Variation in target and distractor heterogeneity impacts performance in the centroid task

Vivian T. Lu Department of Cognitive Sciences, University of California, Irvine, Irvine, California, USA

Charles E. Wright Department of Cognitive Sciences, University of California, Irvine, Irvine, California, USA

Charles Chubb Department of Cognitive Sciences, University of California, Irvine, Irvine, California, USA

George Sperling Department of Cognitive Sciences, University of California, Irvine, Irvine, California, USA

In a selective centroid task, the participant views a brief cloud of items of different types—some of which are targets, the others distractors—and strives to mouse-click the centroid of the target items, ignoring the distractors. Advantages of the centroid task are that multiple target types can appear in the same display and that influence functions, which estimate the weight of each stimulus type in the cloud on the perceived centroid for each participant, can be obtained easily and efficiently. Here we document the strong, negative impact on performance that results when the participant is instructed to attend to target dots that consist of two or more levels of a single feature dimension, even when those levels differ categorically from those of the distractor dots. The results also show a smaller, but still observable decrement in performance that results when there is heterogeneity in the distractor dots.

Introduction

When separated from a group of purple-shirted friends in a large crowd (think Main Street, USA at Disneyland), you might be able to locate the group as the center of a cluster of those purple-colored shirts. This is an example of feature-based attention, a mechanism that makes features of interest—in this example, the color purple—more salient, which in turn makes instances of the feature easier to identify or group together. What would happen, however, if the members of your group were wearing shirts that were varying shades of purple rather than all the same color? Perhaps in this case the process of locating the group might be more difficult even if no one else in the crowd were wearing a purple shirt. Similarly, if everyone else in the crowd were wearing yellow shirts, the localization task would presumable be much easier than if they were wearing a variety of colors. Duncan and Humphreys (1989) began the investigation of these effects in visual search, showing that both distractor heterogeneity and increased similarity between targets and distractors lead to reduced performance.

Increasing the similarity between targets and distractors, increasing target heterogeneity, and increasing distractor heterogeneity are all examples of manipulations that make it harder for observers to attend only to targets and to ignore distractors (Bravo & Nakayama, 1992; Maljkovic & Nakayama, 1994; Nagy & Thomas, 2003; Nagy, Neriani, & Young, 2005; Buetti, Cronin, Madison, Wang, & Lleras, 2016). This breakdown of feature-based selective attention is often understood using a filter analogy, in which attention functions like a filter that is preferentially selective for the features defining target versus distractor items. The more similar an item is to those features, the more salience it will have after passing through the attention filter; distractor items that are not similar to the target features should receive low salience. Using this analogy, task difficulty mirrors achievable filter selectivity.

Ideally, an attention filter should allow only target items to pass through (to subsequent processing) with high salience. However, in a task in which there is variation in both target and distractor items, the participant is unlikely to be able to achieve such an ideal attention filter. In this case, either of two possible

Citation: Lu, V. T., Wright, C. E., Chubb, C., & Sperling, G. (2019). Variation in target and distractor heterogeneity impacts performance in the centroid task. *Journal of Vision*, *19*(4):21, 1–13, https://doi.org/10.1167/19.4.21.

https://doi.org/10.1167/19.4.21 Received July 12, 2018; published April 18, 2019 ISSN 1534-7362

problems may emerge: Some types of distractor items may pass through with sufficient salience to alter subsequent computations, and some types of target items may receive lower salience than others.

The current study uses stimuli in which target and distractor items are small, equiluminant dots that differ from each other in hue but are all slightly brighter than the background gray. Previous studies (Sun, Chubb, Wright, & Sperling, 2016a) have documented that observers can achieve highly effective attention filters selective for specific hues. In this domain, filters that must capture heterogeneous targets require a broad hue *passband*—i.e., that region of the hue circle through which hues pass with high amplitude or, in our terms, high salience. Distractor heterogeneity requires broad, flat hue *stop bands*: the regions on either side of the passband through which hues pass with low amplitude/salience. Increasing target/distractor similarity requires that the filter have sharper transitions between the passband and the stop bands. This explanation arrives at the same conclusion as the studies already cited: We expect feature-based attention tasks to increase in difficulty to the extent that they require the participant to achieve attention filters with wide passbands, wide stop bands, and sharp transitions.

The purpose of the current experiment was to look for evidence that either supports or undermines the filter analogy of feature-based attention by observing how performance is affected by separate manipulations of both target and distractor heterogeneity within the centroid paradigm.

In the centroid paradigm, participants view a brief presentation of a cloud of items varying in one or more features and then use a mouse to indicate the perceived centroid (the center of mass) of all the target items, ignoring the distractor items (Sun et al., 2016b). When presented with a group of items, people naturally tend to find the center of mass of the group. For example, McGowan, Kowler, Sharma, and Chubb (1998), Baud-Bovy and Soechting (2001), Friedenberg and Liby (2002), and Drew, Chubb, and Sperling (2010) have found that participants are able to easily find the center of mass of a cloud of target items. As it is an automatic response that participants make, it seemed reasonable to use it as a measure of performance. Participants would not have to learn a new skill in order to perform the task, nor would the results be applicable only within the setting of the task. For this experiment, the stimulus cloud consisted of 12 items: six targets, each taking one of three distinct reddish hues, and six distractors, each taking one of three distinct greenish hues. For this research, an important advantage of the centroid paradigm is that it supports the efficient estimation of an influence function—an estimate of an observer's attention filter for a task—which is directly analogous to a filter characteristic and characterizes how well the observer was able to attend to each target type and ignore each distractor type. The centroid paradigm also allows us to study the effect of target heterogeneity within a trial instead of between trials. Previous visual-search experiments have been able to study target heterogeneity only by varying the target type across trials, since those tasks rely on the presence of only one or no target in each trial. Because the centroid paradigm allows us to have multiple tokens of the target and multiple target types within a trial, we are able to test the participants' selectivity for each target type when multiple types were present. If having heterogeneity in target types does affect performance, we should be able to observe it easily based on the attention filters estimated for each task. More importantly, we would expect an ideal observer to deploy the same filter for all tasks, since the set of target hues they were meant to attend to was the same across trials throughout the experiment. If this supposition holds, it should be reflected in the similarity of the influence functions across the tasks.

Each participant was tested in four different experimental conditions in which displays differed in target and distractor heterogeneity. Varying both target and distractor heterogeneity allowed us to study both the target passbands and the distractor stop bands. Any display in which target dots were heterogeneous contained two dots with each of the three reddish target hues; similarly, any display in which distractor dots were heterogeneous contained two dots with each of the three greenish distractor hues. Any display in which target dots were homogeneous contained six dots all of the same hue, randomly chosen from the three reddish target hues; similarly, any display in which target dots were homogeneous contained six dots all of the same hue, randomly chosen from the three greenish distractor hues. Because luminance for all dots was kept constant, the only feature that the participants could use to differentiate the dots from one another was the hue.

