Salience Determines Attentional Orienting in Visual Selection

Benchí Wang  
Ministry of Education, Guangzhou, China, and South China Normal University

Jan Theeuwes  
Vrije Universiteit Amsterdam

Recently, the signal-suppression account was proposed, positing that salient stimuli automatically produce a bottom-up salience signal that can be suppressed via top-down control processes. Evidence for this hybrid account came from a capture-probe paradigm that showed that while searching for a specific shape, observers suppressed the location of the irrelevant color singleton. Here we replicate these findings but also show that this occurs only for search arrays with 4 elements. For larger array sizes when both target and distractor singleton are salient, there is no evidence for suppression; instead, and consistent with the stimulus-driven account, there is clear evidence that the salient distractor captured attention. The current study shows that the relative salience of items in the display is a crucial factor in attentional control. In displays with a few heterogeneous items, top-down suppression is possible. However, in larger displays in which both target and distractor singletons are salient, no top-down suppression is observed. We conclude that the signal-suppression account cannot resolve the long-standing debate regarding stimulus-driven and goal-driven attentional capture.

**Public Significance Statement**

This study replicated the critical findings (i.e., suppression effect) reported by Gaspelin, Leonard, and Luck, 2015, when only four elements were presented on the display. Yet with larger search arrays (six and 10 items), the target and distractor singleton become more salient, and then there was no sign of any suppression; instead, and consistent with the stimulus-driven account, there is clear evidence that the salient distractor captured attention. We argued that signal suppression (inhibitory processes) occurs only in displays with a limited number of nonsalient elements allowing for a peculiar (most likely serial) search strategy. The current experiment provides the boundary conditions of when top-down suppression is effective.

**Keywords:** visual selection, attentional capture, signal suppression, salience

It is well known that salient distractors can distract us and interfere with our ongoing task (Theeuwes, 1991, 1992). According to the stimulus-driven account, physically salient objects capture attention, regardless of the intentions of the observers (Theeuwes, 2010). This effect was demonstrated in the so-called additional singleton paradigm (Theeuwes, 1991, 1992), in which observers searched for a shape singleton (e.g., a diamond shape among circles) while an irrelevant color singleton was present. The results showed that the irrelevant color singleton captured attention, even though it was never relevant for the search goal.

The stimulus-driven account was challenged by research demonstrating that if observers search for a specific shape (e.g., a diamond shape among circle, square, and triangle distractors), the capture by the irrelevant singleton can be avoided (Bacon & Egeth, 1994; Leber & Egeth, 2006). It was argued that instead of searching for any salient singleton (called the singleton detection mode), a top-down set for searching for a specific shape (called the feature search mode) could prevent attentional capture (Bacon et al., 1994). Clearly, these studies demonstrate that under some circumstances capture can be prevented.

In the capture debate, recently a new and appealing view was suggested that implicated a major step forward in resolving the capture debate (Gaspelin & Luck, 2018b). Rather than simply observing less capture when observers adopt a feature search mode, the new account provides an explanation why capture is smaller or even absent when the feature search mode is engaged. According to the signal suppression account (Gaspelin et al., 2015;
Sawaki & Luck, 2010), when observers engage the top-down feature search mode, salient distractors still generate a large bottom-up signal, but the signal can be proactively suppressed in a top-down way, thereby preventing a shift of attention toward the salient distractor. In other words, consistent with stimulus-driven accounts (Theeuwes, 1992, 2010), it is argued that physically salient stimuli have the ability to attract attention; yet, consistent with goal-driven theories (Folk & Remington, 2010; Folk, Remington, & Johnston, 1992), inhibitory processes can suppress these salient signals if participants exert top-down control. As such, the signal suppression account seems to be able to resolve a long-standing conflict between stimulus-driven and goal-driven theories of attentional capture.

