs were numerically faster on singleton-present trials (926 ms) than singleton-absent trials (936 ms). A preplanned t-test indicated that this effect was not statistically significant, $t(23) = 1.16$, $p = .256$, $d_z = 0.24$. Mean error rates also did not differ significantly on singleton-present trials (1.2%) and singleton-absent trials (1.3%), $t(23) = 0.30$, $p = .769$, $d_z = 0.06$. Altogether, there was no evidence of attentional capture in either mean RT or manual error rates. This pattern of results is inconsistent with the stimulus-driven accounts and replicates the well-established pattern that inducing feature search mode can eliminate singleton presence costs (Bacon & Egeth, 1994; Leber & Egeth, 2006b).

Probe Trials

Table 1 provides a complete summary of probe results. Participants reported an average of 1.6 letters and 62.3% of these letters were actually present. This is similar to previous studies using the capture-probe paradigm (Gaspelin et al., 2015; Gaspelin & Luck, 2018a, 2018b).

Singleton Suppression Effects

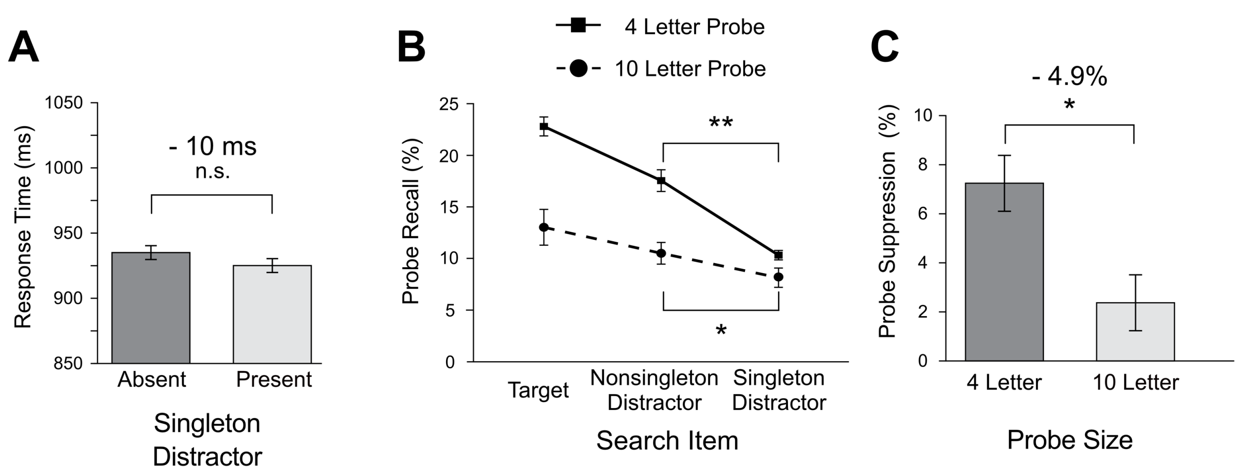
The key question in this study is whether the highly salient color singleton would be suppressed or would capture attention. If the singleton captures attention, participants should be *more* likely to report the singleton distractor than the average nonsingleton distractor. If the singleton is suppressed, participants should be *less* likely to report the singleton distractor than the average nonsingleton distractor. As can be seen in Figure 3B, the singleton distractor was less likely to be reported than the average nonsingleton distractor at both probe sizes (Gaspelin et al., 2015; Gaspelin & Luck, 2018a, 2018b). To further demonstrate this, *probe suppression effects* were computed as difference scores that subtract probe recall accuracy of the singleton distractor from that of the average nonsingleton distractor for each probe size condition (Figure 3C). A positive value indicates that participants were less likely to recall the singleton distractor than the average nonsingleton distractor (i.e., suppression), whereas a negative value indicates that participants were more likely to recall the singleton distractor than the average nonsingleton distractor (i.e., capture). As can be seen, there was no evidence of capture and instead the singleton was suppressed.

Table 1. Mean Probe Recall for all Conditions and Experiments.

		Singleton Absent		Singleton Present		
	Probe Size	Target	Nonsingleton distractor	Target	Nonsingleton distractor	Singleton distractor
Exp. 1	4 Letter	24.2% (2.5%)	18.0% (1.7%)	22.8% (2.5%)	17.6% (1.6%)	10.3% (1.4%)
	10 Letter	12.8% (1.1%)	10.4% (0.8%)	13.0% (1.5%)	10.5% (0.8%)	8.1% (1.0%)
Exp. 2	4 Letter	16.4% (2.2%)	13.3% (1.4%)	14.6% (2.0%)	14.1% (2.1%)	8.8% (1.6%)
	10 Letter	10.4% (1.5%)	7.9% (1.0%)	10.0% (2.1%)	7.6% (1.0%)	6.7% (1.3%)
Exp. 3	4 Letter	20.0% (2.7%)	17.7% (2.0%)	23.4% (3.2%)	19.0% (2.4%)	11.9% (1.9%)
	10 Letter	11.6% (1.6%)	8.6% (1.1%)	12.5% (1.7%)	9.5% (1.2%)	5.0% (1.0%)
Exp. 4	4 Letter	20.3% (1.2%)	18.8% (1.3%)	18.7% (1.7%)	18.1% (1.3%)	14.4% (1.4%)
	10 Letter	11.3% (1.0%)	10.3% (0.6%)	12.7% (1.1%)	10.2% (0.7%)	10.2% (1.2%)

Note. Standard errors of the mean are provided in parentheses. Figure 3B depicts probe recall accuracy on singleton-present trials for both probe size conditions (four-letter probe and ten-letter probe).

Figure 3. *Results for Experiment 1.*



Note. (A) Mean response time (RT) on search trials as a function of singleton presence. (B) Probe recall accuracy on singleton-present trials as a function of probe display size. (C) Probe suppression effects as a function of probe display size. All error bars in all figures represent within-subject standard error of the mean (Cousineau, 2005; Morey, 2008). ** $p < .001$. * $p < .05$.

Preplanned t-tests were conducted to assess probe suppression effects across probe size conditions. Probe suppression effects were larger on four-letter probe trials than ten-letter probe trials, $t(23) = 3.10$, $p = .005$, $d_z = 0.63$. One-sample t-tests then compared each probe suppression effect against zero. This revealed a significant probe suppression effect in both the four-letter probe condition (7.4%), $t(23) = 4.57$, $p < .001$, $d_z = 0.93$, and the ten-letter probe condition (2.5%), $t(23) = 2.62$, $p = .015$, $d_z = 0.54$. Altogether, the probe results are quite clear. There was no evidence that the singleton reliably captured attention. Instead, probe suppression effects were apparent in both probe size conditions. Importantly, this probe suppression effect was larger in the four-letter probe condition than the ten-letter probe condition, which is consistent with a floor effect.

Search Item Analysis

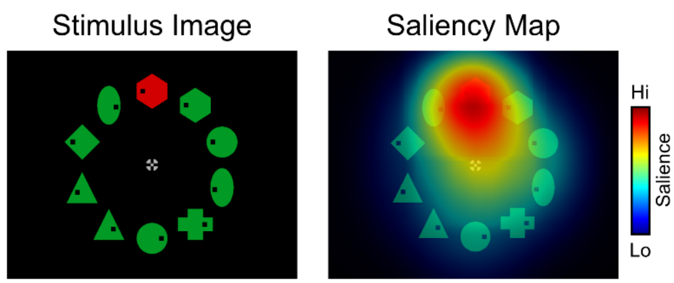
We also performed an analysis comparing probe performance for each search item (target, nonsingleton distractor, and singleton distractor) between both probe size conditions. To be clear, neither stimulus-driven nor signal suppression accounts make any predictions about how probe set size will influence target enhancement and these analyses were included for the sake of completeness. First, we conducted a two-way repeated-measures ANOVA with factors of search item (target, nonsingleton, and singleton distractor) and probe size (four-letter and ten-letter). There was a main effect of search item, $F(2, 46) = 22.09$, $p < .001$, $\eta_p^2 = .468$, indicating that probe recall accuracy was higher for targets (18.2%) and nonsingleton distractors (14.2%) than singleton distractors (9.4%). There was also a main effect of probe size, $F(1, 23) = 58.92$, $p < .001$, $\eta_p^2 = .707$, indicating that probe recall accuracy was higher on four-letter probe trials (17.1%) than the ten-letter probe trials (10.7%). Critically, we observed a significant interaction between search item and probe size, $F(2, 46) = 9.19$, $p < .001$, $\eta_p^2 = .254$.

Paired-sample t-tests compared each search item (target, nonsingleton distractor, and singleton distractor) across probe size conditions (four-letter vs. ten-letter). Probe recall accuracy for the nonsingleton distractor was much higher on four-letter probe trials (17.7%) than ten-letter probe trials (10.6%), $t(23) = 6.63$, $p < .001$, $d_z = 1.35$. Similarly, probe recall accuracy for the target was also higher on four-letter probe trials (23.0%) than ten-letter probe trials (13.3%), $t(23) = 5.61$, $p < .001$, $d_z = 1.15$. Finally, probe recall accuracy for the singleton distractor was slightly higher for four-letter probe trials (10.5%) than ten-letter probe trials (8.3%), $t(23) = 2.21$, $p = .037$, $d_z = 0.45$. In summary, these results indicate that probe accuracy was generally higher in the four-letter probe condition than ten-letter probe condition.

Saliency Maps

An important question is whether the color singleton was truly salient. To examine this, saliency maps of the search stimuli were generated and this allowed us to directly estimate the salience of the singleton (see also Chang et al., in press). The Image Signature Toolbox (Hou et al., 2012) was chosen because it performs similar to human observers in its ability to detect popout stimuli in artificial images, unlike many other saliency models (Kotseruba et al., 2020). Sample images of search arrays from Experiment 1 were generated (1000 images total; 250 for each counterbalancing condition) and then the images were processed using the Image Signature Toolbox in Matlab. The default settings were used, with the exception of the mapWidth() parameter, which was adjusted to accommodate the image resolution (1920 x 1080). The output was a series of saliency maps that resemble heatmaps depicting the relative distribution of salience across the image (Figure 4). In each saliency map, circular regions of interest (1.8° in diameter) were defined around each search item. Within each region of interest, a mean salience score was calculated by averaging the pixels in the heat map.

Figure 4. *Using Saliency Maps to Estimate Singleton Salience.*



Note. Stimulus images (*N* = 1,000) were analyzed using the Image Signature Toolbox to determine if the singleton was truly salient. In this example, the saliency map indicated a high salience score at the location of the color singleton.

The salience scores at each item were then used to develop two metrics of salience. The first metric was the *global saliency index* (GSI), which was computed as the difference between the mean salience score at the singleton distractor minus the mean salience score at the average of all nonsingleton distractor locations (Kotseruba et al., 2020). The difference score was then divided by the sum of the salience scores of all search items. Thus, the GSI is a normalized and continuous metric ranging from -1 to 1. A positive value indicates that the singleton is more salient than the other items, whereas a negative value indicates that the singleton is less salient than the other items. A zero would indicate the singleton is equally salient as other items. The average GSI from all of the images was 0.68 (SE = 0.004), indicating the singleton had a salience score that was higher than the other search items.

The second metric to assess salience was the *singleton win rate*, which is the percentage of images in which the singleton had the highest salience score of all items in the saliency map. In Experiment 1, the singleton win rate was 100%, indicating that the singleton had the highest saliency score of all items in every single heat map. This again confirms that the singleton was salient in Experiment 1.

Discussion

Experiment 1 assessed whether a highly salient color singleton could capture attention. The salience of the color singleton was enhanced by increasing the overall set size of the search displays to 10 items. Yet, no evidence of capture was observed in either mean RTs on search trials or letter recall on probe trials. Instead, we observed clear evidence of a probe suppression effect: Participants were less likely to recall the letter at the singleton distractor location than the average nonsingleton distractor location. These data are consistent with the signal suppression hypothesis, which suggests that singletons are proactively suppressed to prevent attentional capture.

Importantly, we manipulated the number of probe letters, which allowed us to assess potential floor effects in probe recall accuracy. Overall recall accuracy on the probe task was lower in the ten-letter probe condition (10.5%) than the four-letter probe condition (16.9%). Probe suppression effects were also lower in the ten-letter probe condition (2.4%) than the four-letter probe condition (7.2%). Importantly, this pattern of results cannot be explained by differences in physical saliency of the singleton, which should have been equally salient in both probe size conditions because the same search displays were used.

EXPERIMENT 2

In Experiment 1, a highly salient singleton distractor did not produce any evidence of attentional capture. Instead, the singleton was proactively suppressed below baseline levels, despite using the same search array set size of 10 as Wang and Theeuwes (2020). It is possible, however, that the color singleton was not salient enough to overpower suppression, even at a set size of 10. Therefore, Experiment 2 made the singleton even more salient by tripling the number of search items (Figure 5A). The methods were identical to Experiment 1, except that 20 additional nonsingleton distractors appeared in an annulus just outside of the actual search array (see Theeuwes, 2004). Thus, the singleton distractor now appeared amongst 29 homogenously colored items, meaning that it should be even more salient than in Experiment 1

The key predictions are identical to Experiment 1. According to the stimulus-driven account, the singleton distractor should automatically capture attention because it is highly salient and this should result in a probe capture effect. According to the signal suppression hypothesis, however, the color singleton should be suppressed, resulting in probe suppression effects similar to those in Experiment 1. Additionally, probe suppression effects should be greater in the four-letter probe condition than the ten-letter probe condition, due to floor effects in probe recall accuracy.

Method

All methods were identical to Experiment 1, except as follows. First, a new sample of 24 students from State University of New York at Binghamton volunteered for course credit (14 females and 10 males; mean age = 18.7 years).

Second, the set size of the search displays was increased to 30 items (Figure 5A). Each search array consisted of an inner ring of 10 shapes (similar to Experiment 1) and an outer ring of 20 distractor shapes. The new outer ring appeared 5.9° from fixation. The center-to-center distance of shapes in the inner ring and outer ring was 2.4°. The target and singleton distractor could not appear in the outer ring. Probe letters also could not appear in the outer ring. Thus, the outer ring consisted of "dummy" nonsingleton distractors meant to boost the salience of the singleton distractor. This approach of only presenting search stimuli in the inner ring ensures that all search items and probe letters were equidistant from central fixation. Thus, any observed differences in search performance or probe recall accuracy could not be due to differences in the physical distance of search items from fixation (i.e., there were no trials where the target appeared closer to fixation than the singleton distractor or vice versa). To encourage feature search mode, the outer ring consisted of heterogenous shapes: four triangles, four hexagons, four ovals, four crosses, and four diamonds/circles (whichever was not the target shape for the participant). The shapes were the same size and colors as Experiment 1.

Results

Trials with an RT less than 200 ms or greater than 2500 ms (1.6% of trials) were excluded from all search-trial analyses. Additionally, trials with an incorrect response (2.0%) were excluded from search-trial RT analyses.

Search Trials

If the singleton distractor captures attention, mean RTs should be slower on singleton-present trials than singleton-absent trials. As depicted in Figure 5B, no such pattern was observed. Mean RTs were nearly identical on singleton-present trials (1011 ms) and singleton-absent trials (1009 ms), $t(23) = 0.26$, $p = .799$, $d_z = 0.05$. Mean error rates also did not differ significantly between singleton-present trials (1.5%) and singleton-absent trials (1.2%), $t(23) = 1.61$, $p = .122$, $d_z = 0.33$. Altogether, the results from search trials indicate no evidence of attentional capture, replicating Experiment 1.