One reason for using hues instead of other features is that we know that participants are good at discriminating stimuli based on hue. In fact, they are better at finding the centroid of a target hue than of a target luminance or a target saturation (Sun et al., 2016b). In the past, we have not asked participants to find the centroid of multiple target hues, but we do expect them to perform well in this task. D'Zmura (1991) concluded that participants are able to easily distinguish between target and distractor hues if the hues lie on separate sides of a dividing line in color space. Furthermore, Bauer, Jolicoeur, and Cowan (1996) found that if the target hue is colinear with two distractor hues, search performance suffers, but it then improves as the distractor hues move farther away from the target hues. Because our target and distractor hues lie on opposite

	Target dots Homogenous	Target dots Heterogenous
Distractor dots Homogenous	Task T1D1: 1 target type 1 distractor type	Task T3D1: 3 target types 1 distractor type
Distractor dots Heterogenous	Task T1D3: 1 target type 3 distractor types	Task T3D3: 3 target types 3 distractor types

Table 1. Four tasks varied between blocks in the experiment.

sides of a dividing line in color space, and because the targets are not colinear with the distractors, we expect our participants to be able to perform well in this experiment. We will also be able to observe how well hues that lie closer to the divider are distinguished from one another as targets or distractors. We may find that as the distance in color space between target hues and distractor hues increases, participants will be able to distinguish them better, assign more weight to the targets, and perform better at the task.

The four task conditions are summarized in Table 1. In the T3D3 condition, target and distractor dots were both heterogeneous. In the T3D1 condition, target dots were heterogeneous and distractor dots were homogeneous. In the T1D3 condition, target dots were homogeneous and distractor dots were heterogeneous. And in the T1D1 condition, both target and distractor dots were homogeneous. An unusual aspect of the design of this experiment is that, although trial-to-trial target/distractor heterogeneity varied across the four tasks shown in Table 1, across the trials in a block the probabilities associated with target and distractor heterogeneity did not change. This is true because the color used for any set of homogeneous dots was randomly chosen from among the same set of three colors that appeared simultaneously in heterogeneous dot sets. Thus, until the stimulus appeared, all of the target and distractor colors were equally likely to appear on a trial in each task.

We chose this design because if, according to the analogy, attention filters are endogenous—that is, if they must be specified before the observer views a stimulus cloud—then the same filter should be used in both the T1D1 and T3D3 tasks. The key assumptions required for this to be true are that in a given task condition, the subject needs to use a fixed filter whose sensitivity to different types of items does not depend on whether those items actually occur in a given display, and this filter must remain fixed across the brief duration of a stimulus display in each of the experimental conditions tested. If selectivity varies across tasks, it would indicate that the simple filter analogy does not sufficiently describe how the attention mechanism operates. Including the T1D3 and T3D1 tasks should allow us assess the whether any breakdown in the overall comparison is due to heterogeneity in the targets or the distractors.

Methods

Participants

There were eight participants (four women, four men), whose ages ranged from 20 to 30 years. All participants had normal or corrected-to-normal vision. Four of the participants were chosen because they had extensive prior experience with the centroid task. The other four had little to no experience; they were either completely new to the task or had practiced it for only a few hours. The procedure was approved by the University of California, Irvine, Institutional Review Board.

Materials

Participants ran the experiment on a Mac computer (iMac10,1) running MATLAB's (MathWorks, Natick, MA) Psychtoolbox package (Brainard, 1997). The stimuli were displayed on an LED monitor—integrally part of the computer—with a resolution of $1{,}920 \times 1{,}080$ pixels. Stimuli were viewed from a distance of 70 cm.

Stimuli

The stimulus region was 640×640 pixels (visual angle: 12.85°), centered on the display. Dots were squares of 17×17 pixels (visual angle: 0.34°) whose locations on the display were drawn from a bivariate Gaussian distribution centered on the middle of the display area. The standard deviation of these locations was 100 pixels (visual angle: 2.00°). If any of the dots overlapped or fell outside of the display area, the whole stimulus was thrown out and regenerated. In all trials, the cloud of dots was presented for 150 ms then disappeared. After an additional 33 ms, a mask made of colored dots arranged in jittered rows and columns appeared. The mask consisted of 10 rows and 10 columns of dots, each colored with one of the six colors used in the experiment.

The six targets and six distractors were distinguished by their colors, which we selected from a set of eight equiluminant hues (the "Hue" set studied by Sun et al., 2016a) that were equally spaced around an ellipse in color space. Coordinates for each color in CIE 1931 xy color space are listed in Table 2. Colors 6, 7, and 8 were

Role	Color		x-coordinate	y-coordinate
Filler	1 Violet		0.2755	0.2264
	2 Blue		0.2586	0.3000
Distractors	3 Green		0.2787	0.3765
	4 Yellow-green		0.3493	0.4757
Filler	5 Yellow		0.4307	0.5026
	6 Orange		0.4308	0.4154
Targets	7 Red		0.3762	0.3218
	8 Magenta		0.3173	0.2707
Background	Gray		0.3257	0.3478

Table 2. An example for a set of CIE 1931 xy color-space values for the eight colors used for one subject. *Notes*: Colors 1 and 5 were omitted for this experiment. Colors in red ink are targets, whereas colors in green ink are distractors.

the target hues, whereas Colors 2, 3, and 4 were the distractors; Colors 1 and 5 were excluded to create a clear division between the targets and distractors. Red (Color 7) and green (Color 3) were chosen as the center hues in the target and distractor sets, respectively, because they are hues on the L–M axis in the DKL model of color space (Derrington, Lennie, & Krauskopf, 1983), commonly used in color research; and during pilot sessions, these were the colors for which participants exhibited the highest selectivity.

The actual colors for each participant were individually calibrated to be equiluminant to one another, so that hue, not luminance, would guide the selection process. The luminance to which these stimuli were matched was chosen to be slightly brighter (48.11 cd/m^2) than the gray background (47.87 cd/m^2), to aid in their precise spatial localization. Subjective equiluminance of the stimuli was achieved using a minimum-motion paradigm (Antis & Cavanagh, 1983; Lu & Sperling, 2001; Herrera et al., 2013). The hues themselves were chosen in the same ways as those used by Sun et al. (2016a), who demonstrated that the attention filters for selecting one hue from among the others were similarly selective for all eight hues in the color circle—i.e., the eight hues functioned equally well as targets.