Gaspelin et al. (2015) provided compelling evidence for this notion in an innovative study in which the additional singleton task was combined with a letter probe task. In 70% of trials, participants searched for a target shape while ignoring a color singleton. In 30% of trials, letter probes were briefly presented inside the search elements, and participants were required to report as many letters as possible. In conditions in which participants used the singleton detection mode, Gaspelin et al. (2015) showed that participants reported more letters when these were presented inside the singleton distractor than presented inside the nonsingleton distractors. This was considered as a clear evidence for attentional capture by the salient distractor (see also Kim & Cave, 1999 for similar results using a probe task). Critically, however, when observers engaged the feature search mode, the accuracy for the letter inside the irrelevant singleton distractor was reduced below the accuracy observed for letters inside the nonsingleton distractors. This was considered as decisive evidence for active suppression of the singleton-distractor location. This result of subbaseline suppression was considered to be the essential finding for resolving the 25-year-old debate regarding the role of top-down and bottom-up control of attentional selection (Gaspelin & Luck, 2018b, 2018c).

Unlike the signal-suppression account of Gaspelin et al. (2015) that assumes the capture is prevented because the salience signal of the singleton-distractor is suppressed, an alternative account known as rapid disengagement (Theeuwes, 2010) claims that spatial attention is initially captured by the singleton distractor and then rapidly disengaged from it. According to this account, top-down disengagement occurs following the initial shift of spatial attention to the singleton distractor location. Gaspelin et al. (2015) recognized that the rapid disengagement account was a viable explanation for their results (Experiments 2 and 3) because the probe was presented 200 ms after the search display, providing plenty of time for disengagement to occur. Yet they addressed these issues in their final Experiment 4 in which the letters were presented simultaneously with the search display and found basically the same results. It rules out the rapid disengagement account and firmly establishes their signal-suppression account.

Even though the final experiment of Gaspelin et al. (2015) seems to be decisive in the debate, there is one major concern. In their experiment (and also in a recent study by Chang and Egeth, 2019, and a replication with electroencephalographic recording by Gaspelin & Luck, 2018a), the display consisted only of four elements (a target, a singleton distractor, and two nonsingleton distractors), rendering both the target and the distractor nonsalient. In a study published in 2004, Theeuwes (2004) had already shown that the number of elements in the display is critical for the salience signal of both the target and distractor. This is especially the case in displays in which distractor shapes are heterogeneous (e.g., squares, hexagons and circles), which is typically done to induce the feature search mode preventing the target from popping out. The critical message of Theeuwes (2004) was that with a few heterogeneous elements in the display, nothing in the display is salient enough to capture attention. Not surprisingly that in displays like these (i.e., displays used by Gaspelin et al., 2015 having only a few heterogeneous items), the distractor no longer captures attention, not because of a top-down search mode labeled as feature search but simply because the singleton distractor is no longer salient.

Previous research has shown that the saliency of items depends on two factors: local feature contrast (Nothdurft, 1993) and distractor-distractor similarity (Duncan & Humphreys, 1989). Local feature contrast refers to how different an item is from nearby items (Nothdurft, 1993). In small display sizes, when items are equally spaced around the fixation point, they are relatively far apart, reducing the local feature contrast. In addition, to force participants to engage in feature search, the display needs to be heterogeneous (e.g., squares, hexagons, and circles), which negatively affects the distractor-distractor similarity (Duncan et al., 1989). It was shown that distractor heterogeneity reduces search efficiency, resulting in serial search (Duncan et al., 1989). It is clear that both factors play an important role when engaging in feature search among only a few items. Even though a distractor in these displays with only a few items may be unique and as such should be considered as a singleton, it does not necessarily mean that the distractor is also salient. If none of the display items are salient, it is not surprising that no capture is observed.

Given these concerns, we aimed at replicating the critical Experiment 4 of Gaspelin et al. (2015) but now with various search array sizes (four, six, and ten elements). When the number of elements is relatively large, it is expected that both the target and the singleton distractor are salient enough because they will stand out from the background (see Theeuwes, 2004). As in Gaspelin et al. (2015), we also used heterogeneous displays to induced the feature search mode, and we presented the probes simultaneously with the search display. According to the signal-suppression account, there should be active suppression of the salient singleton distractor not only when the singleton distractor is relatively nonsalient (with search array size four, as in Gaspelin et al., 2015) but also when the distractor is relatively salient (with search array sizes six and 10). According to the stimulus-driven account (Theeuwes, 1992, 2004, 2010), if a distractor is salient enough, it should always capture attention, regardless of the search mode used.