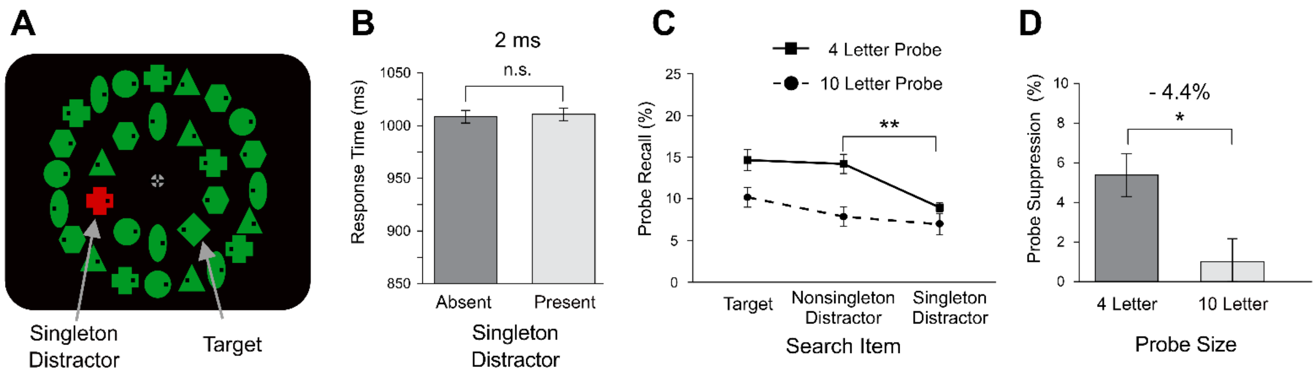
Probe Trials

Figure 5C depicts probe recall accuracy on singleton-present trials in the four-letter probe and ten-letter probe conditions. Table 1 depicts a complete summary of probe results. Participants reported an average of 1.2 letters and 60.5% of these letters were actually present.

Singleton Suppression Effects

In the four-letter probe condition, the singleton distractor was less likely to be reported than the average nonsingleton distractor: a probe suppression effect. In the ten-letter probe condition, no probe suppression effect was

Figure 5. *Stimuli and Results for Experiment 2.*



Note. (A) A search display for Experiment 2. (B) Mean response time (RT) on search trials as a function singleton presence. (C) Probe recall accuracy on singleton-present trials as a function of probe display size. (D) Probe suppression effects as a function of probe display size. ** $p < .001$. * $p < .05$.

observed. We again computed probe suppression effects, which are a difference score of probe recall accuracy at the singleton distractor location and the average nonsingleton distractor location (Figure 5D). One-sample t-tests compared each probe suppression effect against zero. This revealed a significant probe suppression effect in the four-letter probe condition (5.4%), $t(23) = 3.13$, $p = .005$, $d_z = 0.64$, but not the ten-letter probe condition (1.0%), $t(23) = 1.06$, $p = .299$, $d_z = 0.22$. A preplanned paired-samples t-test demonstrated that the probe suppression effect was significantly larger in the four-letter probe condition than the ten-letter probe condition, $t(23) = 2.68$, $p = .013$, $d_z = 0.55$. These results replicate the basic pattern of results from Experiment 1. There was no evidence that the highly salient color singleton captured attention. Importantly, the probe suppression effect was larger in the four-letter probe condition than the ten-letter probe condition, which is consistent with a floor effect on recall accuracy.

Search Item Analysis

We also performed an additional analysis comparing probe performance on singleton-present trials. As in Experiment 1, we had no a priori hypotheses about how probe set size would influence target report. First, we conducted a two-way repeated-measures ANOVA with factors of search item (target, nonsingleton, vs. singleton distractor) and probe size (four-letter vs. ten-letter). There was a main effect of search item, $F(2, 46) = 6.71$, $p = .003$, $\eta_p^2 = .192$, indicating that probe recall accuracy was higher for targets (12.3%) and nonsingleton distractors (10.9%) than singleton distractors (7.7%). There was also a main effect of probe size, $F(1, 23) = 30.55$, $p < .001$, $\eta_p^2 = .552$, indicating that probe recall accuracy was higher on four-letter probe trials (12.5%) than the ten-letter probe trials (8.1%). We observed a trend toward an interaction between search item and probe size, $F(2, 46) = 2.68$, $p = .079$, $\eta_p^2 = .065$.

Paired-samples t-tests compared each search item (target, nonsingleton distractor, and singleton distractor) at each probe size (four-letter vs. ten-letter). Probe recall accuracy for the nonsingleton distractor was much higher on four-letter probe trials (14.1%) than ten-letter probe trials (7.6%), $t(23) = 4.70$, $p < .001$, $d_z = 0.96$. Similarly, probe recall accuracy for the target was also higher on four-letter probe trials (14.6%) than ten-letter probe trials (10.0%), $t(23) = 3.37$, $p = .003$, $d_z = 0.69$. Finally, probe recall accuracy for the singleton distractor did not differ significantly between four-letter probe trials (8.7%) and ten-letter probe trials (6.7%), $t(23) = 1.56$, $p = .132$, $d_z = 0.32$.

Saliency Maps

The salience of the color singleton was verified in the same manner as Experiment 1. We generated 1000 stimulus images and analyzed them using the Image Signature Toolbox. The mean GSI across the stimulus images was 0.83 (SE = 0.002), indicating that the color singleton was highly salient. The GSI was significantly higher in Experiment 2 (0.83) than in Experiment 1 (0.68), $t(1998) = 37.12$, $p < .001$, $d_z = 1.66$, which suggests that increasing the set size from 10 to 30 items did actually improve the salience of the color singleton. The singleton win rate was 100%, indicating that the singleton had the highest salience score in all images. Altogether, the results suggest that the singleton was salient.

Discussion

In Experiment 2, the salience of the singleton distractor was increased by including 20 additional nonsingleton distractor items at dummy locations. This should have massively boosted the salience of the color singleton. Nonetheless, there was still no evidence of capture. There was no singleton presence cost on mean RT on search trials. There was also no probe capture effect on probe trials. Instead, there was a probe suppression effect: the singleton distractor was less likely to be reported than the average nonsingleton distractor. This is consistent with the signal suppression hypothesis, but inconsistent with stimulus-driven accounts.

As in Experiment 1, we independently manipulated probe size (4 vs. 10) while holding the number of items in the search display constant. Importantly, probe suppression effects were smaller in the ten-letter probe condition (0.9%) than the four-letter probe condition (5.3%), even though the singleton was equally salient in both conditions. This result is consistent with a floor effect: low overall probe accuracy makes it difficult to observe underlying suppression effects.

EXPERIMENT 3

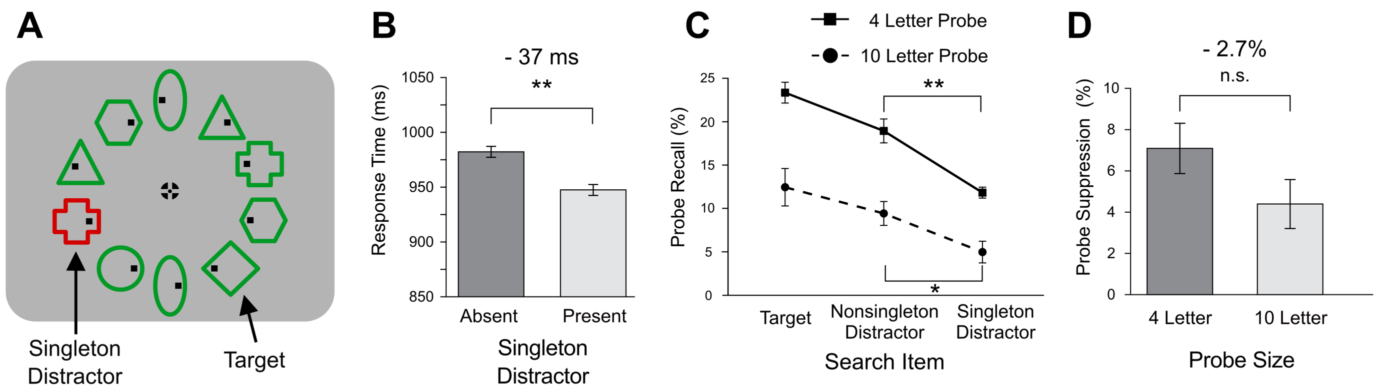
In Experiments 1 and 2, there was no evidence that the color singleton was able to capture attention. The data instead suggested that the singleton distractor was proactively suppressed below baseline levels. This pattern of results differs from Wang and Theeuwes (2020), who reported a 1.9% probe capture effect from the singleton distractor. A notable difference is that the current study used filled shapes on a black background (similar to Gaspelin et al., 2015), whereas Wang and Theeuwes (2020) used unfilled shapes on a gray background. It is conceivable that unfilled shapes somehow give the color singleton a competitive advantage. For example, perhaps by weakening top-down guidance to the target, the color singleton can more strongly compete for attentional allocation.

Experiment 3 modified the stimuli to be similar to Wang and Theeuwes (2020). The stimuli were the same as Experiment 1, except that the shapes were changed to unfilled outlines that appeared on a gray background (Figure 6A). The key predictions are identical to Experiments 1 and 2. According to a stimulus-driven account, the singleton should automatically capture attention, yielding a probe capture effect. According to the signal suppression hypothesis, the singleton should be proactively suppressed, yielding a probe suppression effect. Suppression effects should be particularly strong at the low probe set size.

Method

All methods were identical to Experiment 1, except for the following. First, a new sample of 24 students from State University of New York at Binghamton volunteered for course credit (14 females and 10 males; mean age = 20.8 years). Second, the stimuli were adapted to be similar to those used by Wang and Theeuwes (2020). As shown in Figure 6A, the search shapes were changed to unfilled shapes (0.2° in thickness) instead of filled shapes. Also, shapes now appeared on a gray background that was photometrically isoluminant with the colors of the

Figure 6. *Stimuli and Results for Experiment 3.*



Note. (A) A search display for Experiment 3. (B) Mean response time (RT) on search trials as a function singleton presence. (C) Probe recall accuracy on singleton-present trials as a function of probe display size. (D) Probe suppression effects as a function of probe display size. ** $p < .001$. * $p < .05$.

shapes (30.0 cd/m2). The fixation cross (which was previously gray) was changed to black.

Results

Trials with an RT less than 200 ms or greater than 2500 ms (1.3% of trials) were excluded from all search-trial analyses. Additionally, trials with an incorrect response (2.6%) were excluded from search-trial RT analyses.

Search Trials

As depicted in Figure 6B, mean RT was faster when the singleton distractor was present than absent: a singleton presence benefit. This pattern of results has previously been taken to suggest that the singleton was suppressed, speeding search when it is present (Chang & Egeth, 2019; Gaspelin et al., 2015). A preplanned t-test indicated that mean RTs were significantly faster on singleton-present trials (946 ms) than singleton-absent trials (984 ms), $t(23) = 4.77$, $p < .001$, $dz = 0.97$: a 38-ms singleton presence benefit. Mean error rates did not differ significantly on singleton-present trials (2.1%) and singleton-absent trials (1.9%), $t(23) = 0.54$, $p = .597$, $dz = 0.11$. Altogether, the results on search trials replicate Experiments 1 and 2. There was no evidence of capture in either mean RT or mean error rates.

Probe Trials

Figure 6C depicts probe recall accuracy on singleton-present trials for the four-letter probe and ten-letter probe conditions. Probe suppression effects are depicted in Figure 6D. A complete summary of probe results is depicted in Table 1. Participants reported an average of 1.2 letters per trial and 71.2% of these letters were actually present.

Singleton Suppression Effects

As shown in Figure 6C, in both probe size conditions, the singleton distractor was less likely to be reported than the average nonsingleton distractor. As in Experiments 1 and 2, preplanned one-sample t-tests compared these probe suppression effects against zero. This revealed a significant probe suppression effect in both the four-letter probe condition (7.1%), $t(23) = 4.28$, $p < .001$, $dz = 0.87$, and the ten-letter probe condition (4.5%), $t(23) = 3.55$, $p = .002$, $dz = 0.73$. Although probe suppression effects were numerically larger in the four-letter probe condition than the ten-letter probe condition, a preplanned paired-samples t-test demonstrated this effect was not statistically significant, $t(23) = 1.64$, $p = .115$, $d_z = 0.34$. Consistent with Experiments 1 and 2, there was no evidence that the singleton captured attention. Instead, we observed probe suppression effects indicating that the singleton was suppressed below the baseline level of the nonsingleton distractors and did not capture attention.

Search Item Analysis

An additional analysis compared probe performance on singleton-present trials. First, we conducted a two-way repeated-measures ANOVA with factors of search item (target, nonsingleton, vs. singleton distractor) and probe size (four-letter vs. ten-letter). There was a main effect of search item, $F(2, 46) = 20.24$, $p < .001$, $\eta_p^2 = .446$, indicating that probe recall accuracy was higher for targets (18.2%) and nonsingleton distractors (14.3%) than singleton distractors (8.7%). There was also a main effect of probe size, $F(1, 23) = 32.13$, $p < .001$, $\eta_p^2 = .565$, indicating that probe recall accuracy was higher on four-letter probe trials (18.3%) than the ten-letter probe trials (9.1%). There was a nonsignificant trend toward an interaction between search item and probe size, $F(2, 46) = 2.43$, $p = .099$, $\eta_p^2 = .056$.

Paired-samples t-tests compared each search item (target, nonsingleton distractor, and singleton distractor) across probe size conditions (four-letter vs. ten-letter), consistent with Experiments 1 and 2. Probe recall accuracy for the nonsingleton distractor was much higher on four-letter probe trials (19.0%) than ten-letter probe trials (9.5%), $t(23) = 6.37$, $p < .001$, $d_z = 1.30$. Similarly, probe recall accuracy for the target was also higher on four-letter probe trials (23.4%) than ten-letter probe trials (12.5%), $t(23) = 4.53$, $p < .001$, $d_z = 0.92$. Finally, probe recall accuracy for the singleton distractor was higher for four-letter probe trials (11.9%) than ten-letter probe trials (5.0%), $t(23) = 4.01$, $p < .001$, $d_z = 0.82$.

These results are consistent with Experiments 1 and 2. Probe recall accuracy was higher for four-letter than ten-letter probes. This pattern of probe suppression across probe sizes is consistent with the notion that increasing the number of nonsingleton distractors leads to difficulty assessing probe suppression at higher probe sizes (i.e., 10-letter probes).

Saliency Maps

The salience of the color singleton was assessed using the Image Signature Toolbox (Hou et al., 2012). The GSI was 0.17 (SE = 0.008) and the singleton win rate was 34%, indicating that the color singleton was moderately salient. We also compared saliency maps for Experiments 1 and 3, which were identical except for the use of filled shapes (Exp. 1) versus unfilled shapes (Exp. 3). Interestingly, the GSI in Experiment 3 (0.17) was much smaller than in Experiment 1 (0.68), $t(1998) = 59.96$, $p < .001$, $d_z = 2.68$. Additionally, the singleton win rate in Experiment 3 (34%) was also lower than in Experiment 1 (100%). Thus, by more closely aligning our stimuli with those used by Wang and Theeuwes (2020), we actually decreased the salience of the singleton distractor. In other words, using unfilled shapes on a gray background seemed to reduce the contrast of color singleton with the background and other search items.

Discussion

Experiment 3 used stimuli that were more similar to Wang and Theeuwes (2020) than Experiments 1 and 2. Search stimuli were unfilled outlines of shapes that appeared on a gray background. However, we found no evidence of attentional capture by the highly salient color singleton. On search trials, mean RTs were actually faster when the singleton was present than when it was absent (a singleton presence benefit). Critically, there were probe suppression effects in both probe size conditions. Altogether, the current results are consistent with the signal suppression account in that a highly salient color singleton failed to capture attention.

EXPERIMENT 4

In Experiments 1–3, there was no evidence that the color singleton captured attention. This pattern of results differs from Wang and Theeuwes (2020), who reported a 1.9% probe capture effect from the singleton distractor at a set size of 10. Experiment 4 was therefore a direct replication that used the same search stimuli as Wang and Theeuwes (2020). That is, the search arrays were adapted to use the same shapes, stimulus colors, and display sizes. The key predictions are identical to Experiments 1–3. According to a stimulus-driven account, the singleton should capture attention and lead to a probe capture effect. According to the signal suppression hypothesis, the singleton should be proactively suppressed and lead to a probe suppression effect. This probe suppression effect should be larger in the four-letter probe condition, where potential floor effects are minimized.