Procedure

Participants were briefly presented with a cloud of dots and then asked to click on the center of mass of the target dots. The sequence and timing of the displays used in the experiment are presented in Figure 1.

The experiment was divided into the four tasks summarized in Table 1. In the T1D3 and T1D1 tasks (in which only one target hue was presented on a given trial), one of the three target hues (orange, red, and magenta; Colors 6, 7, and 8) was randomly chosen for all six target dots on a trial. Likewise, in the T3D1 and T1D1 tasks, one of the three distractor hues (blue, green, and yellow-green; Colors 2, 3, and 4) was randomly chosen for all six distractor dots on a trial (Figure 2). All blocks in each task consisted of 90 trials that were categorized as one of three trial types—singleton (one target dot present), target only (six target dots present), and full set (targets and distractors present). Task 1 had six blocks, Tasks 2 and 3 had eight blocks each, and Task 4 had 10 blocks. The number of blocks in each task varied due to the number of hues that appeared in each trial in each task. Because all six colors were presented in each trial in Task 1, participants were able to provide enough data to calculate sufficiently precise influence functions for each color in only six blocks. In the other tasks, not all hues were presented on every trial, so we needed to have participants complete more trials, spread out across more blocks, in order to have enough data to provide sufficient precision in the influence functions. Each participant completed each of the four tasks twice in a randomized order determined by a Latin square.

Analysis

To decompose the response errors, we followed the procedures described by Sun et al. (2016b) to derive an influence function *f* from each participant's data in

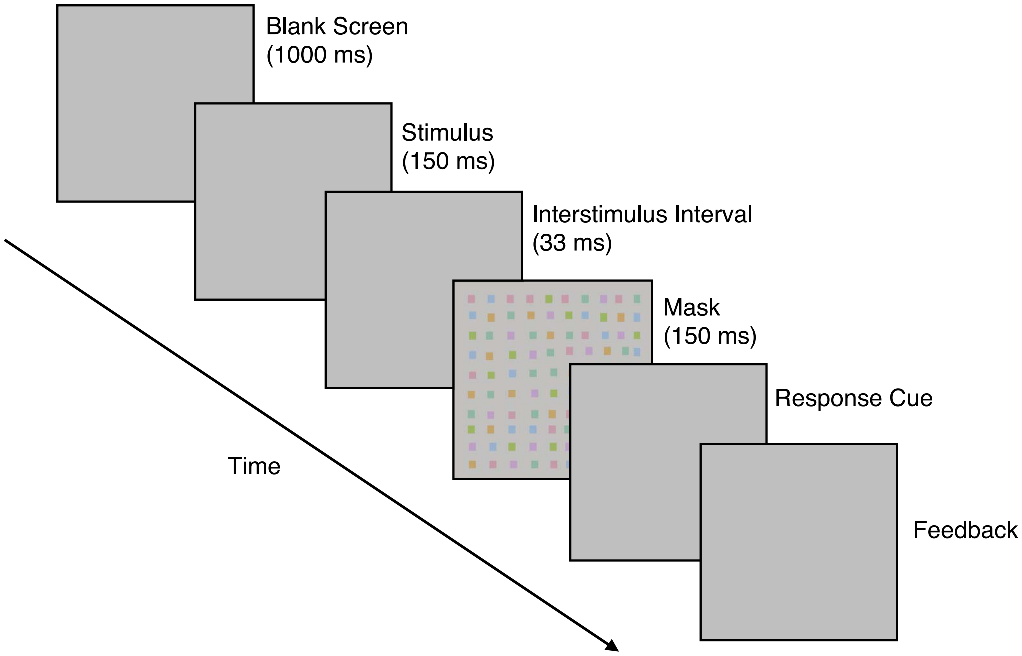

Figure 1. Sequence of display events on a trial. Stimulus displays were presented for 150 ms, followed by a blank screen for 33 ms and a mask for another 150 ms. Once the cross appeared, participants were able to move the mouse to adjust its position to the perceived centroid of the target dots, which was selected using a mouse click. Participants were shown a feedback screen that displayed the original stimulus cloud, a cross to show the location of the response, and a set of concentric circles to show the true location of the target centroid.

each condition. The first step in these analyses generates estimates of the observer's attention filter f_φ. An observer's attention filter is the vector of weights (one for each of the eight hues used in our stimuli) used by the observer when performing a task with a particular target filter φ. For tasks with a single target, the target filter takes the values 1 for the target hue and 0 for the distractor hue(s). For the tasks with three target hues, the target filter takes the values 1/3 for each of those hues and 0 for the distractor hue(s).

With target filter $\varphi(c)$, the correct response T on a given trial has x- and y-coordinates

$$T_x = \frac{\sum_i \varphi(C_i) x_i}{\sum_i \varphi(c_i)} \text{ and } T_y = \frac{\sum_i \varphi(c_i) y_i}{\sum_i \varphi(c_i)}, \quad (1)$$

where the sum is over all squares i in the display, C_i is the hue of square i, and x_i and y_i are the x- and y-coordinates of its location. Typically, however, the response of the observer deviates from this target location.

We assume that the x- and y-coordinates of the observer's response on trial t are given by

$$R_{t,x} = \mu_{t,x} + Q_{t,x} \text{ and } R_{t,y} = \mu_{t,y} + Q_{t,y}, \quad (2)$$

where $Q_{t,x}$ and $Q_{t,y}$ are independent, normally distributed random variables with mean 0 and some standard deviation σ, and for some function $f_\varphi(C)$ we have

$$\mu_{t,x} = \frac{\sum_i f_\varphi(C_{t,i}) x_{t,i}}{\sum_i f_\varphi(c_{t,i})} \text{ and } \mu_{t,y} = \frac{\sum_i f_\varphi(c_{t,i}) y_{t,i}}{\sum_i f_\varphi(c_{t,i})}. \quad (3)$$

In Equation 3, $h_{t,i}$, $x_{t,i}$, and $y_{t,i}$ are the hue and x- and y-coordinates of the ith square in the stimulus on trial t, and $f_\varphi(c)$ is the attention filter that the observer uses to perform the task.

The influence function $f(c)$ shows how much weight dots of hue c exerted on the responses produced by the participant in a condition and is therefore an estimate of the participant's attention filter for that particular hue. Additionally, we were able to characterize the performance of each participant in each condition by calculating two measures described by Sun et al. (2016b): selectivity ratio and efficiency.

The selectivity ratio summarizing the influence function f is defined as the sum of $f(c)$ across all target hues c divided by the mean of $|f(c)|$ across all distractor hues c. Taking the logarithm (base 10) of selectivity ratios is useful because the resulting scale is closer to equal interval. If the log10 selectivity ratio in a condition is 1, then its target hues have 10 times as much weight on the participant's response as the distractor hues do.