**Method**

The study was approved by both the Ethical Review Committee of the Vrije Universiteit Amsterdam and the Ethical Review Committee of Zhejiang Normal University.

**Participants**

Seventy-two undergraduates (eight men and 64 women; with a mean age of 19.2 ± 1.1 years) were recruited from Zhejiang
Normal University in China for monetary compensation. All participants provided written informed consent before the study and reported normal color vision and normal or corrected-to-normal visual acuity. Participants were equally and randomly allocated to the different search array size conditions. Sample size was predetermined based on previous studies (Gaspelin et al., 2015; Gaspelin & Luck, 2019). According to the power analysis described by Gaspelin and Luck (2019; p. 8), “We estimated the population effect size and standard deviation by pooling the probe suppression effects (the difference between singleton distractor location and non-singleton distractor location in probe task) across participants from three previous experiments that were similar in methodology to the current experiment (Gaspelin et al., 2015, Experiments 2–4), yielding an N of 72 participants. The probe suppression effects in the pooled data were quite robust, with an effect size of $dz = .97$. Thus, to achieve a power of 95% and an alpha of 5% with this effect size, a sample size of 16 participants would be needed.” Because our experiment involved a between-subjects design, we adopted a sample size of 24 subjects. If there is active suppression, this large sample size should be enough to detect it.

**Apparatus and Stimuli**

Participants were seated in a dimly lit laboratory, 63 cm away from the liquid crystal display color monitor with their chin on a chinrest. The background was gray (red-green-blue [RGB] = 128, 128, 128). As illustrated in Figure 1A, the primary search display for search array size four contained one circle with a radius of 0.7°, one diamond (subtended by 1.6°), one square (subtended by 1.6°), and one hexagon (subtended by 1.6°); for search array size six, there was one circle, one diamond, two squares, and two hexagons; for search array size 10, there was one circle, one diamond, four squares, and four hexagons. Those display elements were colored in red (RGB = 255, 0, 0) or green (RGB = 0, 255, 0), and were centered 3.0° from the fixation (a black cross, 0.5° × 0.5°, RGB = 0, 0, 0), containing a 0.2° black dot located 0.2° left or right from the center of the element. On search-probe trials, a white (RGB = 255, 255, 255) uppercase letter (0.75° tall) was presented in Simhei typeface at the center of each search element. All possible letters from the English alphabet were presented subsequently as a response display. Stimulus presentation and response registration were controlled by custom scripts written in Python 2.7.

**Procedure and Design**

The procedure and task were virtually identical to that of Experiment 4 of Gaspelin et al. (2015) except that we varied search array sizes (four, six, and 10 elements) between participants. On each trial, a fixation cross was presented for 500 ms, followed by a primary search display in which participants were asked to search for a specific shape (for half of the participants a circle shape; for the other half, a diamond shape). Participants were required to keep fixation at the cross throughout a trial. The search target was presented in each trial and appeared equally often at each location. A uniquely colored singleton was used as the salient distractor in half of the trials, with a different color as the target and other display elements (red or green balanced between participants).

On search-only trials (two thirds of the trials), the search array was presented for 3,000 ms or until participants responded (as shown in Figure 1B). Participants were required to indicate

---

**Figure 1.** (A) Example of search displays for different search array sizes. (B) The procedure in search-only task, in which participants were required to search for a specific shape (for different participants either a circle or a diamond shape) and to indicate the position (i.e., left or right) of the black dot inside. (C) The procedure in search-probe task, in which the search display and the probe letters were presented simultaneously for 100 ms. Participants were required to memorize the letters and to recall them by pressing the corresponding key on the keyboard as accurately as possible. See the online article for the color version of this figure.
whether the dot was on the left or right side of the target shape (i.e., the specific shape, circle or diamond) by pressing the left or right key on the keyboard using left hand as fast as possible, respectively. Responses were speeded and feedback, “You did not respond; please respond as fast as possible” or “Incorrect response; please focus on the task,” was given when participants did not respond or responded incorrectly, respectively.

On search-probe trials (one third of the trials), as shown in Figure 1C, the search array and the probe display (containing to-be-memorized letters) appeared simultaneously for 100 ms. For those trials, participants did not have to respond to the search array but had to attend and memorize as many letters as possible. For each trial, the letters were selected randomly, without replacement, from the English alphabet. Then a response display was presented until that participants responded. Participants had to recall as many letters as possible by pressing the corresponding letter keys on the keyboard without time pressure, and only accuracy was emphasized. Once the letter was selected, it turned into red. When they finished their response, they pressed the space key to continue.

Participants were first trained for a small number of trials to make sure they understood the task before testing started. Different search array sizes were tested between participants. For search array size four, participants completed four blocks with each containing 96 trials (a total of 384 trials). For search array size six, they completed three blocks with each containing 144 trials (a total of 432 trials). For search array size 10, they completed four blocks with each containing 120 trials (a total of 480 trials). The search-only and search-probe trials were mixed within blocks.

### Results

**Search Array Size Four**

**Search-only condition.** Trials (1.7%) on which the response times (RTs) were slower than 2,000 ms were removed from analysis. Mean RTs are presented in Figure 2A, left panel. Mean RTs were the same for distractor present (831 ms) and absent conditions (834 ms), $t(23) = 0.63, p = .53$, Cohen’s $d = 0.21$, $BF_{01} = 3.89$. The same pattern of results was observed for error rates: distractor present (1.1%) versus distractor absent (1.4%), $t(23) = 0.93, p = .36$, Cohen’s $d = 0.02$, $BF_{01} = 3.15$. These findings are consistent with those of Gaspelin et al., 2015, showing no attentional capture for search array size four when a feature search mode is used.

**Search-probe condition.** We first calculated the proportion of probes that were recalled when they were presented at the target location, at the singleton distractor location, and at each nonsingleton distractor location, and then averaged across the nonsingleton distractor locations, to provide the probe recall accuracy for each location. As illustrated in Figure 2B, left panel, the probe recall accuracy was significantly lower when the probe was presented at the singleton distractor location (19.4%) than at the nonsingleton distractor location (22.8%), $t(23) = 2.53, p = .02$, Cohen’s $d = 0.94$, providing an exact replication of the critical finding of Gaspelin et al., 2015.

For the remaining conditions, a repeated-measures ANOVA on probe recall accuracy with the factors of distractor condition...
(singleton distractor present and absent) and probe type (target and nonsingleton distractor) was conducted. No significant main effects nor significant interactions were observed, all $Fs < 1.96$, all $ps > .18$.

**Search Array Size Six**

**Search-only condition.** Trials (1.2%) on which the RTs were slower than 2,000 ms were removed from analysis. Mean RTs are presented in Figure 2A, middle panel. Unlike the results of array size four, with increasing the array size, singleton distractor started capturing attention because mean RTs were higher when the distractor was present (886 ms) than when it was absent (869 ms), $t(23) = 2.39$, $p = .03$, Cohen’s $d = 0.97$. There was no effect on mean error rates, $t(23) = 0.89$, $p = .38$, Cohen’s $d = 0.02$, BF$_{10} = 3.26$.

**Search-probe condition.** As shown in Figure 2B, middle panel, the probe recall accuracy was basically identical for probes presented at the singleton distractor location (19.3%) compared with nonsingleton distractor location (20.4%), $t(23) = 1.05$, $p = .31$, Cohen’s $d = 0.28$, BF$_{10} = 2.86$.

For the remaining conditions, a repeated-measures ANOVA on probe recall accuracy with the factors of distractor condition (singleton distractor present and absent) and probe type (target and nonsingleton distractor) was conducted. No significant main effects or significant interactions were observed, all $Fs < 2.73$, all $ps > .11$.

**Search Array Size Ten**

**Search-only condition.** Trials (3.3%) on which the RTs were slower than 2,000 ms were removed from analysis. Mean RTs are presented in Figure 2A, right panel. For search array size 10, there was a strong attentional capture because RTs were slower when the distractor was present (988 ms) than when it was absent (945 ms), $t(23) = 5.07$, $p < .01$, Cohen’s $d = 2.21$. There was no effect on mean error rates, $t(23) = 0.95$, $p = .35$, Cohen’s $d = 0.01$, BF$_{10} = 3.11$.

**Search-probe condition.** Unlike what was found for array size four, now there is a reversed pattern because probe recall accuracy was higher when the probe was presented at the singleton distractor location (18.1%) than at the nonsingleton distractor location (16.2%), $t(23) = 2.76$, $p = .01$, Cohen’s $d = 0.51$. Note that recall accuracy for probes presented at the target location was higher in the distractor absent condition than in the distractor present condition, suggesting that when attention is allocated to the singleton distractor, fewer resources are available for processing the target (see similar results in Experiment 1 in Gaspelin et al., 2015 when participants used the singleton detection mode to find the target).

For the remaining conditions, a repeated-measures ANOVA on probe recall accuracy with the factors of distractor condition (singleton distractor present and absent) and probe type (target and nonsingleton distractor) was conducted. The recall accuracy was higher for probes presented at the target location than at the nonsingleton distractor location, $F(1, 23) = 14.87$, $p < .01$, $\eta^2_p = .39$. Moreover, the recall accuracy was also higher on singleton distractor present versus absent trials, $F(1, 23) = 7.89$, $p = .01$, $\eta^2_p = .26$. There was no significant interaction, $F(1, 23) < 0.01$, $p = .99$, $\eta^2_p < .01$.

**Analysis Across Different Search Array Sizes**

For each subject, we first calculated the attentional capture effect in the search-only task and the distractor suppression effect in the search-probe task by using the mean RTs on distractor present trials minus that on distractor absent trials and the recall accuracy for probes presented at nonsingleton distractor location minus that for probes presented at singleton distractor location, respectively. Then we compared them across different search array sizes.

As illustrated in Figure 3A, there was a suppression effect for search array size four (3.4%), but no such an effect for search array size 6 (1.1%), and a reversed pattern for search array size 10.

![Figure 3](image_url) **Figure 3.** (A) The suppression effect (reflected by the recall accuracy for probes presented at the nonsingleton distractor location minus that at the singleton distractor location) in the search-probe task and the attentional capture (reflected by the mean response times on distractor present trials minus that on distractor absent trials) in the search-only task for different search array sizes. (B) The correlation between the suppression effect in the search-probe task and the attentional capture in the search-only task across different participants. **" $p < .01$. See the online article for the color version of this figure."
Attentional capture was not found for search array size four (−3 ms) but was present for search array size six (17 ms) and for search array size 10 (43 ms). We compared search array sizes four and 10 and found that the capture effect was larger for search array size 10 than that for size four, t(46) = 4.63, p < .001, Cohen’s d = 1.34; and the suppression effect was smaller for search array size 10 than that for size four, t(46) = 3.49, p = .001, Cohen’s d = 1.01. We also calculated the correlation between the suppression effect and the attentional capture across 72 different participants, with a significant Spearman correlation, r = −0.31, p = .01, BF_{10} = 0.25 (same for Pearson correlation, r = −0.29, p = .01, BF_{10} = 0.35; see Figure 3B). Overall, it indicates that the suppression effect observed in the probe task is negatively related to the capture effect in the search task.

**Discussion**

For search array size four, we perfectly replicated the results of Gaspelin et al. (2015) showing suppression of the singleton distractor location. Yet for array size six, there was no evidence for the suppression of the singleton distractor location. Critically, for search array size 10, the effect was reversed, now showing enhanced processing of probes that were presented at the singleton distractor location. This latter finding suggests that for larger search array sizes, singleton distractor captured attention, even though participants had to use the feature-search mode to find the target. Indeed, in this experiment, the target shape was kept constant and was presented among a heterogeneous set of distractor shapes, resulting in feature-search mode (i.e., shape-based search) and eliminating the possibility of searching for something unique (i.e., discouraging the singleton detection mode).

The current findings suggest that the signal-suppression account (Gaspelin et al., 2015), which assumes that salient-but-irrelevant singletons can be suppressed when participants engage the top-down feature search mode, is limited in scope because it seems to work only for arrays with a few nonsalient items. As such, this account alone cannot resolve the conflict between stimulus-driven and goal-driven theories of attentional capture (Folk & Remington, 2010; Folk et al., 1992; Theeuwes, 1992, 2010). Note that the current findings also showed that when both target and singleton distractors were sufficiently salient, there was attentional capture, regardless of what search mode was used (see also Theeuwes, 2004 for a similar argument).

The correlation that we reported between the suppression effect and the capture effect is compelling. Overall, it indicates that at an individual level, the suppression effect that is observed in the probe task is significantly correlated with the capture effect in the search task. This implies that, consistent with the stimulus-driven account, if the singleton distractor is salient enough, attentional resources are initially allocated to its location, resulting in a higher recall accuracy for probes presented at the singleton distractor than at nonsingleton distractor locations.

It is possible that there is suppression of the singleton distractor location after spatial attention has been initially captured by the singleton distractor, as has been suggested by the rapid disengagement account of Theeuwes (2010). This is consistent with Kim et al. (1999), who also used the additional singleton paradigm in combination with a probe-detection task. They showed that 60 ms after display onset, probe RTs at the singleton distractor location were about 20 ms faster than at the target location, indicating initial capture. Yet at an stimulus onset asymmetry of 150 ms, this pattern was reversed: The probe RTs at the target location were about 15 ms faster than at the distractor location, indicating that attention was rapidly disengaged from the singleton-distractor location and redirected to the target location.

It should be noted that recently Gaspelin and Luck (2018a) provided neural evidence for the signal-suppression account. Specifically, they focused on inhibition-related component of the event-related potential (ERP) signal called distractor positivity (P_{D}; Hickey, Di Lollo, & McDonald, 2009). It has been shown that stimuli that fail to capture attention elicit a P_{D} component (e.g., Burra & Kerzel, 2014; Eimer & Kiss, 2008; Feldmann-Wüstefeld, Uengoer, & Schubö, 2015; Gaspar & McDonald, 2014). Gaspar and Luck (2018a) showed that across participants, the magnitude of the suppression in a task as we used here was related to the magnitude of the P_{D}. This elegantly connected the behavioral suppression to the neural measures of the suppression. Even though these results are convincing, it should be noted that also in their ERP study, Gaspelin and Luck (2018a) used only heterogeneous search arrays of four elements, which as shown here, renders none of the elements salient enough to say anything about the suppression of salient distractors.

Recently Barras and Kerzel (2016) also investigated this issue and used eight instead of four items. Important for the present discussion, when using eight elements, there was no sign of a P_{D}, even when participants engaged in the feature search mode. It suggests that if the P_{D} is a marker of suppression, this can be found only in sparse displays in which none of the elements are salient enough to generate a saliency signal. In another more recent ERP study, Kerzel and Burra (2020) revisited this issue but now explicitly focused on small search displays. In fact, the displays were identical to those of Gaspelin et al. (2015) and Gaspelin and Luck (2018a). This study shows that the P_{D} does not represent distractor suppression but instead is the result of a paradoxical flip of the contralateral voltage difference because of a peculiar search strategy that is used in small displays (Kerzel et al., 2020).

The current study indicates that the relative salience of items in the display is crucial in the capture-suppression debate. With a few heterogeneous items on display, it is possible to obtain top-down suppression (Chang et al., 2019; Gaspelin et al., 2015; Gaspelin & Luck, 2018a). However, in displays in which there is enough local feature contrast rendering both target and distractor singletons truly salient, no top-down suppression is observed. Instead, and consistent with the stimulus-driven account, the most salient item in the display captures attention.

In summary, consistent with Gaspelin et al. (2015; Gaspelin & Luck, 2018a), when engaged in feature search and the display consists of only four elements, there is below-baseline suppression of the singleton distractor. However, adding elements to the display renders both target and distractor singleton salient, resulting in attentional capture by the distractor, even when participants engage in the feature search mode. We conclude that signal-suppression account is limited in scope and cannot resolve the debate between stimulus-driven and goal-driven theories. If anything, the below-baseline suppression observed by Gaspelin et al. (2015) is the result of some idiosyncratic (most likely serial) search strategy that can operate only in displays containing a limited number of nonsalient elements.
References


Received March 7, 2020
Revision received May 1, 2020
Accepted May 2, 2020