Method

All methods were similar to Experiment 1, except as follows. First, a new sample of 40 participants from State University of New York at Binghamton volunteered for course credit (27 females and 13 males; mean age = 19.7 years). This sample size was chosen based upon a post hoc power analysis of the probe capture effect (d_z = 0.51) reported by Wang and Theeuwes (2020). The observed power at their sample size of 24 was 0.67, suggesting that the study was slightly underpowered. A sample size of 37 participants would be needed to achieve at least 85% power.[1]

Second, the search stimuli were adapted to be identical to Wang and Theeuwes (2020). As shown in Figure 7A, shapes were either red (RGB: 255, 0 , 0; 63.4 cd/m^2, *x* = .641, *y* = .331) or green (RGB: 0, 255, 0; 226.0 cd/m^2, *x* = .304, *y* = .639,) on a gray background (RGB: 128, 128, 128; 58.3 cd/m^2, *x* = .302, *y* = .322). These colors were not photometrically isoluminant. Each search display contained four squares (1.6° x 1.6°), four hexagons (1.6° x 1.6°), one diamond (1.6° x 1.6°), and one circle (radius of 0.7°). The shapes were arranged equidistant from the fixation point on an imaginary circle with an eccentricity of 3.0°. All shapes were unfilled, with an outline thickness of 0.1°. Additionally, the postmasks (#) after the probe display were removed from the trial progression. Thus, after the probe display appeared, a probe response prompt appeared 100 ms later.[2]

Results

Trials with an RT less than 200 ms or greater than 2500 ms (0.4% of trials) were excluded from all search-trial analyses. Additionally, trials with an incorrect response (1.9%) were excluded from search-trial RT analyses.

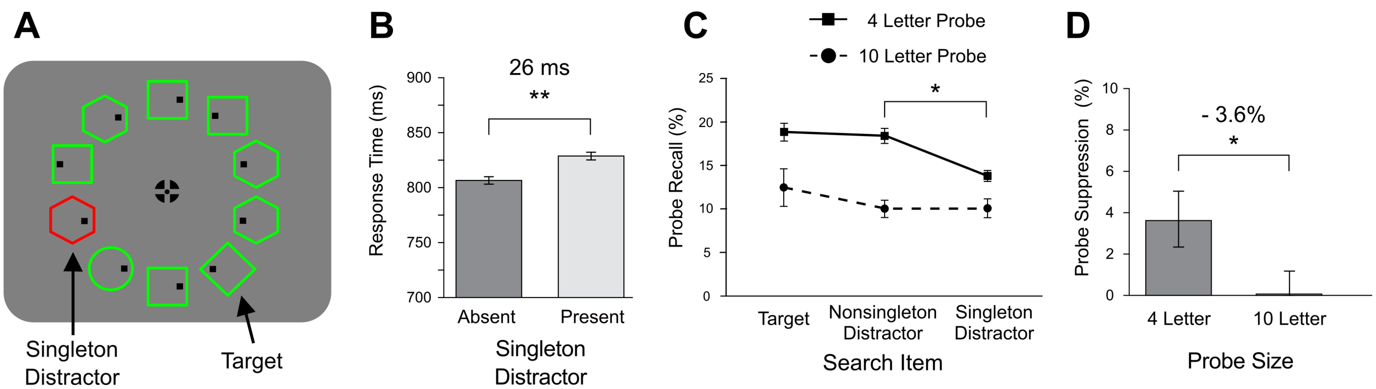
Search Trials

As depicted in Figure 7B, mean RT was slower when the singleton distractor was present than absent: a singleton presence cost. A preplanned within-subject t-test indicated that mean RTs were significantly slower on singleton-present trials (830 ms) than singleton-absent trials (804 ms), $t(39) = 4.64$, $p < .001$, dz = 0.73. Additionally, error rates were slightly higher on singleton-present trials (2.0%) than singleton-absent trials (1.6%), $t(39) = 2.32$, $p = .026$, dz = 0.37. The 26-ms singleton presence cost replicates Wang and Theeuwes (2020). However, singleton presence costs are not, by themselves, a definitive indication of attentional capture. The

[1] The issue of low power does not apply to Experiments 1–3. For example, the probe suppression effect in the four-letter probe condition of Experiment 1 (d_z = 0.93) would yield 85% power at a sample size of 13 participants.

[2] After completing the experiment, we contacted the authors who kindly provided us with screenshots of their stimuli. The only visible difference between our stimuli and their stimuli was the beginning position of the notional circle array of shapes. Our search locations were generated beginning from 0° angular position (top position). Their search stimuli began at a position slightly offset from the top position.

Figure 7. *Stimuli and Results for Experiment 4.*



Note. (A) A search display for Experiment 4. (B) Mean response time (RT) on search trials as a function singleton presence. (C) Probe recall accuracy on singleton-present trials as a function of probe display size. (D) Probe suppression effects as a function of probe display size. ** $p < .001$. * $p < .05$.

presence of the singleton could slow an attentional shift to the target without an actual attentional shift to the singleton itself. These filtering costs have been observed in previous studies of attentional capture (Becker, 2007; Folk & Remington, 1998). Therefore, the following section will assess attentional allocation to the singleton item using probe recall accuracy.

Probe Trials

Figure 7C depicts probe recall accuracy on singleton-present trials for the four-letter probe and ten-letter probe conditions. Figure 7D depicts probe suppression effects for each probe size condition. Table 1 provides a complete summary of probe results. Participants reported an average of 1.6 letters per trial and 65.4% of these letters were actually present.

Singleton Suppression Effects

As shown in Figure 7C, neither condition produced a probe capture effect. For ten-letter probe trials, there was a 0.0% probe capture effect, t(39) = 0.01, p = .990, dz = 0.00. Thus, we did not replicate the probe capture effect reported by Wang and Theeuwes (2020). For four-letter probe trials, there was a 3.6% probe suppression effect (Gaspelin et al., 2015; Gaspelin & Luck, 2018a, 2018b), t(39) = 2.18, p = .036, dz = 0.34. A preplanned paired-sample t-test indicated that the probe suppression effect in the four-letter probe condition was larger than the ten-letter probe condition, t(39) = 2.11, p = .042, dz = 0.33. Altogether, the results were consistent with Experiments 1–3. There was no evidence that the singleton captured attention in probe recall accuracy. Probe suppression effects were enhanced in the four-letter probe condition compared to the ten-letter probe condition, consistent with a floor effect.

Search Item Analysis

We also performed an analysis comparing probe performance on singleton-present trials. First, we conducted a two-way repeated-measures ANOVA with factors of search item (target, nonsingleton, vs. singleton distractor) and probe size (four-letter vs. ten-letter). There was a main effect of search item, $F(2, 78) = 4.37$, $p = .016$, $\eta_p^2 = .078$, indicating that probe recall accuracy was higher for targets (15.7%) and nonsingleton distractors (14.1%) than singleton distractors (12.3%). There was also a main effect of probe size, $F(1, 39) = 47.17$, $p < .001$, $\eta_p^2 = .536$, indicating that probe recall accuracy was higher on four-letter probe trials (17.1%) than the ten-letter probe trials (11.0%). There was not a significant interaction between search item and probe size, $F(2, 78) = 2.06$, $p = .135$, $\eta_p^2 = .026$.

Preplanned paired-samples t-tests compared each search item (target, nonsingleton distractor, and singleton distractor) across probe size (four-letter vs. ten-letter) as in Experiments 1–3. Probe recall accuracy for the nonsingleton distractor was much higher on four-letter probe trials (18.1%) than ten-letter probe trials (10.2%), $t(39) = 7.14$, $p < .001$, dz = 1.13. Similarly, probe recall accuracy for the target was also higher on four-letter probe trials (18.7%) than ten-letter probe trials (12.7%), $t(39) = 3.87$, $p < .001$, dz = 0.61. Finally, probe recall accuracy for the singleton distractor was higher for four-letter probe trials (14.4%) than ten-letter probe trials (10.2%), $t(39) = 3.07$, $p = .004$, dz = 0.49.

Saliency Maps

The salience of the color singleton was examined using the Image Signature Toolbox (Hou et al., 2012). The GSI was -0.12 (SE = 0.010). This negative value indicates that the singleton distractor had an average saliency score that was lower than other search items. Additionally, the singleton win rate was 6% in the saliency maps, which is slightly less than chance (10%). Thus, the Image Signature Toolbox seemed to indicate that the color singleton was less salient than the average search item.

Further inspection of the saliency maps indicated that the circle item received a high salience score. On every trial, one item was a circle in Wang and Theeuwes (2020). If the GSI is recalculated for the circle (instead of the color singleton), the GSI was 0.64 (SE = 0.003) and the win rate was 100%. This pattern suggests that the circle item, which was a curvilinear singleton, might have been highly salient. This issue does not apply to the stimuli from Experiments 1–3, which used oval distractors to prevent the circle from popping out and maximized shape heterogeneity amongst nontargets.

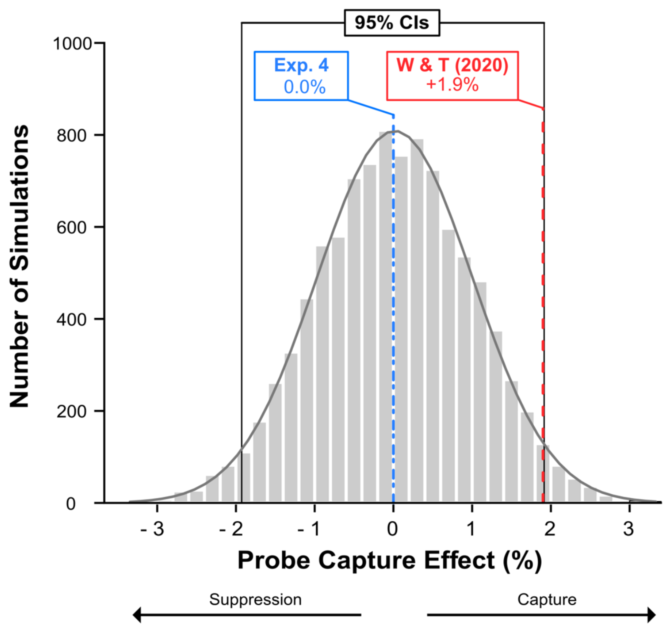
Bootstrapping Simulation

Experiment 4 failed to replicate the probe capture effect observed by Wang and Theeuwes (2020). We therefore conducted a bootstrapping analysis (Efron & Tibshirani, 1993) to assess the relative likelihood of obtaining a probe capture effect from the data of Experiment 4. For each simulation, a random subset of 24 subjects, matching the sample size of Wang and Theeuwes (2020), was selected from the full sample of 40 participants in Experiment 4. This random subsampling was repeated 10,000 times. For each subsampling, a mean probe capture effect was calculated in the ten-letter probe condition. As shown in Figure 8, the mean of the resulting distribution was 0.0%, which is the probe capture effect observed in Experiment 4. Interestingly, the 1.9% probe capture effect observed by Wang and Theeuwes (2020) falls just within the two-tailed 95% confidence interval. In other words, approximately 2.6% of the simulations resulted in a probe capture effect equal to or greater than that observed by Wang and Theeuwes (2020). Thus, although relatively rare, it is possible that to obtain the Type I error of a small probe capture effect when the underlying population distribution is truly null.

Discussion

Experiment 4 used the same search stimuli as Wang and Theeuwes (2020). However, it failed to replicate the key finding from this study: there was a 0.0% probe capture effect in the 10-letter probe condition. A bootstrapping analysis showed that it is possible to obtain a Type I error of a probe capture effect at lower sample sizes. Ultimately, future research may be needed to definitively ascertain why conflicting results were found. In any case, it is important to highlight that Experiment 4 did replicate the same floor effect pattern as Experiments 1–3 with smaller probe suppression effects at high probe set sizes.

Figure 8. *A Bootstrapping Simulation of Experiment 4.*



Note. The probe suppression effects in the ten-letter probe condition of Experiment 4 (N = 40) were resampled to a random subset 24 participants in 10,000 simulations. This histogram depicts the frequency of probe capture effects at each resampling. The probe capture effect observed by Wang and Theeuwes (2020) falls just within the two-tailed 95% confidence interval.

GENERAL DISCUSSION

A longstanding debate has been whether physically salient items can automatically capture attention. Recent evidence has suggested a potential resolution. According to the signal suppression hypothesis, salient items do attract attention, but can be proactively inhibited to prevent attentional capture (Gaspelin & Luck, 2018c). Although this potential resolution has gained much recent support, it has recently been challenged by stimulus-driven accounts. Wang and Theeuwes (2020) used a similar capture-probe paradigm to Gaspelin et al. (2015) but increased the salience of the color singleton by increasing the set size of the search display. At high set sizes, the singleton produced a probe capture effect, which was taken as evidence that salient items cannot be suppressed. An important caveat is that the manipulation of set size was confounded with the number of probe letters. This potentially problematic because high probe set sizes can yield floor-like effects that make probe suppression effects difficult to observe. The current study therefore introduced a new approach that allows one to assess probe suppression effects at large set sizes.

Experiment 1 used a capture-probe paradigm with a set size of 10 items (as opposed to 4 or 6 items; Gaspelin et al., 2015). Importantly, the number of probe letters was manipulated independently. On ten-letter probe trials, all search items were probed (as in Wang & Theeuwes, 2020). On four-letter probe trials, only 4 locations were probed (as in Gaspelin et al., 2015). This new four-letter probe condition should boost overall probe recall accuracy and eliminate any potential floor effects that make suppression effects difficult to observe. The results indicated no evidence of attentional capture by the color singleton. On search trials, there was no singleton presence cost on manual RT. On probe trials, probe recall accuracy was reduced for singleton distractor compared to the average nonsingleton distractor. Critically, the magnitude of the probe suppression effect was larger on the four-letter probe trials (7.4%) than the ten-letter probe trials (2.5%). These results suggest that highly salient color singletons can be suppressed, but that suppression effects will be difficult to observe at large probe set sizes.

Experiment 2 further increased the salience of the color singleton by increasing the set size to 30 items. This experiment produced the same basic pattern of results as Experiment 1. There was no evidence of capture by the color singleton. Importantly, probe suppression effects were larger in the four-letter probe condition (5.3%) than the ten-letter probe condition (0.9%). These results again suggest that highly salient color singletons can be suppressed, but that large probe set sizes can make suppression effects difficult to observe. Experiments 3 and 4 adapted the search stimuli to be similar to Wang and Theeuwes (2020). Experiment 3 was a conceptual replication that used unfilled shapes on a gray background, but the same shapes and colors of Experiment 1. Experiment 4 was a direct replication that used the exact same search stimuli (shape sizes, colors, eccentricities) as Wang and Theeuwes (2020). Neither experiment replicated the probe capture effect observed in the previous study. Instead, both experiments suggested that the singleton was suppressed and that the magnitude of suppression effects was greater at low probe set sizes.

Altogether, the current study indicates that color singletons can be suppressed, even when highly salient. In all experiments, the probe results indicated that the color singleton was inhibited below baseline levels of processing. Thus, we did not merely observe a lack of capture by the color singleton (i.e., a null effect). Rather, we observed a significant reduction in attentional allocation to the singleton compared to other search items. This pattern of results is consistent with the signal suppression hypothesis, which proposes that salient stimuli can be proactively suppressed to prevent capture. This result is inconsistent with stimulus-driven accounts which predict that attention should be automatically allocated to salient items.

Floor effects are certainly part of the problem with the capture-probe paradigm used by Wang and Theeuwes (2020). Across four experiments, probe suppression effects were consistently reduced in the ten-letter probe condition compared to the four-letter probe condition, even though the singleton was equally salient in both conditions. Furthermore, a simulation of probe performance revealed that suppression effects are guaranteed to be reduced as the probe set size increases due to a capacity limitation in the number of items an individual can recall in the probe task (see Online Supplementary Materials). Future studies should therefore be careful to avoid exceptionally high probe set sizes in the capture-probe paradigm. This may involve probing a subset of items as in the current study or employing an alternate probing technique (e.g., see Chang & Egeth, 2019).

Floor effects, however, may not be the entire story as to why Wang and Theeuwes (2020) found a small-but-significant probe capture effect. None of the four experiments in the current study replicated the probe capture effect, even at a probe set size of 10. Experiment 4 even tried to directly replicate this study by using the exact same stimuli (shapes, sizes, luminance, etc.) and an increased sample size to detect the probe capture effect. Yet no probe suppression effect was observed. One possibility, therefore, is that the probe capture effect was a spurious result. The bootstrapping simulation certainly suggested that it would be possible to obtain a small probe capture effect, with a smaller sample size.

It is also possible, however, that the opposing results are due to some aspect of experimental design. One potential issue is that the search displays of Wang and Theeuwes (2020) may have promoted a kind of singleton detection mode (Bacon & Egeth, 1994). Each search display contained two unique shapes (circle and diamond) amongst eight relatively homogenous shapes (four squares and four hexagons). Participants may have simplified search by first locating the two shape singletons (circle and diamond) before searching for the specific target shape. This strategy of searching for popout stimuli may have caused attentional capture by the salient color singleton. This strategy seems particularly likely in the circle target condition, where the target was the only curvilinear shape and was slightly smaller than the other shapes. In fact, the salience maps (Experiment 4) actually suggested that the circle was highly salient. This same problem is also prominent in Theeuwes (2004) in which three unique shapes (diamond, triangle, square) were presented amongst 17 relatively small circle distractors. Although it was suggested this type of display encourages a feature search, it seems possible that participants could drastically reduce the set size by locating the three popout shapes by using singleton detection mode.

Eliminating the possible use of singleton detection mode is tricky but important in studies of attentional capture. Previous studies have suggested that, when the target has a specific feature (e.g., shape) but is also salient, participants can choose whether to search for a specific feature (feature-search mode) or instead search for the salient item (singleton detection mode; Leber & Egeth, 2006b) (see also Cosman & Vecera, 2013; Leber & Egeth, 2006a). This “subject-option mode” makes it difficult to ascertain whether observed capture by salient distractors is due to bottom salience or a top-down attentional set for salience. Thus, it is imperative to ensure that the target is not salient in attentional capture studies. For example, Experiments 1-3 of the current study were careful to prevent singleton detection mode by including several different distractor shapes to improve distractor heterogeneity and including curvilinear distractors to prevent the circle target from becoming salient. Future studies will certainly be needed to explore this issue in depth.

The current study introduced a new technique to verify the salience of the color singleton via saliency maps. This is important to avoid potentially circular logic: One cannot test whether salient items capture attention, and also use capture effects to determine whether an item was salient (see Chang et al., in press). Our basic approach was to generate saliency maps of our stimuli using the Image Signature Toolbox (Hou et al., 2012) and then examine whether the singleton was more salient than other search items. This approach yielded several interesting findings. First, the singleton was ranked as highly salient in Experiments 1–3, even though behavioral performance showed that the singleton was suppressed. This directly opposes the predictions of a stimulus-driven account. Second, the saliency maps confirmed that increasing the set size of search arrays did increase the relative salience of the color singleton, consistent with the claims of Wang and Theeuwes (2020). The GSI was lower in Experiment 1 (set size 10; GSI = 0.68) than Experiment 2 (set size 30; GSI = 0.83). Finally, the saliency maps also suggested that using unfilled shapes on a gray background can actually reduce the salience of the color singleton (see Experiments 3 and 4). In summary, saliency maps provided several useful insights in this study and seem like a potentially useful tool for future studies.

An important aspect of the current design was that the manipulation of probe size occurred randomly by trial, which should have prevented any strategic differences in attentional deployment between probe size conditions. Alternatively, manipulating probe size between subjects or in a blocked design would allow participants to predict the upcoming probe size and this could possibly cause participants to use different strategies for each probe size condition. For example, a consistent low probe set size might cause participants to distribute the attentional window broadly in an attempt to report all four probe letters, which would be practically impossible in the ten-letter condition due to limitations of visual working memory (Luck & Vogel, 2013). In the current design, the upcoming probe size was unpredictable on each trial, which effectively prevents any a priori knowledge of the upcoming probe display size and thus eliminates any such possibility. This may, however, be an important issue for future research using probe techniques.

One additional line of evidence supporting the signal suppression hypothesis has come from studies of the distractor positivity (P_D) event-related potential component, which have shown that the P_D component seems to occur in the absence of attentional capture by salient items (Berggren & Eimer, 2020; Cosman et al., 2018; Feldmann-Wüstefeld et al., 2020; Gaspar et al., 2016; Gaspar & McDonald, 2014; Gaspelin & Luck, 2018a; Sawaki & Luck, 2010). Recently, an alternative account of the P_D component has been proposed, which also criticizes the low set sizes of commonly employed in the capture-probe paradigm (Kerzel & Burra, 2020). This model does not propose that low set sizes reduce the salience of the singleton. Rather, it proposes that the P_D component may reflect contingent capture by the target-colored distractor contralateral to the singleton distractor. Although this alternative account can potentially explain the P_D component in Gaspelin and Luck (2018a), it seems unable to explain why the P_D component has been observed in many studies using relatively large set sizes (Feldmann-Wüstefeld et al., 2020; Gaspar et al., 2016; Gaspar & McDonald, 2014; Sawaki & Luck, 2010). In any case, future research may help draw more definitive conclusions about the relationship of the P_D component to the set size of the search display.

In summary, the current study demonstrated that even highly salient color singleton distractors can be suppressed, supporting the signal suppression hypothesis (Gaspelin & Luck, 2018c, 2019; Luck et al., 2021) and refuting stimulus-driven accounts of attentional capture. Furthermore, the current study introduced a new probe technique to overcome floor effects at exceptionally large set sizes. When the probe set size is kept low, there is no indication that the singleton distractor captures attention.

Acknowledgments

We would like to thank Steve Luck, Eric Ruthruff, and Howard Egeth for their valuable feedback.

Online Supplemental Material

Supplemental data to this article can be accessed online at https://osf.io/bwg2s under "Online Supplementary Material."

References

Bacon, W. F., & Egeth, H. E. (1994). Overriding stimulus-driven attentional capture. *Perception & Psychophysics*, *55*(5), 485–496. https://doi.org/10.3758/BF03205306

Becker, S. I. (2007). Irrelevant singletons in pop-out search: Attentional capture or filtering costs? *Journal of Experimental Psychology: Human Perception and Performance*, *33*(4), 764–787. https://doi.org/10.1037/0096-1523.33.4.764

Berggren, N., & Eimer, M. (2020). Spatial filtering restricts the attentional window during both singleton and feature-based visual search. *Attention, Perception & Psychophysics*, *82*(5), 2360–2378. https://doi.org/10.3758/s13414-020-01977-5

Chang, S., & Egeth, H. E. (2019). Enhancement and Suppression Flexibly Guide Attention. *Psychological Science*. https://doi.org/10.1177/0956797619878813

Chang, S., Niebur, E., & Egeth, H. (in press). Standing out in a small crowd: The role of display size in attracting attention. *Visual Cognition*.

Cosman, J. D., Lowe, K. A., Woodman, G. F., & Schall, J. D. (2018). Prefrontal Control of Visual Distraction. *Current Biology*, *28*(3), 414-420.e3. https://doi.org/10.1016/j.cub.2017.12.023

Cosman, J. D., & Vecera, S. P. (2013). Context-dependent control over attentional capture. *Journal of Experimental Psychology: Human Perception and Performance*, *39*(3), 836–848. https://doi.org/10.1037/a0030027

Cousineau, D. (2005). Confidence intervals in within-subject designs: A simpler solution to Loftus and Masson's method. *Tutorials in Quantitative Methods for Psychology*, *1*(1), 42–45. https://doi.org/10.20982/TQMP.01.1.P042

Efron, B., & Tibshirani, R. J. (1993). *An Introduction to the Bootstrap* (1st ed.). Chapman and Hall/CRC.

Egeth, H. E., Leonard, C. J., & Leber, A. B. (2010). Why salience is not enough: Reflections on top-down selection in vision. In *Acta Psychologica* (Vol. 135). https://doi.org/10.1016/j.actpsy.2010.05.012

Feldmann-Wüstefeld, T., Busch, N. A., & Schubö, A. (2020). Failed Suppression of Salient Stimuli Precedes Behavioral Errors. *Journal of Cognitive Neuroscience*, *32*(2), 367–377. https://doi.org/10.1162/jocn_a_01502

Folk, C. L., Leber, A. B., & Egeth, H. E. (2002). Made you blink! Contingent attentional capture produces a spatial blink. *Perception and Psychophysics*, *64*(5), 741–753. https://doi.org/10.3758/BF03194741

Folk, C. L., & Remington, R. W. (1998). Selectivity in distraction by irrelevant featural singletons: Evidence for two forms of attentional capture. *Journal of Experimental Psychology: Human Perception and Performance*, *24*(3), 847–858.

Gaspar, J. M., Christie, G. J., Prime, D. J., Jolicœur, P., & McDonald, J. J. (2016). Inability to suppress salient distractors predicts low visual working memory capacity. *Proceedings of the National Academy of Sciences*, *113*(13), 3693. https://doi.org/10.1073/pnas.1523471113

Gaspar, J. M., & McDonald, J. J. (2014). Suppression of salient objects prevents distraction in visual search.

The Journal of Neuroscience, *34*(16), 5658–5666. https://doi.org/10.1523/JNEUROSCI.4161-13.2014

Gaspelin, N., Gaspar, J. M., & Luck, S. J. (2019). Oculomotor inhibition of salient distractors: Voluntary inhibition cannot override selection history. *Visual Cognition*, *27*(3–4), 227–246. https://doi.org/10.1080/13506285.2019.1600090

Gaspelin, N., Leonard, C. J., & Luck, S. J. (2015). Direct Evidence for Active Suppression of Salient-but-Irrelevant Sensory Inputs. *Psychological Science*, *26*(11), 1740–1750. https://doi.org/10.1177/0956797615597913

Gaspelin, N., Leonard, C. J., & Luck, S. J. (2017). Suppression of overt attentional capture by salient-but-irrelevant color singletons. *Attention, Perception, and Psychophysics*, *79*(1), 45–62. https://doi.org/10.3758/s13414-016-1209-1

Gaspelin, N., & Luck, S. J. (2018a). Combined electrophysiological and behavioral evidence for the suppression of salient distractors. *Journal of Cognitive Neuroscience*, *30*(9), 1265–1280. https://doi.org/10.1162/jocn_a_01279

Gaspelin, N., & Luck, S. J. (2018b). Distinguishing among potential mechanisms of singleton suppression. *Journal of Experimental Psychology: Human Perception and Performance*, *44*(4), 626–644. https://doi.org/10.1037/xhp0000484

Gaspelin, N., & Luck, S. J. (2018c). The Role of Inhibition in Avoiding Distraction by Salient Stimuli. *Trends in Cognitive Sciences*, *22*(1), 79–92. https://doi.org/10.1016/j.tics.2017.11.001

Gaspelin, N., & Luck, S. J. (2019). Inhibition as a potential resolution to the attentional capture debate. *Current Opinion in Psychology*, *29*, 12–18. https://doi.org/10.1016/j.copsyc.2018.10.013

Hou, X., Harel, J., & Koch, C. (2012). Image Signature: Highlighting Sparse Salient Regions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *34*(1), 194–201. https://doi.org/10.1109/TPAMI.2011.146

Kerzel, D., & Burra, N. (2020). Capture by context elements, not attentional suppression of distractors, explains the PD with small search displays. *Journal of Cognitive Neuroscience*, *32*(6), 1170–1183.

Kotseruba, I., Wloka, C., Rasouli, A., & Tsotsos, J. K. (2020). Do Saliency Models Detect Odd-One-Out Targets? New Datasets and Evaluations. *ArXiv Preprint ArXiv:2005.06583*.

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and ANOVAs. *Frontiers in Psychology*, *4*. https://doi.org/10.3389/fpsyg.2013.00863

Lamy, D., & Egeth, H. E. (2003). Attentional capture in singleton-detection and feature-search modes. *Journal of Experimental Psychology. Human Perception and Performance*, *29*(5), 1003–1020. https://doi.org/10.1037/0096-1523.29.5.1003

Leber, A. B., & Egeth, H. E. (2006a). Attention on autopilot: Past experience and attentional set. *Visual Cognition*, *14*(4–8), 565–583. https://doi.org/10.1080/13506280500193438

Leber, A. B., & Egeth, H. E. (2006b). It's under control: Top-down search strategies can override attentional capture. *Psychonomic Bulletin & Review*, *13*(1), 132–138.

Lien, M.-C., Ruthruff, E., Goodin, Z., & Remington, R. W. (2008). Contingent attentional capture by top-down control settings: Converging evidence from event-related potentials. *Journal of Experimental Psychology: Human Perception and Performance*, *34*(3), 509–530. https://doi.org/10.1037/0096-1523.34.3.509

Lien, M.-C., Ruthruff, E., & Johnston, J. C. (2010). Attentional capture with rapidly changing attentional control settings. *Journal of Experimental Psychology: Human Perception and Performance*, *36*(1), 1–16. https://doi.org/10.1037/a0015875

Luck, S. J., Gaspelin, N., Folk, C. L., Remington, R. W., & Theeuwes, J. (2021). Progress toward resolving the attentional capture debate. *Visual Cognition*, *29*(1), 1–21. https://doi.org/10.1080/13506285.2020.1848949

Luck, S. J., & Vogel, E. K. (2013). Visual working memory capacity: From psychophysics and neurobiology to individual differences. *Trends in Cognitive Sciences*, *17*(8), 391–400. https://doi.org/10.1016/j.tics.2013.06.006

Maxwell, J. W., Gaspelin, N., & Ruthruff, E. (2020). No identification of abrupt onsets that capture attention: Evidence against a unified model of spatial attention. *Psychological Research*. https://doi.org/10.1007/s00426-020-01367-4

Mordkoff, J. T. (2019). A Simple Method for Removing Bias From a Popular Measure of Standardized Effect Size: Adjusted Partial Eta Squared. *Advances in Methods and Practices in Psychological Science*, *2*(3), 228–232. https://doi.org/10.1177/2515245919855053

Morey, R. D. (2008). Confidence Intervals from Normalized Data: A correction to Cousineau (2005). *Tutorials in Quantitative Methods for Psychology*. https://doi.org/10.20982/tqmp.04.2.p061

Ruthruff, E., & Gaspelin, N. (2018). Immunity to attentional capture at ignored locations. *Attention, Perception, and Psychophysics*, *80*(2), 325–336. https://doi.org/10.3758/s13414-017-1440-4

Sawaki, R., & Luck, S. J. (2010). Capture versus suppression of attention by salient singletons: Electrophysiological evidence for an automatic attend-to-me signal. *Attention, Perception, & Psychophysics*, *72*(6), 1455–1470. https://doi.org/10.3758/APP

Stilwell, B. T., Bahle, B., & Vecera, S. P. (2019). Feature-based statistical regularities of distractors modulate attentional capture. *Journal of Experimental Psychology: Human Perception and Performance*, *45*(3), 419–433. https://doi.org/10.1037/xhp0000613

Theeuwes, J. (1991). Exogenous and endogenous control of attention: The effect of visual onsets and offsets. *Perception & Psychophysics*, *49*(1), 83–90. https://doi.org/10.3758/BF03211619

Theeuwes, J. (1992). Perceptual selectivity for color and form. *Perception & Psychophysics*, *51*(6), 599–606. https://doi.org/10.3758/BF03211656

Theeuwes, J. (2004). Top-down search strategies cannot override attentional capture. *Psychonomic Bulletin and Review*, *11*(1), 65–70. https://doi.org/10.3758/BF03206462

Vatterott, D. B., Mozer, M. C., & Vecera, S. P. (2018). Rejecting salient distractors: Generalization from experience. *Attention, Perception, and Psychophysics*, *80*(2), 485–499. https://doi.org/10.3758/s13414-017-1465-8

Wang, B., & Theeuwes, J. (2018a). Statistical regularities modulate attentional capture. *Journal of Experimental Psychology: Human Perception and Performance*, *44*(1), 13–17. https://doi.org/10.1037/xhp0000472

Wang, B., & Theeuwes, J. (2018b). How to inhibit a distractor location? Statistical learning versus active, top-down suppression. *Attention, Perception, & Psychophysics*, *80*(4), 860–870. https://doi.org/10.3758/s13414-018-1493-z

Wang, B., & Theeuwes, J. (2020). Salience Determines Attentional Orienting in Visual Selection. *Journal of Experimental Psychology: Human Perception and Performance*. https://doi.org/10.1037/xhp0000796

Weaver, M. D., van Zoest, W., & Hickey, C. (2017). A temporal dependency account of attentional inhibition in oculomotor control. *NeuroImage*, *147*, 880–894. https://doi.org/10.1016/j.neuroimage.2016.11.004

Won, B.-Y., & Geng, J. J. (2020). Passive exposure attenuates distraction during visual search. *Journal of Experimental Psychology: General*. APA PsycInfo. https://doi.org/10.1037/xge0000760

Won, B.-Y., Kosoyan, M., & Geng, J. J. (2019). Evidence for second-order singleton suppression based on probabilistic expectations. *Journal of Experimental Psychology: Human Perception and Performance*, *45*(1), 125–138. https://doi.org/10.1037/xhp0000594

Yantis, S., & Jonides, J. (1984). Abrupt visual onsets and selective attention: Evidence from visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *10*(5), 601–621. https://doi.org/10.1037/0096-1523.10.5.601