Efficiency reflects the proportion of dots that would need to be processed by an ideal observer (using the same influence function as the participant) to achieve the same level of response error as the participant. Specifically, the ideal observer is presented with the

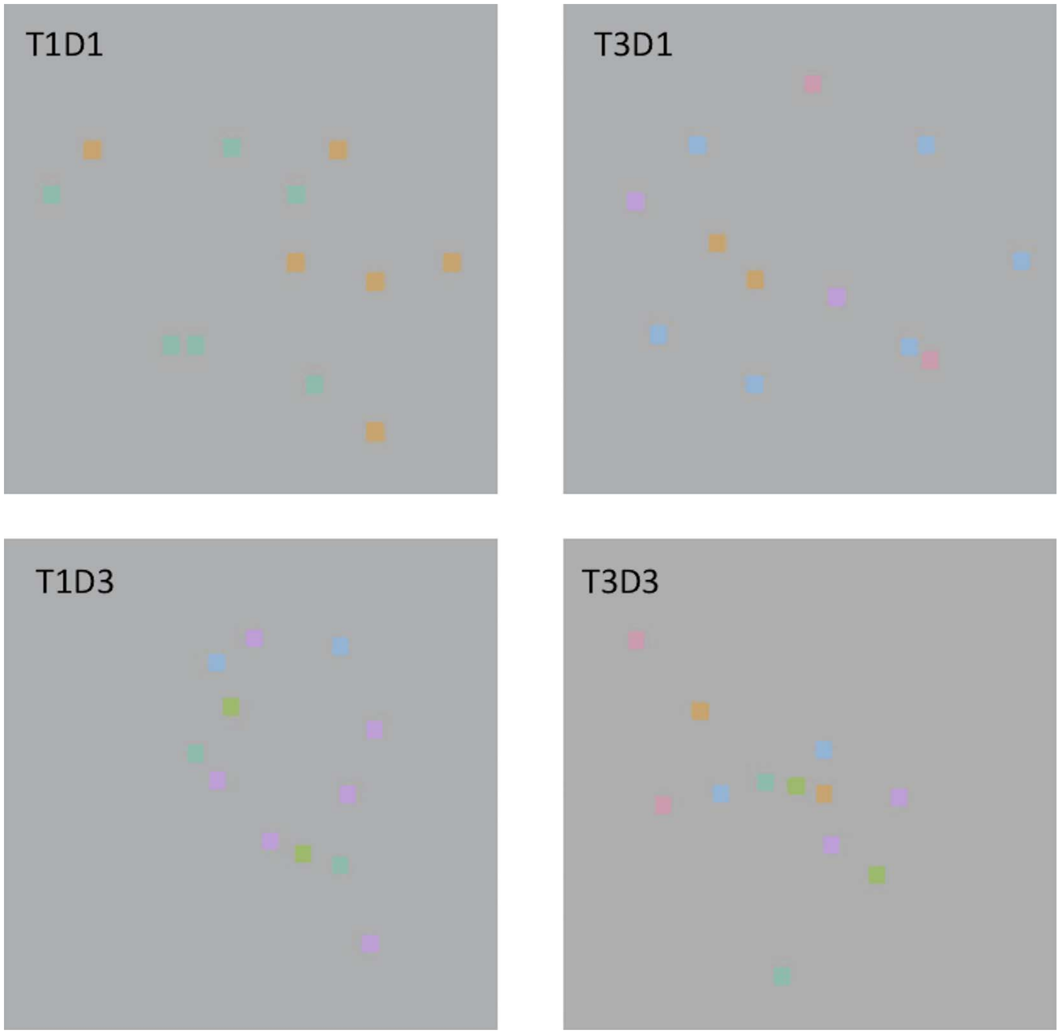

Figure 2. Sample stimuli for each of the four tasks. In all four tasks, orange, red, and magenta were the target colors.

same sequence of stimuli as was presented to the participant. On each simulated trial, dots are removed independently from each stimulus display with some probability p; then the remaining dots are given weights (according to their different types) by the influence function derived for the participant; finally, the centroid of the (decimated and filter-weighted) dot cloud is computed. The probability p is adjusted until the ideal observer's error matches the estimate of the participant's response error derived from the data. Efficiency is then $1 - p$. In each trial, there are a total of 12 dots present, so an efficiency of 91.7% means that on average, the ideal observer's accuracy matches the participant's best when 91.7%, or 11, of the 12 dots were included in the centroid calculation. Because participants' responses inevitably include error other than that due to missed stimulus items, efficiency can be understood as a lower bound on the number of items processed by a participant under the assumption that the participant uses the model-derived influence function.

Results

Mean errors (distance between response and centroid, measured in pixels) for full-set and target-only trials are displayed in Table 3. Participants performed significantly better in target-only trials than in full-set trials for all four tasks. For homogeneous targets, the increase in error was over 20%; for heterogeneous targets, the increase was 34% with homogeneous distractors and 58% with heterogeneous distractors. From this, we can conclude that there was substantial room for performance improvement in the full-set trials for any of the four tasks. We also found that there was a cost of having heterogeneous instead of homogeneous targets. Even without distractors present, the error in the target-only trials with heterogeneous targets was greater than the error in the same trials with homogeneous tasks, $\Delta = 17.2 - 15.9 = 1.3$, $t(7) = 5.379$, $p = 0.001$, Bayes factor (BF) $= 40.199$.[1]

Task		Full set error	Target-only error	Significance
Homogenous targets	Homogenous distractors	**19.1** [16.0 22.1]	**Δ = 3.1** **15.9** [13.2 18.7]	$t(7) = 6.015$ $p = 0.0005$* BF = 68.8
	Heterogenous distractors	**19.8** [14.7 25.0]	**Δ = 3.9**	$t(7) = 2.998$ $p = 0.02$* BF = 3.8
Heterogenous targets	Homogenous distractors	**23.1** [19.4 26.8]	**Δ = 5.9** **17.2** [13.8 20.2]	$t(7) = 4.654$ $p = 0.002$* BF = 20.8
	Heterogenous distractors	**27.2** [23.4 30.9]	**Δ = 10.0**	$t(7) = 10.204$ $p = 0.0005$* BF = 1110

Table 3. Mean centroid response error for each trial type in each task, averaged across participants, with confidence interval for each mean in brackets. *Notes*: Errors from the target-only conditions are pooled together in tasks T1D1 and T1D3, and then in tasks T3D1 and T3D3, since distractor heterogeneity did not affect these conditions. In all tasks, error in the full-set trials was significantly greater than in the target-only trials. From this table, we may conclude that participants were not performing at ceiling in any of the tasks in the full set trials. Note also that there was a cost of target heterogeneity even for the for-target-only conditions.

