UNIVERSITY OF CALIFORNIA,
IRVINE

Grouping Principles in Centroid Tasks:
The influence of bottom-up attention on selective centroid judgements

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Cognitive Science

by

Vivian Thi Lu

Dissertation Committee:
Professor Charles E. Wright, Chair
Professor Charles Chubb
Professor Thomas Michael D'zmura

2019

DEDICATION

To my parents

for their love and support, especially during trying times

and to my sisters

for their equally important antics and laughter

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I would like to express the deepest appreciation to my committee chair, Professor Charles E. Wright, who had inspired me to pursue a doctorate in this field of study. Without his guidance and persistent help, this dissertation would not have been possible.

I would like to thank my committee members, Professor Charles Chubb and Professor Mike D'Zmura, for their time and their guidance while I designed my experiments and analyzed data.

A heartfelt 'thank you' to my cohort for their support and camaraderie as we progressed through graduate school together. If not for them, I would have felt lost and unsure of my abilities as a researcher. Thank you to my best friend and ever-patient partner for encouraging me to do my best every day. Through his example, he reminded me to persevere.

The text of this dissertation is a reprint of the material as it appears in Journal of Vision. The co-authors listed in this publication directed and supervised research which forms the basis for the dissertation. Financial support was provided by the University of California, Irvine, the Associate Dean's Award, and the Academic Senate's Council on Research, Computing, and Libraries Grant.

CURRICULUM VITAE

Vivian Thi Lu

Education

University of California- Irvine, School of Social Science, Irvine, CA

Ph.D. in Cognitive Sciences **2019**

Dissertation: Grouping Principles in Centroid Tasks: The influence of bottom-up attention on selective centroid judgements

Chair: Professor Charles E. Wright

Masters of Arts in Psychology **2017**

Bachelor of Science in Cognitive Science (with concentration in Experimental Psychology) **2015**

Awards and Honors

Associate Dean's Award **March 2019**

Academic Senate's Council on Research, Computing, and Libraries Grant **July 2018**

Magna Cum Laude **June 2015**

Campus-Wide Honors **June 2015**

Honors in Cognitive Sciences **June 2015**

Presentations and Publications

Psychonomics Society Meeting, Boston, Massachusetts **November, 2016**

Title: Target-item Heterogeneity Undermines Performance in Estimating Centroids

Undergraduate Research Opportunities Program Symposium, UCI **May, 2015**

Title: A Study of Luminance in Feature-Based Attention

Journal of Vision

Vivian T. Lu, Charles E. Wright, Charles Chubb, George Sperling; Variation in target and distractor heterogeneity impacts performance in the centroid task. *Journal of Vision* 2019;19(4):21. doi: 10.1167/19.4.21.

Experience

Lab of Dr. Charles E. Wright, Department of Cognitive Sciences

Researcher **June 2012-June 2019**

- Designed novel experiments in MATLAB to probe stimuli and observe how differences in features affect selectivity for target objects in a display. Projects involved programming, troubleshooting, and organizing schedules with participants to collect data.
- Wrote code in MATLAB to organize data and to generate regression models to find which factors in the experiment influenced the data most. Analysis codes utilized t-tests, contrasts, and ANOVAS to test for significance of the factors.
- Trained research assistants in coding, troubleshooting, data collection, and data analysis in MATLAB

Lab of Dr. Naomi Sugie, Department of Criminology, Law, and Society

UX Designer **September 2017 – June 2019**

- Created wireframes and interactive prototypes with a coding team for the interface of a mobile application designed to assist individuals recently released from incarceration find employment and reintegrate into society.

- Developed collaboration with leaders of reintegration programs in Orange County to conduct user studies and interviews to determine the usability of the application.
- Summarized qualitative data from user studies in Excel into succinct reports and implemented changes to the design of the application based on feedback given.

Huntley-Fenner Advisors (Irvine, California)
Human Factors Research Intern **June 2018 - September 2018**

- Organized and analyzed customer reviews from Amazon and Target using Excel pivot tables. Conducted omnibus tests to identify significant differences between Amazon and Target users.
- Aggregated data sets from the Consumer Product Safety Commission databases using their built-in query tool.
- Made recommendations for organization of data and reported summaries and interpretations to principal consultant.

ABSTRACT OF THE DISSERTATION

Grouping Principles in Centroid Tasks:
The influence of bottom-up attention on selective centroid judgements

By

Vivian Thi Lu

Doctor of Philosophy in Cognitive Sciences

University of California, