

# Dimension-specific signal modulation in visual search: Evidence from inter-stimulus surround suppression

Louis K. H. Chan

Psychology Unit, Hong Kong Baptist University,  
Hong Kong



William G. Hayward

Department of Psychology, University of Hong Kong,  
Hong Kong



A fundamental task for the visual system is to determine where to attend next. In general, attention is guided by visual saliency. Computational models suggest that saliency values are estimated through an iterative process in which each visual item suppresses each other item's saliency, especially for those with close proximity. To investigate this proposal, we tested the effect of two salient distractors on visual search for a size target. While fixing the target-to-distractor distance, we manipulated the distance between two distractors. If two salient distractors suppressed each other when they were close together, they should interfere with search less; this was exactly what we found. However, we observed such a distance effect only for distractors of the same dimension (e.g., both defined in color) but not for those of different dimensions (e.g., one defined in color and the other in shape), displaying specificity to a perceptual dimension. Therefore, we conclude that saliency in visual search is calculated through a surround suppression process that occurs at a dimension-specific level.

Keywords: signal modulation, visual search, surround suppression

Citation: Chan, L. K. H., & Hayward, W. G. (2012). Dimension-specific signal modulation in visual search: Evidence from inter-stimulus surround suppression. *Journal of Vision*, 12(4):10, 1–10, <http://www.journalofvision.org/content/12/4/10>, doi:10.1167/12.4.10.

## Introduction

Information from the outside world is continually entering our minds, but we do not have the capacity to fully process all of it. As such, attention's role is to intelligently allocate mental resources, so that inputs that are of higher potential importance are processed with priority. Top-down and bottom-up factors combine to determine this process, so that locations that are known to be important, or perceptual events that are unusual, receive focal attention. This study investigates the mechanism by which perceptual inputs are converted into a guide for allocating focal attention.

Traditional models of visual search (e.g., Treisman & Sato, 1990; Wolfe, 1994) propose a two-stage architecture as the basis for search performance. The initial stage is that sensory properties are initially registered in dimension-specific channels (e.g., color, orientation); these signals then converge on a *saliency map* that codes the presence of relevant visual information and represents the priority in which each spatial location should be focally attended (Duncan & Humphreys, 1989; Itti & Koch, 2000; Wolfe, 1994). The proposal that bottom-up saliency guides attention is supported by the phenomenon of attentional capture, which is exemplified in the additional singleton paradigm (Theeuwes, 1991, 1992). In this paradigm, observers search for a target among distractors, during which time a singleton distractor may appear

(hereafter, we refer to a singleton distractor as a “singleton” only). Both the target and the singleton have a unique feature, and so both are salient. Importantly, when (and only when) the singleton is more salient than the target, its presence slows search. This shows that attention is preferentially allocated to the most salient objects in a visual scene.

Many proposals from the psychological literature suggest that perceptual contrasts (Nothdurft, 1991, 1992) underlie calculations of saliency levels (e.g., Treisman & Sato, 1990; Wolfe, 1994). Although this proposal may suffice for explaining the effect of a singleton on search performance, it may be problematic in situations where local contrasts are abundant across a scene, since it seems implausible that all these contrasts could attract focal attention. Biased competition models (e.g., Duncan & Humphreys, 1989) represent a slightly different case, because through competition the purpose of normalization may be achieved. Such models propose that competition occurs between object representations, so that a salient shape oddball can compete with a color oddball. Another group of models, primarily motivated by known neurophysiology (Huang & Grossberg, 2010; Itti & Koch, 2000), addresses this problem differently and proposes a normalization step between the contrast and saliency representations, so that in no case are saliency values high across all locations. This normalization step may take the form of center-surround modulation in biological systems. It is generally suggested that normalization occurs at a

dimension-specific stage, so that whereas same-dimension visual features suppress each other, different-dimension features do not.

An analysis of existing evidence appears to suggest that interaction between stimuli occurs primarily within a dimension-specific level. For example, Treisman (1988) showed that only within-dimension distractor heterogeneity interfered with target detection, but across-dimension heterogeneity did not. Similarly, attentional capture is usually much larger when the singleton is defined within the target dimension than within another dimension (e.g., Kumada, 1999). The general observation is that visual search performance is much more affected by stimulus variations within a single dimension. Consistent with this suggestion, Krummenacher, Müller, and Heller (2002) found that a dual-target advantage exceeded a race condition only when the two targets were near and were defined across different dimensions; it did not when they were defined within the same dimension. Violation of a race condition means that the response to the two stimuli presented together is faster than would be expected by taking the faster responses to either stimulus, indicating summative processing. Their interpretation of this result is that whereas two activation peaks of different dimensional origins add together on the saliency map, two peaks from the same dimension have first modulated each other before they arrive at the saliency map stage.

The aim of the present study was to gather evidence on the normalization process that affects saliency calculation. Our hypothesis is that such a process takes place at a dimension-specific stage in the form of center-surround modulation, so that high-contrast visual features suppress their neighbors on a dimensional contrast map. To formulate a test for visual saliency, we made use of Theeuwes' (1991, 1992) additional-singleton paradigm. It is known that a distractor interferes with search only if it is more salient than the target. By extending this rationale, we examined the effect of surround suppression on distractor saliency by looking at the extent to which the distractors interfered with search.

In the experiments reported in this study, we made use of a variation of the additional singleton paradigm, in which two “singleton” distractors were used. We measured the size of attentional capture when we varied the distance between the two singletons. According to surround suppression, when two singletons defined by the same dimension are close together, they should suppress each other strongly. This should lead to a reduced capture. In contrast, when two singletons are far apart or are of different dimensions, they should produce a stronger capture. In [Experiment 1](#), we examined variations in inter-singleton distance with two same-dimension singletons. In [Experiment 2](#), we examined the same effect for two different-dimension singletons. Both experiments measured response time (RT). In [Experiment 3](#), we replicated the tests by measuring accuracy within a brief display experiment.

## Experiment 1

In [Experiment 1](#), effects of inter-singleton distance on the size of attentional capture were investigated. We manipulated the distance between two singletons (close, intermediate, or far), defined by unique colors or unique shapes. We studied the case for same-dimension singletons in this experiment, such that two singletons were either a diamond and a pentagon or a red and a green circle. We used a size target (a large circle), so that both singleton pairs were across-dimension singletons. In the visual search task, participants had to first detect the target and then report the orientation of a line inside it by making a button press. Target-singleton distance was held constant (refer to [Figure 1](#) for an illustration).

## Methods

### Participants

Twenty members of the University of Hong Kong community (15 females) participated in [Experiment 1](#). Data from one participant were replaced due to a high error rate, and another participant was replaced due to exceptionally slow RTs. All reported normal or corrected-to-normal vision. Each of them was paid HK\$20 for a 20-min session.

### Stimuli

Stimuli were presented on a 20" CRT monitor with effective viewing dimensions of 39.1 cm by 29.3 cm. Viewing distance was 70 cm, fixed using a chin rest. Upon a dark background, distractors were circles of diameter  $2.02^\circ$ , border thickness  $0.12^\circ$ , and a light gray color. The line inside was  $1^\circ$  long and was colored gray. The target was a 130% scaled version of a distractor stimulus. If a pair of singletons was present, they were either a diamond and a pentagon of diagonal length  $2.53^\circ$  and a pentagon of a height  $2.10^\circ$  or a red and a green circle of the same size as other



Figure 1. Illustrations of search displays. The left panel illustrates search displays in [Experiment 1](#) and in same-dimension singleton trials of [Experiment 3](#). The right panel illustrates search displays in [Experiment 2](#) and in different-dimension singleton trials of [Experiment 3](#). A dark color depicts a red (or green) color. Figures not drawn in scale. Refer to text for details.

distractors. For the target, the line inside was either vertical or horizontal; for the singleton and the distractors, the line inside was  $22.5^\circ$  rotated either clockwise or anti-clockwise from either the vertical or the horizontal. The fixation dot was a  $0.3^\circ$  diameter circle filled in white. The set size was always 12. Search items were randomly positioned within two  $1.6^\circ$  apart,  $15.2^\circ$  wide,  $22.8^\circ$  tall regions, one on each side, with 6 items in each region. There was a minimum edge-to-edge distance of  $0.8^\circ$  between each item. The target–singleton distance was fixed at  $10^\circ$ , and the inter-singleton distance could be  $4^\circ$ ,  $8^\circ$ , or  $12^\circ$ , according to the experimental conditions. The target and the singleton were always located within the same region (i.e., same hemifield), because previous studies showed that distractor interference was stronger at the same side of the target (e.g., Caputo & Guerra, 1998; Mounts, 2000a, 2000b; Mounts & Gavett, 2004).

### Procedure

At the beginning of each trial, a fixation dot appeared for 1000 ms, followed by the search display. The search display contained a size target (a larger circle); participants detected it and responded to the orientation of a line inside it. The left shift key was pressed for a vertical line and the right shift key was pressed for a horizontal line. Responses were made as quickly and as accurately as possible. If the participant made an error, a 1000-Hz sine tone was played for 200 ms. Each trial concluded with a 500-ms inter-trial interval.

In two fifths of the trials, there was no singleton distractor in the search array; in the other trials, two singletons were close, intermediate, or far apart from each other (one fifth of the trials each). These trials were mixed within blocks. Each singleton pair (diamond and pentagon or red and green) was used in alternate blocks, counter-balanced across participants. Each participant performed 11 blocks of 45 trials. The first block served as a practice block and was not analyzed. The first 5 trials in each block were unanalyzed warm-up trials.

## Results and discussion

Trials with RTs shorter than 200 ms and the slowest 1% of trials were excluded from the analysis, as were incorrect responses. Error rates were low (2.3% of the trials) and showed no trend toward a speed–accuracy trade-off. RTs are plotted in Figure 2.

A preliminary analysis that included singleton dimension (two shape distractors vs. two color distractors) as a variable showed neither a main effect,  $F(1,19) = 0.88$ ,  $p > 0.3$ , nor an interaction effect with inter-singleton distance,  $F(3,57) = 1.51$ ,  $p > 0.2$ . This variable was therefore excluded from subsequent analyses. In this experiment, two same-dimension singletons were presented at different distances from each other. Planned contrasts showed no

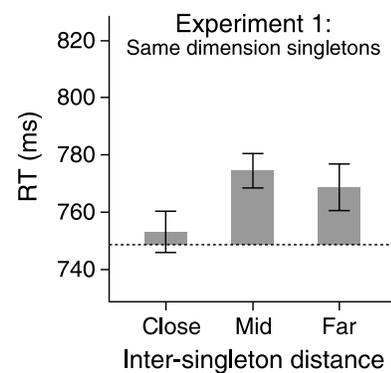


Figure 2. Results of Experiment 1 in terms of RTs. Data bars show RT interference caused by singletons. Error bars show  $\pm 1$  standard error of the mean.

significant attentional capture for a close inter-singleton distance,  $F(1,19) = 0.31$ ,  $p > 0.5$ , but significant attentional capture was found at intermediate,  $F(1,19) = 15.05$ ,  $p < 0.01$ , and far distances,  $F(1,19) = 5.25$ ,  $p < 0.04$ . By subtracting RT with each inter-singleton distance from that of singleton-absent trials, an analysis of variance was performed on capture size. This analysis showed a significant distance effect,  $F(2,38) = 3.34$ ,  $p < 0.05$ . Planned contrasts showed that attentional capture was weaker at a close distance than at an intermediate distance,  $F(1,19) = 7.71$ ,  $p < 0.02$ , and was marginally weaker than that at a far distance,  $F(1,19) = 2.45$ ,  $p = 0.13$ .

The present results are best explained by assuming surround suppression between visual items. In this case, nearby singletons suppressed each other, resulting in lower saliency and reduced attentional capture. Since the target–singleton distance was held constant, the inter-singleton distance effect must reflect mutual interaction between two singletons. A potential issue was that when two singletons were near, they might be perceptually grouped and rejected together, resulting in a weaker capture. It means that the present data might not reflect a surround suppression process. We discuss this possibility in the General discussion section.

## Experiment 2

In Experiment 2, we examined the distance effect between two different-dimension singletons. The experiment was identical to Experiment 1 except for the pairing of singletons. For instance, a red singleton was now paired with a diamond, and a green singleton was paired with a pentagon. If surround suppression had a dimension-specific locus, then it should not occur across dimensions, and as such, attentional capture would be expected at all inter-singleton distances, and its size should not show any distance effects.

## Methods

### Participants

Fourteen members of the University of Hong Kong community (10 females) participated in [Experiment 2](#). Data from one participant was replaced due to a high error rate. All reported normal or corrected-to-normal vision. Each of them was paid HK\$20 for a 20-min session.

### Stimuli and procedure

The stimuli and procedure were identical to [Experiment 1](#) except that a red singleton was now paired with a diamond and a green singleton was paired with a pentagon.

## Results and discussion

Trials were screened as in [Experiment 1](#). Error rates were low (3.0% of the trials). Planned contrasts showed significant attentional capture at all inter-singleton distances, close:  $F(1,13) = 8.44$ ,  $p < 0.02$ , intermediate:  $F(1,13) = 25.72$ ,  $p < 0.001$ , far:  $F(1,13) = 12.38$ ,  $p < 0.01$ . An analysis of variance showed an insignificant distance effect on capture size,  $F(2,26) = 1.88$ ,  $p = 0.17$ . Planned contrasts showed that a close singleton yielded faster RT than an intermediate one,  $F(1,13) = 5.57$ ,  $p < 0.04$ , but it was associated with less accurate responses,  $F(1,13) = 5.83$ ,  $p < 0.04$ , indicative of a speed–accuracy trade-off. No other RT comparisons reached significance (all  $p > 0.3$ ).

Since the short range effect in this experiment was contaminated by a speed–accuracy trade-off, it is not clear whether mutual suppression has occurred across dimensions. In order to compensate for the speed–accuracy trade-off, we analyzed data of [Experiment 2](#) in terms of inverse efficiency, which is computed as RT divided by accuracy (Townsend & Ashby, 1983). Inverse efficiency data are plotted in [Figure 3](#). Consistent with analyses based on RT, significant attentional capture was observed

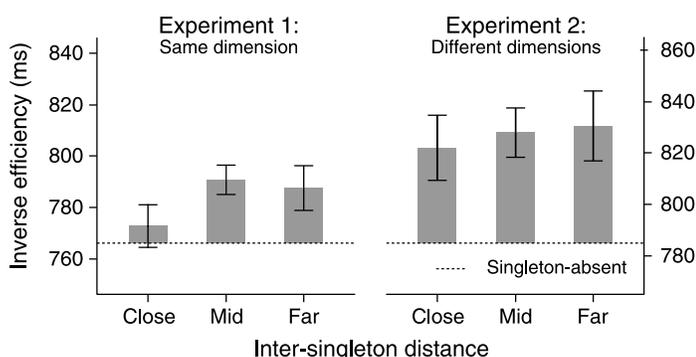


Figure 3. Results of [Experiments 1](#) and [2](#) in terms of inverse efficiency. Data bars show inverse efficiency impairments caused by singletons. Error bars show  $\pm 1$  standard error of the mean.

at all inter-singleton distances, close:  $F(1,13) = 7.26$ ,  $p < 0.02$ , intermediate:  $F(1,13) = 17.00$ ,  $p < 0.01$ , far:  $F(1,13) = 10.00$ ,  $p < 0.01$ . There is no longer any evidence for any distance effect on capture size,  $F(2,26) = 0.30$ ,  $p > 0.7$ , and capture size at intermediate and far inter-singleton distances was not larger than that at a close distance, intermediate:  $F(1,13) = 0.31$ ,  $p > 0.5$ , far:  $F(1,13) = 0.44$ ,  $p > 0.5$ . Importantly, as it can be seen in [Figure 3](#), a sharp decrease in capture size was absent when two singletons came near from an intermediate distance. This result stands in contrast to the case for [Experiment 1](#), in which two near singletons were associated with a significant drop in capture size in terms of both RT and inverse efficiency (both  $p < 0.04$ ; see [footnote 1](#) for the data analysis of [Experiment 1](#) in terms of inverse efficiency), as compared to two singletons further apart. Therefore, it appears to signal a qualitative difference between the inter-item modulation processes within and across dimensions.

## Experiment 3

Although a difference in the degree of distance effects between [Experiments 1](#) and [2](#) was apparent, the existence of a speed–accuracy trade-off in [Experiment 2](#) potentially threatened the reliability of this result. The aim of [Experiment 3](#) was therefore to determine whether differential distance effects would be observable when speed–accuracy trade-off was carefully avoided. To achieve this purpose, we measured response accuracy under a brief display situation. We determined the stimulus exposure by using a staircase procedure in a calibration block for each participant, so that accuracy was set at a suitable level.

To make possible a statistical comparison between distance effects within and across dimensions, in [Experiment 3](#), we included both same-dimension and different-dimension singleton conditions inside a within-subject manipulation. Therefore, we mixed both singleton conditions within blocks and measured the size of attentional capture at each inter-singleton distance.

## Methods

### Participants

Thirty-three members of the University of Hong Kong community (22 females) participated in [Experiment 3](#). All reported normal or corrected-to-normal vision. Each of them was paid HK\$40 for a 40-min session.

### Stimuli

The stimuli were based on [Experiments 1](#) and [2](#) with the following changes. First, the size of the entire search

display was reduced to 80% of the original, so that the entire search area had shrunk from two  $1.6^\circ$  apart,  $15.2^\circ \times 22.8^\circ$  regions to two  $1.3^\circ$  apart,  $12.2^\circ \times 18.3^\circ$  regions. All visual stimuli, as well as the target–singleton and inter-singleton distances, were scaled down by the same ratio. The display duration was set by using a staircase procedure as described in the [Procedure](#) section below. On each trial, the search display was followed by a mask display, which was composed of randomly oriented bars centered at each search item. There were eight bars centered at each item, in which four bars had a normal size (same size as those inside each distractor) and four bars had a larger size (130% scaled). We used mask bars of two different sizes for each item in order to ensure that they can effectively mask both distractor and target stimuli (because the targets have a size of 130%). These masking bars were white in color.

### Procedure

The procedure of this experiment was based on [Experiments 1 and 2](#). In this experiment, the fixation display was shortened to 500 ms. The search display exposure was set according to data obtained in a calibration block. The search screen was followed by a mask display, which appeared for 500 ms. A response was only allowed after the offset of the mask; this prevented quick, anticipation responses.

The first block of each experimental session was a calibration block. The procedure of each trial in this block was identical to that in other blocks except that the search display duration was adaptively set by a staircase routine. In other subsequent blocks, the duration was fixed. The staircase procedure started with a 720-ms display duration, a duration that was deemed easy by all participants. Before the first reversal, each correct response leads to a reduction of exposure duration (77.88% of the previous duration). After the first reversal, the staircase followed a 2-down, 1-up (for 21 participants) or a 3-down, 1-up (for 12 participants) adaptive procedure,<sup>2</sup> and the average exposure level of those trials after the sixth reversal was used for the rest of the experimental session. For the 21 participants that were given a 2-down, 1-up staircase, target eccentricity was constrained between  $6.4^\circ$  and  $12^\circ$ .

In 40% of the trials, there were no singleton distractors in the search array; in other trials, two singletons were either close, intermediate, or far apart; each distance constitutes 20% of the trials. Four singleton pairs, red–green, diamond–pentagon, red–diamond, and green–pentagon, were factorially crossed with each singleton condition (singleton-absent, close, intermediate, far), and these trials were mixed within blocks. Each participant performed 7 blocks of 85 trials. The first block served as the calibration block and was not analyzed. The first 5 trials in each block were unanalyzed warm-up trials.

## Results and discussion

Accuracy data are plotted in [Figure 4](#). First, for trials of same-dimension singletons, planned contrasts showed no attentional capture (in terms of accuracy drop from a singleton-absent condition) at a close inter-singleton distance,  $F(1,32) = 0.01$ ,  $p > 0.9$ , but capture was significant at an intermediate distance,  $F(1,32) = 6.67$ ,  $p < 0.02$ , and at a far distance,  $F(1,32) = 7.90$ ,  $p < 0.01$ . For trials of different-dimension singletons, capture was significant at a short distance,  $F(1,32) = 4.55$ ,  $p < 0.05$ , marginally significant at an intermediate distance,  $F(1,32) = 3.62$ ,  $p = 0.07$ , and significant at a far distance,  $F(1,32) = 12.99$ ,  $p < 0.01$ . This pattern of results mirrors those we observed in [Experiments 1 and 2](#).

Of primary interest is whether distance effects differed between same- and different-dimension trials. A two-way repeated-measures analysis of variance was performed with inter-singleton dimensional consistency (i.e., whether or not the singletons were of the same dimension) and inter-singleton distance as independent variables and capture size as the dependent variable. Results showed an insignificant interaction between dimensional consistency and distance,  $F(2,64) = 2.05$ ,  $p = 0.14$ , which showed that a simple main effect of inter-singleton distance was only found in same-dimension trials,  $F(2,64) = 3.61$ ,  $p < 0.04$ , but not in different-dimension trials,  $F(2,64) = 1.09$ ,  $p > 0.3$ . Planned contrasts showed consistent results with this analysis. For same-dimension trials, significantly weaker capture was observed at a close distance than at intermediate,  $F(1,32) = 5.4$ ,  $p < 0.03$ , and far distances,  $F(1,32) = 8.02$ ,  $p < 0.01$ . However, for different-dimension trials, capture was not reduced at a short distance, short vs. intermediate:  $F(1,32) = 0.17$ ,  $p > 0.6$ , short vs. far:  $F(1,32) = 0.90$ ,  $p > 0.3$ , suggesting that surround suppression was not occurring across dimensions.

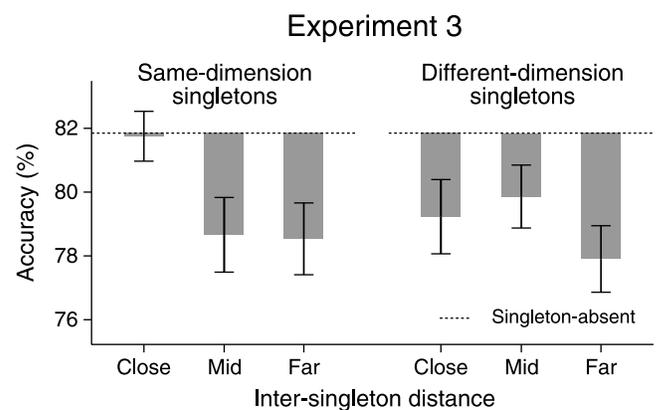


Figure 4. Results of [Experiment 3](#). Data bars show accuracy drop caused by singletons. Error bars show  $\pm 1$  standard error of the mean.

Although the interaction between dimensional consistency and inter-singleton distance was only marginally significant, further analysis showed that such an effect did reach a statistical threshold when only close and intermediate distances were taken into account,  $F(1,32) = 4.30$ ,  $p < 0.05$ . This strongly suggests that surround suppression has a dimension-specific locus at least for its short-range mechanism. In both [Experiments 2](#) and [3](#), a modest, insignificant trend of a further increase of capture size at a further distance was apparent. This could be indicative of a longer range mechanism that occurs at a dimension-general level or an artifact of random data sampling. The current data are indecisive to this question. However, we deem the former option unlikely because such a trend was not observed within a dimension ([Experiments 1](#) and [3](#)).

## General discussion

Prioritizing where to attend is one of the major goals of the visual system, and the calculation of visual saliency plays an important role in this process. One question regarding saliency calculation is how dimension-specific local contrasts are transformed into a dimension-general saliency representation, so that when perceptual contrasts are abundant across a scene, they could be regarded as non-salient. In this study, we propose that surround suppression between visual items can resolve this question as an implementation of a normalization process. We hypothesize that surround suppression acts on local contrast representations and thus has a dimension-specific locus.

The present results show two findings that are consistent with this hypothesis. First, two nearby visual items mutually suppress the saliency of each other. Nearness of two singletons reduced attentional capture to zero, in terms of RT ([Experiment 1](#)) and accuracy ([Experiment 3](#)). Second, mutual suppression is larger and is more robust within a perceptual dimension. Evidence for short-range surround suppression was only found when singletons were defined in the same dimension ([Experiments 1](#) and [3](#)), and short-range surround suppression was not found when singletons were defined in different dimensions ([Experiments 2](#) and [3](#)). Taken together, these experiments found evidence for surround suppression, which appeared to be dimension-specific.

These results point to a dimension-specific center-surround operation associated with each high-contrast feature. This finding is consistent with a saliency-based visual search mechanism that includes a dimension-specific normalization process (e.g., Itti & Koch, 2000; Koch & Ullman, 1985), which is carried out by a center-surround process. The hypothesized neural effect of this process is that while isolated high-contrast items would produce an activation peak, low-contrast and crowded

high-contrast items would result in a lower degree of activation. We reason that this conversion bridges the different computational goals for a feature map structure and a saliency map structure, in which the former serves to provide a perceptual representation of the current scene while the latter serves to provide an operational guide for deploying attention.

This interpretation of our results has important implications for visual search models. Let us consider Treisman's (1988) findings again to illustrate this point. She reported that distractor heterogeneity impaired target detection only within a dimension but not across dimensions. According to her original explanation, this result was due to (i) dimension-specific processing modules that operate largely independently and (ii) target detection that is driven by these modules. However, following a later phase of theoretical development, Treisman (Treisman & Sato, 1990) and other theorists (Itti & Koch, 2000; Wolfe, 1994) proposed that visual search is based on a saliency map that integrates information from all dimensions. This proposal is incompatible with the idea of module-specific target detection. For these later visual search theories to cope with Treisman's original data, one has to make additional assumptions, such as assuming top-down attention can select one dimension and filter away distractions from irrelevant dimensions. However, in a recent study, we have shown that none of these attempts were particularly successful (Chan & Hayward, 2012). For instance, neither a search-type explanation (Chan & Hayward, 2009), a dimension-weighting account (Müller, Heller, & Ziegler, 1995), nor appealing to high target saliency can resolve these theoretical limitations.

The current account of surround suppression however presents a solution to explain Treisman's (1988) data, even with the assumption that search is driven by a dimension-general saliency map. For instance, as heterogeneous distractors suppress each other at a dimension-specific stage, only a target that occurs within the same representational space would suffer from such suppression. As such, only within-dimension heterogeneity can have an effect before activations are integrated onto the saliency map. Within the saliency map, it is possible for across-dimension signals to interfere with the target. However, mutual suppression between distractors at an earlier stage should have already attenuated these signals to a lower degree before their arrival. This explains why across-dimension heterogeneity could only produce modest impairment to visual search, and the target from another dimension enters the saliency map without suffering from much interference.

At the same time, it presents a different case for attentional capture, in which interference is also present across dimensions (e.g., Theeuwes, 1992). In attentional capture, the capture distractor is generally the only oddball (or one of a few oddballs) within its dimension, with the other distractors being homogeneous. There is no other signal suppressing its activation. Therefore, the

capture distractor can always arrive at the saliency map with a considerable activation strength, exerting an influence on search priority.

Our proposal of dimension-specific spatial inhibition may be compared with recent literature that has identified a center-surround profile for attentional selection. Researchers have demonstrated that a distractor impairs target identification more as it moves closer to the target (Caputo & Guerra, 1998; Mathôt, Hickey, & Theeuwes, 2010; Mounts, 2000b), that visual search for a target shows center facilitation and surround inhibition as measured by a stimulus probe (Cave & Zimmerman, 1997; Mounts, 2000a), and that a distance effect occurs for identification of two targets, in which performance is lower at a closer distance (Bahcall & Kowler, 1999; Wei, Lü, Müller, & Zhou, 2008). Common to these observations is a target-driven surround suppression, following selection of a target, a distractor, a probe, or another target. This phenomenon is generally explained in terms of attentional selection of a target being achieved through suppression of the neural responses of its nearby distractors (Luck, Girelli, McDermott, & Ford, 1997). Physiological evidence is in line with this suggestion (Chelazzi, Duncan, Miller, & Desimone, 1998; Chelazzi, Miller, Duncan, & Desimone, 1993; Luck, Chelazzi, Hillyard, & Desimone, 1997; Moran & Desimone, 1985). To what extent do the present findings reflect the same attentional phenomenon as in these reports?

Whereas the similarity between our results and those of previous reports is that both have shown a surround suppression effect, a major difference is that we found surround suppression between two non-targets (Experiments 1 and 3). This effect cannot come from spatial suppression around a selected target. However, although a distractor may successfully capture attention, full processing of a visual item seems to be limited to attentional selection of the target (Remington & Folk, 2001). Furthermore, Cave and Zimmerman (1997) showed that the degree to which target-driven suppression occurs is correlated with the confusability between the target and distractors, indicating that attentional suppression is engaged upon demand. Therefore, target-driven suppression seems to be activated upon attentional selection rather than occurring for any item upon which attention might dwell only briefly (such as salient distractors). In terms of our findings, it seems unlikely that the observers selected the singletons, as they were highly dissimilar to the target. In addition, the capture caused by singletons was smaller than 50 ms, which matches the attentional dwell time commonly found for distractor rejections (Treisman & Gelade, 1980; Wolfe, 1994; Wolfe, Cave, & Franzel, 1989). Attentional dwell time associated with target processing is usually longer than 200 ms (e.g., Duncan, Ward, & Shapiro, 1994; Moore, Egeth, Berglan, & Luck, 1996; Theeuwes, Godijn, & Pratt, 2004).<sup>3</sup> Taken together, it seems that the surround suppression we report here reflects different mechanisms from the previously

reported phenomenon of target-driven suppression. Instead, it has a stimulus-driven locus, it occurs for each instance of a visual feature, and its suppressive strength is determined by the local contrast level of the corresponding visual feature.

The relationship between the present findings and the previous literature regarding an attentional surround can be embraced within a biased competition view of visual attention. According to this view (Desimone & Duncan, 1995; Duncan & Humphreys, 1989; Luck, Girelli et al., 1997), visual objects compete with each other for access to visual awareness and memory. Whereas visual saliency of the objects would determine the competition outcome, the role of attention is to bias the competition in a top-down fashion so as to favor certain task-relevant attributes or spatial locations. Here, an attentional surround can be considered as a mechanism with which such bias is set. However, it should be noted that competition should still occur even for visual objects that receive no top-down attentional bias. In light of the present results, this competition may also occur in the form of suppressive surrounds.

The present conclusion can also be presented within a functional architecture based on the guided search model and the dimension-weighting account (Müller et al., 1995; Wolfe, 1994). According to this framework, visual search is performed according to attentional priority set by a saliency map that receives inputs from each dimension-specific feature map coding local feature contrasts. The dimension-weighting account suggests that when feature map signals are aggregated onto the saliency map, signals are first weighted according to top-down needs as well as search history. Within this framework, the present addition is a mutual suppression mechanism that operates on each feature map, such that high-contrast signals that are close together inhibit each other's strength. This idea is consistent with one formulation of the dimension-weighting account (Krummenacher et al., 2002, p. 1311) and makes explicit the reason why signal coactivation occurs only across dimensions but not within a dimension (Krummenacher et al., 2002). For instance, signals from different feature maps add together on the saliency map, while signals on the same map suppress each other. Considering surround suppression within the dimension-weighting account also clarifies the distinction between a target-driven attentional surround and the stimulus-driven surround suppression process. For instance, whereas a target-driven attentional surround is a consequence of attentional selection (the saliency map only acts to prioritize attention), stimulus-driven surround suppression is a concurrent preattentive process that occurs on the feature maps.

In the present study, the main finding supporting the surround suppression account is that attentional capture was reduced when two singletons were of the same dimension and were near each other. It is important to consider if any alternative explanation exists, which can

threaten the current interpretation. One such explanation can be drawn from perceptual grouping, in which two near singletons may be perceptually grouped and thus are rejected together, resulting in reduced attentional capture. We deem this account unlikely for three reasons. First, the singleton pairs that we used were highly dissimilar and so are unlikely to group. Second, in the same-dimension conditions (Experiments 1 and 3), even though the color and orientation singleton pairs appear to have different degrees of similarity (and so should form groups of different strength), we did not observe different attentional capture size,  $F(3,57) = 1.13$ ,  $p > 0.3$ . Third, even though grouping two singletons may ease rejection of them, the group itself should still capture attention. In the present results, however, capture was virtually non-existent in the critical condition. Therefore, it is clear that a perceptual grouping account is inconsistent with and cannot fully account for the current results.

Taken together, while we acknowledge the previous findings regarding a center-surround profile of attention, we suggest that there is both a theoretical need and an empirical basis to assume an additional stimulus-driven center-surround modulation, which is associated with each high-contrast feature instance. This modulation resides at a dimension-specific level. The functional role of this process is to derive a visual representation that serves to guide attention. By transforming a perceptual representation (feature contrasts) into a functional representation (visual saliency), attention can be guided to important locations without facing a potential computational problem if many locations seem interesting. Thus, the intelligent allocation of mental resources is made possible.

## Acknowledgments

This research was supported by a grant from the Hong Kong Research Grants Council (HKU744209H) to WGH. We thank Ming Hon Chu and Hoi Ying Chu for their assistance in carrying out the research project.

Commercial relationships: none.

Corresponding author: Louis K. H. Chan.

Email: clouis@graduate.hku.hk.

Address: Psychology Unit, Hong Kong Baptist University, 8 On Muk Street, Shek Mun, Shatin, Hong Kong.

## Footnotes

<sup>1</sup>We reanalyzed data of Experiment 1 in terms of inverse efficiency, and the results resembled those of the original RT analysis. In this analysis, planned contrasts showed no attentional capture for two close singletons,  $F(1,19) = 0.56$ ,  $p > 0.4$ , but showed significant capture for

two singletons at intermediate and far distances,  $F(1,19) = 15.75$ ,  $p < 0.01$ ;  $F(1,19) = 5.40$ ,  $p < 0.04$ . Analysis of capture size showed a marginally significant distance effect,  $F(2,38) = 2.30$ ,  $p = 0.11$ . Planned contrasts showed that attentional capture was weaker at a close distance than at an intermediate distance,  $F(1,19) = 5.03$ ,  $p < 0.04$ , and was insignificantly weaker than at a far distance,  $F(1,19) = 1.89$ ,  $p = 0.19$ . Inverse efficiency data are plotted in Figure 3.

<sup>2</sup>The reason for using two staircase settings: we started with the intention to set accuracy near 80% and used a 3-down, 1-up staircase in the calibration block. Presumably due to learning, however, it ended up with an average accuracy of 90.9% for 12 participants. Therefore, we continued our experiment with a 2-down, 1-up staircase, which resulted in an average accuracy of 73.7%. Since the choice of the staircases does not harm the validity of the data, we included data from both settings in our analysis. A mixed-design analysis of variance showed no evidence for any interaction between staircase settings and singleton conditions,  $F(6,186) = 0.61$ ,  $p > 0.7$ ; planned contrasts also showed no difference between capture size of each staircase setting at each singleton condition, all  $p > 0.2$ . Therefore, we pooled data of both staircases together in our main analysis in order to maximize sensitivity.

<sup>3</sup>As suggested by Clayton Hickey, an alternative reason for why capture effects are small here is that the target and distractor were reasonably far from one another. For instance, it was shown that capture effects are larger when target and distractor were close, and this observation is thought to reflect receptive field overlap (Hickey & Theeuwes, 2011). We agree with this description. However, our argument is that even without receptive field overlap (or the degree of receptive field overlap is controlled), attention generally dwells much longer at a target than at an attention-capturing distractor. We speculate that this difference reflects a distinction in selection and attention. Here, we want to argue that while selection triggers target-driven suppression, a distractor does not.

## References

- Bahcall, D. O., & Kowler, E. (1999). Attentional interference at small spatial separations. *Vision Research*, 39, 71–86.
- Caputo, G., & Guerra, S. (1998). Attentional selection by distractor suppression. *Vision Research*, 38, 669–689.
- Cave, K. R., & Zimmerman, J. M. (1997). Flexibility in spatial attention before and after practice. *Psychological Science*, 8, 399–403.
- Chan, L. K. H., & Hayward, W. G. (2009). Feature integration theory revisited: Dissociating feature detection and attentional guidance in visual search. *Journal of Experimental Psychology: Human Perception and Performance*, 35, 119–132.

- Chan, L. K. H., & Hayward, W. G. (2012). *Dimension-specific signal modulation in visual search: Evidence from visual search*. Manuscript in preparation.
- Chelazzi, L., Duncan, J., Miller, E. K., & Desimone, R. (1998). Responses of neurons in inferior temporal cortex during memory-guided visual search. *Journal of Neurophysiology*, *80*, 2918–2940.
- Chelazzi, L., Miller, E. K., Duncan, J., & Desimone, R. (1993). A neural basis for visual search in inferior temporal cortex. *Nature*, *363*, 345–347.
- Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, *18*, 193–222.
- Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological Review*, *96*, 433–458.
- Duncan, J., Ward, R., & Shapiro, K. (1994). Direct measurement of attentional dwell time in human vision. *Nature*, *369*, 313–315.
- Hickey, C., & Theeuwes, J. (2011). Context and competition in the capture of visual attention. *Attention, Perception, & Psychophysics*, *73*, 2053–2064.
- Huang, T.-R., & Grossberg, S. (2010). Cortical dynamics of contextually-cued attentive visual learning and search: Spatial and object evidence accumulation. *Psychological Review*, *117*, 1080–112.
- Itti, L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research*, *40*, 1489–1506.
- Krummenacher, J., Müller, H. J., & Heller, D. (2002). Visual search for dimensionally redundant pop-out targets: Parallel-coactive processing of dimensions is location specific. *Journal of Experimental Psychology: Human Perception and Performance*, *28*, 1303–1322.
- Kumada, T. (1999). Limitations in attending to a feature value for overriding stimulus-driven interference. *Perception & Psychophysics*, *61*, 61–79.
- Luck, S. J., Chelazzi, L., Hillyard, S. A., & Desimone, R. (1997). Neural mechanisms of spatial selective attention in areas V1, V2, and V4 of macaque visual cortex. *Journal of Neurophysiology*, *77*, 24–42.
- Luck, S. J., Girelli, M., McDermott, M. T., & Ford, M. A. (1997). Bridging the gap between monkey neurophysiology and human perception: An ambiguity resolution theory of visual selective attention. *Cognitive Psychology*, *33*, 64–87.
- Mathôt, S., Hickey, C., & Theeuwes, J. (2010). From reorienting of attention to biased competition: Evidence from hemifield effects. *Attention, Perception, & Psychophysics*, *72*, 651–657.
- Moore, C. M., Egeth, H., Berglan, L. R., & Luck, S. J. (1996). Are attentional dwell times inconsistent with serial visual search? *Psychonomic Bulletin & Review*, *3*, 360–365.
- Moran, J., & Desimone, R. (1985). Selective attention gates visual processing in the extrastriate cortex. *Science*, *229*, 782–784.
- Mounts, J. R. W. (2000a). Evidence for suppressive mechanisms in attentional selection: Feature singletons produce inhibitory surrounds. *Perception & Psychophysics*, *62*, 969–983.
- Mounts, J. R. W. (2000b). Attentional capture by abrupt onsets and feature singletons produces inhibitory surrounds. *Perception & Psychophysics*, *62*, 1485–1493.
- Mounts, J. R. W., & Gavett, B. E. (2004). The role of salience in localized attentional interference. *Vision Research*, *44*, 1575–1588.
- Müller, H. J., Heller, D., & Ziegler, J. (1995). Visual search for singleton feature targets within and across feature dimensions. *Perception & Psychophysics*, *57*, 1–17.
- Nothdurft, H. C. (1991). Texture segmentation and pop-out from orientation contrast. *Vision Research*, *31*, 1073–1078.
- Nothdurft, H. C. (1992). Feature analysis and the role of similarity in preattentive vision. *Perception & Psychophysics*, *52*, 355–375.
- Remington, R. W., & Folk, C. L. (2001). A dissociation between attention and selection. *Psychological Science*, *12*, 511–515.
- Theeuwes, J. (1991). Cross-dimensional perceptual selectivity. *Perception & Psychophysics*, *50*, 184–193.
- Theeuwes, J. (1992). Perceptual selectivity for color and form. *Perception & Psychophysics*, *51*, 599–606.
- Theeuwes, J., Godijn, R., & Pratt, J. (2004). A new estimation of the duration of attentional dwell time. *Psychonomic Bulletin & Review*, *11*, 60–64.
- Townsend, J. T., & Ashby, F. G. (1983). *The stochastic modeling of elementary psychological processes*. Cambridge, UK: Cambridge University Press.
- Treisman, A. (1988). Features and objects: The fourteenth Bartlett memorial lecture. *Quarterly Journal of Experimental Psychology A: Human Experimental Psychology*, *40A*, 201–237.
- Treisman, A., & Sato, S. (1990). Conjunction search revisited. *Journal of Experimental Psychology: Human Perception and Performance*, *16*, 459–478.
- Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, *12*, 97–136.

- Wei, P., Lü, J., Müller, H. J., & Zhou, X. (2008). Searching for two feature singletons in the visual scene: The localized attentional interference effect. *Experimental Brain Research*, *185*, 175–188.
- Wolfe, J. M. (1994). Guided search 2.0: A revised model of visual search. *Psychonomic Bulletin & Review*, *1*, 202–238.
- Wolfe, J. M., Cave, K. R., & Franzel, S. L. (1989). Guided search: An alternative to the modified feature integration model for visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *15*, 419–433.