From the response-error data, we estimated the relative amount that each of the items in the display influenced the participant's centroid judgment; these values are referred to as *influence functions*. Figure 3 shows the influence functions averaged across all eight subjects, each panel displaying data for one of the four tasks. Each function reflects the attention filters for the six hues, shown on the x-axis of the plots. Each line in one of the panels in the figure connects the attention-filter estimates for one of the conditions for the task displayed in that panel. The points on each line represent the hues that were presented in the condition. Note that the lines serve only to identify the estimates from a condition; they should not be seen as interpolating the data between the data points. The y-axes represent the weight of each hue in the task. Influence functions are defined only up to an arbitrary multiplicative constant. As a matter of convention, we normalize the weights in any given influence function to sum to 1. Thus, in an ideal filter the weight of each target hue in the T3D3 and T3D1 conditions should be 1/3, whereas the ideal weight of the single target hue in the T1D3 and T1D1 conditions should be 1. In all tasks, the ideal weight of distractors was 0. Error bars on each point represent the 95% confidence intervals consistent with a repeated-measures analysis, with the main effect of participants removed (Morey, 2008; Franz & Loftus, 2012). These were calculated separately for the targets and the distractors, as the number of the number of target and distractor types varied between tasks.

The influence functions in Figure 3 suggest that participants generally were able to base their centroid responses on the target items and ignore the distractors. At the same time, there are clearly systematic differences across tasks that can best be summarized using the selectivity measure, which we will do in the following. Here we note that in the T3D3 and T3D1 tasks, the purple target (rightmost on the horizontal axis in each panel of Figure 3) exerts slightly less weight than the red and orange targets, which suggests that the purple hue was harder to categorize as a red than the orange hue: for the T3D3 task, $t(7) = -3.57$, $p = 0.009$, BF = 7.03; for the T3D1 task, $t(7) = -2.87$, $p = 0.024$, BF = 3.31. However, this asymmetry is not found in tasks T1D3 or T1D1, in which the targets were homogeneous. Also interesting is that, in all tasks, the distance in hue space between the distractor and target in the different conditions did not have a systematic effect on performance. We might have expected worse performance for conditions in which the target was closer to the distractor around the circle of hues, namely when Colors 3 (yellow-green) and 4 (orange) were paired or when Colors 1 (blue-green) and 6 (purple) were paired. Likewise, we would have expected the best performance to emerge when Colors 3 (green) and 7 (red) were paired, but none of these conditions was significantly different from one another. This noneffect of hue pairs confirms that performance in the centroid task is driven not by the saliency of the hues of the target dots but rather by whether or not attention is allotted to the hue. The evidence indicates that none of the hues used in the experiment was more salient than any other.

As noted before, the important differences between tasks are best captured using log10 selectivity and

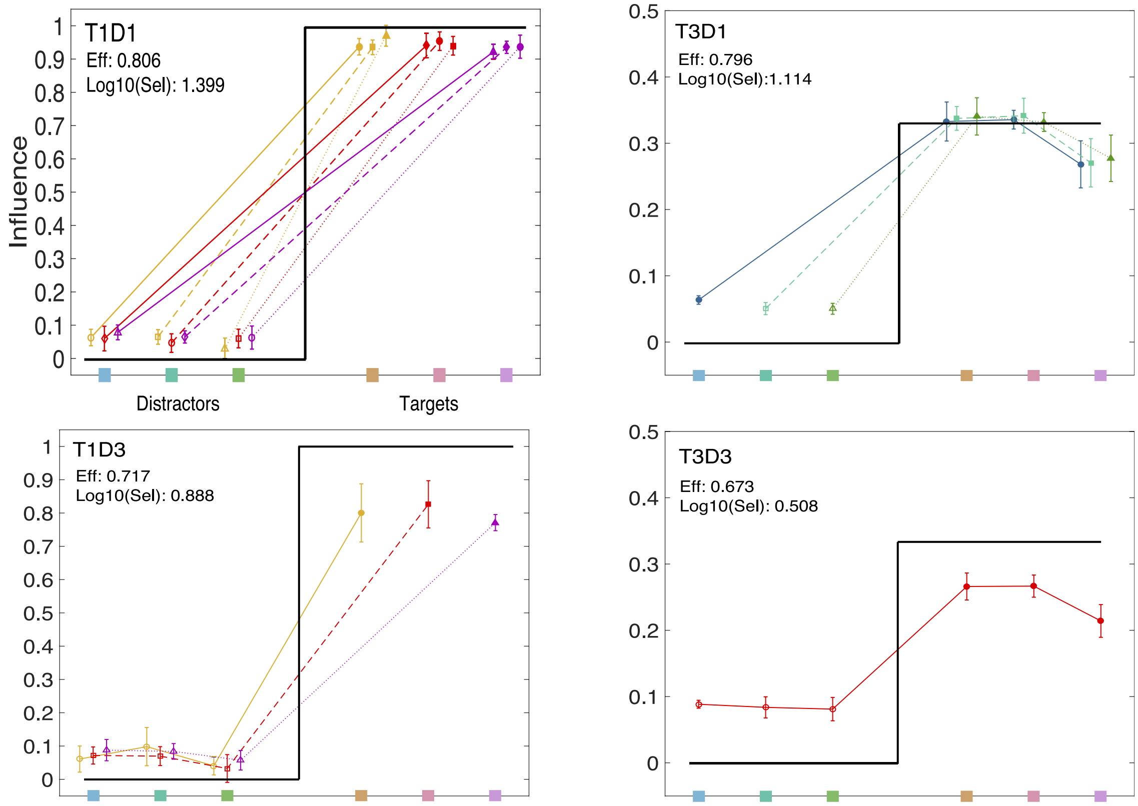

Figure 3. Influence functions, averaged across participants, for each of the four tasks. Task T1D1 consists of one target hue and one distractor hue on each trial, T3D1 consists of one distractor hue and three target colors on each trial, T1D3 consists of one target hue and three distractor hues on each trial, and T3D3 consists of three target hues and three distractor hues on each trial. The colored squares on the horizontal axis show the approximate hue of each stimulus. The height of each data point represents the weight of that hue for one of the stimulus conditions—a combination of target and distractor types—included in the task. The colored lines simply connect the data points from a stimulus condition, so the points on a line indicate which colors were present in the condition, and should not be interpreted to interpolate between those points. The solid black lines display the ideal filter for each task. Error bars reflect a 95% confidence interval for the weight of that color computed with the main effect of subjects removed and then adjusted to eliminate bias as described by Morey (2008) and Franz and Loftus (2012).

efficiency. These measures, averaged across participants in each task, are summarized in Tables 4 and 5.

Looking at both measures, we observe effects of target homogeneity versus heterogeneity—for selectivity, $t(7) = 4.39$, $p = 0.003$, BF = 16.14; for efficiency, $t(7) = 5.799$, $p = 0.0003$, BF = 57.54—and distractor homogeneity versus heterogeneity—for selectivity, $t(7) = 3.674$, $p = 0.008$, BF = 7.84; for efficiency, $t(7) = 2.381$, $p = 0.049$, BF = 1.93—with significantly larger effects of target heterogeneity than distractor heterogeneity: for selectivity, $t(7) = 6.969$, $p = 0.00011$, BF = 143.96; for efficiency, $t(7) = 2.451$, $p = 0.044$, BF = 2.09. There was also little evidence for an interaction between target and distractor homogeneity versus heterogeneity: for selectivity, $t(7) = -0.676$, $p = 0.521$, BF = 0.406; for efficiency, $t(7) = -1.274$, $p = 0.243$, BF = 0.63.

Looking at the effects of skill level, there were significant differences in selectivity between experts and novices in the T3D1, T1D3, and T1D1 tasks. Experts' log10 selectivity ratios, on average, were greater than novices' by 0.6 (four times larger) in the T3D1 task, $t(7) = 3.025$, $p = 0.023$, BF = 3.92; by 0.7 in the T1D1 task, $t(7) = 3.880$, $p = 0.008$, BF = 9.70; and by 1.0 in the T1D1 task, $t(7) = 3.881$, $p = 0.008$, BF = 7.14. There was no significant difference in performance between the expert group and novice group on the T3D3 task—for selectivity, $\Delta = 0.318$, $t(7) = 1.861$, $p = 0.112$, BF = 0.91—but what difference there was did show that experts still performed better than novices on average. There was no significant difference in efficiency between novices and experts in any of the four tasks. There were also no significant interactions, suggesting that the effects of target and distractor homogeneity versus heterogeneity generalize across skill levels. The effects we found for target and distractor heterogeneity are therefore not driven by practice with the centroid task. Experts were more selective for targets overall, but their performance in the heterogeneous conditions worsened as much as the novices'. This suggests that the effects are not task specific.

		Target hues: T1 (1 hue per trial)	Target hues: T3 (3 hues per trial)	Mean	
Distractor hues	D1 (1 hue per trial)	**1.40** (25.1) [1.04 1.76]	**0.89** (7.7) [0.73 1.04]	**1.14** (13.9)	Distractor effect (D1-D3): **0.33** (2.1) [0.12 0.55]
	D3 (3 hues per trial)	**1.11** (13.0) [1.00 1.23]	**0.51** (3.2) [0.28 0.74]	**0.81** (6.471)	
	Mean	**1.26** (18.1)	**0.70** (5.0)	**0.98** (9.5)	
		Target effect (T1-T3): **0.56** (3.6) [0.26 0.86]			Interaction (T3D3 – T3D1) – (T1D3 – T1D1): **0.10** (1.2) [-0.43 0.24]

Table 4. Strength of the influence functions in terms of log10 selectivity for four tasks, together with the confidence intervals of the values averaged across all eight subjects. *Notes*: In all but the T3D3 task, the value shown in each cell is the average of the values determined separately in the conditions within the task. Values in parentheses are the selectivity ratios exponentiated to reverse the log10 transformation.

Discussion

Results revealed that target and distractor heterogeneity both degraded performance in the centroid task, but that target heterogeneity had a larger effect than distractor heterogeneity. Back-transforming the values from Table 4 to their untransformed linear scale, we found that selectivity for T1D1 averaged over subjects and hues was 25.1, and selectivity for T3D3 was 3.22 (Table 4); this is a large difference, a factor of 6.8. The other two conditions were of intermediate selectivity, with selectivity reduced by almost twice as much due to increasing target heterogeneity (i.e., T1D1 vs. T3D1: 3.6) than to increasing distractor heterogeneity (i.e., T1D1 vs. T1D3: 1.9). In the logarithmic

		Target hues: T1 (1 hue per trial)	Target hues: T3 (3 hues per trial)	Mean	
Distractor hues	D1 (1 hue per trial)	**0.81** (9.7) [0.78 0.83]	**0.72** (8.6) [0.69 0.75]	**0.80**	Distractor effect (D1-D3): **0.027** [0.00 0.05]
	D3 (3 hues per trial)	**0.80** (9.6) [0.77 0.83]	**0.67** (8.1) [0.62 0.72]	**0.73**	
	Mean	**0.80**	**0.70**	**0.75**	
		Target effect (T1-T3): **0.11** [0.06 0.15]			Interaction (T3D3 – T3D1) – (T1D3 – T1D1): **-0.03** [-0.10 0.03]

Table 5. Efficiencies for four tasks, together with the ranges of the values averaged across all eight subjects. *Notes*: Just as in Table 4, the value shown for each task (except for T3D3) is the average of the values determined separately in the conditions within the task. Values in parentheses are $N \times$ Efficiency, which gives the lower bound on the estimate of the total number of dots (out of a possible 12) that the participants must have processed on average for the task.

analysis, the interaction is small and not reliable, suggesting that separate effects of target and distractor heterogeneity combine multiplicatively (additively in the log domain) to produce the overall selectivity. This two-factor log-linear model accounts for 99.5% of the variance in log10 selectivity. For efficiency, the two-factor linear model accounts for 97.7% of the systematic variance.

In none of the tasks did the distance between the hues of targets versus distractors around the hue circle have an effect on performance. We also found that in the T3D1 and T3D3 tasks, the weights of the targets were smaller for the purple targets, which suggests that it was more difficult for the participants to categorize purple as a reddish hue than it was for them to categorize orange as a reddish hue; however, this effect was found only when target heterogeneity was high.

Although we have strong evidence that target heterogeneity affects performance, we were concerned that running each task in separate blocks may have allowed participants to optimize their strategy in a way that improved performance when the targets or the distractors were homogeneous. In response to this concern, we had five of the participants complete a control experiment in which trials from all four tasks were mixed together. If participants were in some way optimizing the filter used for each task based on the knowledge of the task, then they would be expected to perform worse when conditions were mixed in the same block. If they were using a fixed response computation throughout the experiment, then their performance in the mixed task should not differ from that in the separately blocked tasks. We found that mixing the conditions did not measurably change performance.

Another concern was that numerosity may have affected performance. Research conducted in our lab suggests that when the tokens of one hue in a stimulus cloud are substantially more numerous than the tokens of any of the types of a set of mixed-hue distractors, participants are able to find the centroid of dots of that more numerous hue with high efficiency and selectivity (Sun, Chubb, Wright, & Sperling, 2018), even when they do not know beforehand what that hue will be. Unlike the top-down effects of selective attention that we have been discussing, this appears to reflect a stimulus-driven (bottom-up) form of selective attention. In the current experiment, the tokens of the target hue were more numerous than tokens of the distractor types in the T1D3 task. Given the results of Sun et al., this numerosity difference might have been expected to produce a bottom-up effect that would increase the salience of the targets, improving performance in that task. Analogously, in the T3D1 task the tokens of the one distractor type were more numerous than those of any of the target types. This numerosity difference might have been expected to increase the salience of the distractors, perhaps harming performance. The presence of either of these effects would have shown up in the analyses as an interaction. However, because we did not find a significant interaction between target and distractor heterogeneity in the current experiment for either log10 selectivity or efficiency, we conclude that if numerosity did have either of these effects, they were not large enough to be discernible in our design.

To assess the implication of these results, consider that a common interpretation of the analogy that attention operates like a filter—one in which the characteristics of the filter are endogenously determined—leads to the expectation that participants would use the same filter, one with a broad passband for reddish stimuli, in all four tasks of this experiment. This interpretation suggests that selectivity ratios and efficiencies should not vary across tasks. Contrary to this expectation, performance was better when targets were homogeneous, even though participants could not predict the target to which they would have to attend on each trial. Similarly, when distractors were homogeneous, they were easier to ignore. This is consistent with findings from previous research demonstrating that search times increased as variation in the distractor group increased (Bundesen & Pedersen, 1983). We conclude that the simple analogy of an endogenous filter cannot solely explain the phenomena of feature-based attention.

One way to explain these results and save the general analogy that attention operates like a filter is to posit a mechanism in which the filter is generated dynamically, starting from an endogenously determined goal, to reflect the actual statistics of the overall scene being processed, or possibly even just local patches of the scene (Foley, 1994; Zenger & Sagi, 1996; Lee, Itti, Koch, & Braun, 1999; Ren & Malik, 2003; Danelljan, Hager, Shahbaz Khan, & Felsberg, 2015). If this suggestion is correct, it may have been important that all of the targets were slightly brighter than the background. We made this choice because experience in our lab suggests that centroid performance for hue targets is reduced when the dots are equiluminant with the background. This luminance increment may have allowed the stimulus items to be discriminated from the background so that their histogram could be computed and an optimal filter constructed for the particular set of target and distractor tokens appearing on that trial.

A different way to understand these results would be grouping, a bottom-up mechanism that can segregate, precategorically, similar items in a visual scene. With low target heterogeneity, the targets would all be clustered into a single group because of their shared reddish hue, and the resulting group could be easily selected. With high target heterogeneity, the targets would be clustered into three groups, each of which would have to be identified separately; their three locations would then need to be maintained until the locations could be merged, presumably by a different process than that used to find the centroid of a selected group of dots, to produce the overall centroid. Error introduced in maintaining or merging the centroids of these separate groups could explain the reduced selectivity and efficiency in the T3D1 and T3D3 tasks. Similarly, when distractors are homogeneous (in the T1D1 and T3D1 tasks), the distractors would be clustered into a single group that would be easier to ignore.

Future work will have to look for evidence to select between these and other possible mechanisms in order to better understand how feature-based attention operates. One important question to address would be how far apart stimulus items can be in feature space without incurring the centroid-calculation costs observed here, and whether this distance depends on the arrangement of the targets and distractors. So, for example, Bauer et al. (1996) have shown that there is a substantial increase in slope for visual search when the target and multiple distractor types are close to being colinear in color space; Sun et al. (2016a) have demonstrated a similar effect in the context of the centroid task.

A final issue that is of particular interest to those of us working with the centroid task is whether the results produced with this procedure reflect the operation of feature-based attention more generally or something particular to the centroid task. We used the centroid task for this research because the manipulation of both target and distractor heterogeneity is quite natural within the context of this task—although it is possible to manipulate within-trial distractor heterogeneity in a visual-search task, this cannot be done for target heterogeneity (Bundesen & Pedersen, 1983; Duncan & Humphreys, 1989; D'Zmura, 1991; Bravo & Nakayama, 1992; Bauer et al., 1996; Nagy & Thomas, 2003; Nagy et al., 2005). The results from this study provide some evidence that these effects of heterogeneity are not limited to this task, as we did not find a significant interaction between practice and target and distractor heterogeneity.

Although it is perhaps not a common task, locating the centroid of spatially separate objects is something that the visual(-motor) system appears to do naturally (McGowan et al., 1998; Baud-Bovy & Soechting, 2001). This tendency to locate the center of mass plays a role in grasping, as shown by Goodale et al. (1994). Participants in that study were asked to grip objects with their thumb and index finger, and the line joining the two contact points tended to pass close to the center of mass of the object. We speculate that the feature-based processes underlying centroid percepts are similar to the parallel search processes ascribed to visual search by Dosher, Han, and Lu (2010) or, more recently, Buetti et al. (2016), wherein the location of target items is easily extracted if the target is dissimilar from the distractors. Also, this parallel process does not rely solely on whether or not an item "pops out" from its distractors but rather involves attentive filtering that is influenced by the observer's goals. Ultimately, however, it may require neurophysiological studies to resolve this question of generality.

Conclusions

In this experiment, we tested the filter analogy for feature-based attention by varying target and distractor heterogeneity and seeing whether this variation affected participants' ability to pick the centroid of the target dots. Assuming the filter analogy could best describe attention, performance would have been consistent across all four tasks. We found, however, that participants achieved the highest log10 selectivity and efficiency for the task with no heterogeneity in the target and distractor groups, and the lowest for the task with the highest heterogeneity in both groups. This indicates that the simple filter analogy does not accurately describe the mechanism that drives attention. Instead, the mechanism that is used may be more similar to a grouping mechanism.

Keywords: feature-based attention, centroid task, target/distractor heterogeneity

Acknowledgments

The authors would like to thank their colleagues in the lab for their help collecting data and suggesting how to interpret the results.

Commercial relationships: none.
Corresponding author: Charles E. Wright.
Email: cewright@uci.edu.
Address: Department of Cognitive Sciences, University of California, Irvine, Irvine, California, USA.

Footnote

[1] The Bayes factor is the Bayesian alternative to classical hypothesis tests based on the ratio of the probability of the alternative hypothesis compared to the null hypothesis (Goodman, 1999). The value was computed with version 0.9.8 of the BayesFactor package, available online at http://pcl.missouri.edu/bf-one-sample (Rouder, Speckman, Sun, Morey & Iverson, 2009).

References

Antis, S. M., & Cavanagh, P. (1983). A minimum motion technique for judging equiluminance. In J. D. Mollon & L. T. Sharpe (Eds.), *Colour vision: Physiology and psychophysics* (pp. 155–166). London, UK: Academic Press.

Baud-Bovy, G., & Soechting, J. (2001). Visual localization of the center of mass of compact, asymmetric, two-dimensional shapes. *Journal of Experimental Psychology: Human Perception and Performance*, *27*, 692–706.

Bauer, B., Jolicoeur, P., & Cowan, W. B. (1996). Visual search for colour targets that are or are not linearly separable from distractors. *Vision Research*, *36*(10), 1439–1466.

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*(4), 433–436, https://doi.org/10.1163/156856897x00357.

Bravo, M. J., & Nakayama, K. (1992). The role of attention in different visual-search tasks. *Attention, Perception, & Psychophysics*, *51*(5), 465–472.

Buetti, S., Cronin, D. A., Madison, A. M., Wang, Z., & Lleras, A. (2016). Towards a better understanding of parallel visual processing in human vision: Evidence for exhaustive analysis of visual information. *Journal of Experimental Psychology: General*, *145*(6), 672–707.

Bundesen, C., & Pedersen, L. F. (1983). Color segregation and visual search. *Perception & Psychophysics*, *33*(5), 487–493.

Danelljan, M., Hager, G., Shahbaz Khan, F., & Felsberg, M. (2015). Convolutional features for correlation filter based visual tracking. In H. J. Trussell (Ed.), *Proceedings of the IEEE International Conference on Computer Vision Workshops* (pp. 621–629). Los Alamitos, CA: IEEE Computer Society.

Derrington, A., Lennie, P., & Krauskopf, J. (1983). Chromatic response properties of parvocellular neurons in the macaque LGN. In J. D. Mollon & L. T. Sharpe (Eds.), *Colour vision: Physiology and psychophysics* (pp. 245–251). London, UK: Academic Press.

Dosher, B., Han, S., & Lu, Z. (2010). Object attention sharpens the tuning of the perceptual template and interacts with task precision. *Journal of Vision*, *10*(7): 242, https://doi.org/10.1167/10.7.242. [Abstract]

Drew, S. A., Chubb, C. F., & Sperling, G. (2010). Precise attention filters for Weber contrast derived from centroid estimations. *Journal of Vision*, *10*(10):20, 1–16, https://doi.org/10.1167/10.10.20. [PubMed] [Article]

Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological Review*, *96*(3), 433–458, https://doi.org/10.1037/0033-295x.96.3.433.

D'Zmura, M. (1991). Color in visual search. *Vision Research*, *31*(6), 951–966.

Foley, J. M. (1994). Human luminance pattern-vision mechanisms: Masking experiments require a new model. *Journal of the Optical Society of America A*, *11*(6), 1710–1719.

Franz, V. H., & Loftus, G. R. (2012). Standard errors and confidence intervals in within-subjects designs: Generalizing Loftus and Masson (1994) and avoiding the biases of alternative accounts. *Psychonomic Bulletin & Review*, *19*(3), 395–404.

Friedenberg, J., & Liby, B. (2002). Perception of two-body center of mass. *Perception & Psychophysics*, *64*(4), 531–539.

Goodale, M. A., Meenan, J. P., Bülthoff, H. H., Nicolle, D. A., Murphy, K. J., & Racicot, C. I. (1994). Separate neural pathways for the visual analysis of object shape in perception and prehension. *Current Biology*, *4*(7), 604–610.

Goodman, S. (1999). Toward evidence-based medical statistics. 2: The Bayes factor. *Annals of Internal Medicine*, *130*(12), 1005–1013.

Herrera, C., Sun, P., Groulx, K., Wright, C., Chubb, C., & Sperling, G. (2013). How do the S-, M- and L-cones contribute to motion luminance assessed using minimum motion? *Journal of Vision*, *13*(9): 1021, https://doi.org/10.1167/13.9.1021. [Abstract]

Lee, D. K., Itti, L., Koch, C., & Braun, J. (1999). Attention activates winner-take-all competition among visual filters. *Nature Neuroscience*, *2*(4), 375–381.

Lu, Z. L., & Sperling, G. (2001). Three-systems theory of human visual motion perception: Review and update. *Journal of the Optical Society of America A*, *18*(9), 2331–2370.

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

McGowan, J. W., Kowler, E., Sharma, A., & Chubb, C. (1998). Saccadic localization of random dot targets. *Vision Research*, *38*(6), 895–909.

Morey, R. D. (2008). Confidence intervals from normalized data: A correction to Cousineau (2005). *Tutorials in Quantitative Methods for Psychology*, *4*, 61–64.

Nagy, A. L., Neriani, K. E., & Young, T. L. (2005).

Effects of target and distractor heterogeneity on search for a color target. *Vision Research*, *45*(14), 1885–1899.

Nagy, A. L., & Thomas, G. (2003). Distractor heterogeneity, attention, and color in visual search. *Vision Research*, *43*(14), 1541–1552.

Ren, X., & Malik, J. (2003). Learning a classification model for segmentation. In F. T. Ulaby (Ed.), *Proceedings: Ninth IEEE International Conference on Computer Vision* (pp. 10–17). Los Alamitos, CA: IEEE Computer Society, https://doi.org/10.1109/iccv2003.1238308.

Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian *t*-tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review*, *16*, 225–237.

Sun, R., Chubb, C., Wright, C. E., & Sperling, G. (2016a). Human attention filters for single colors. *Proceedings of the National Academy of Sciences, USA*, *113*(43), E6712–E6720.

Sun, P., Chubb, C., Wright, C. E., & Sperling, G. (2016b). The centroid paradigm: Quantifying feature-based attention in terms of attention filters. *Attention, Perception, & Psychophysics*, *78*(2), 474–515, https://doi.org/10.3758/s13414-015-0978-2.

Sun, P., Chubb, C., Wright, C. E., & Sperling, G. (2018, Month). *The perceptual capacity of concurrent grouping of colored dots by similarity and by dissimilarity*. Paper presented at the meeting of the Vision Science Society, St. Pete Beach, FL.

Zenger, B., & Sagi, D. (1996). Isolating excitatory and inhibitory nonlinear spatial interactions involved in contrast detection. *Vision Research*, *36*(16), 2497–2513.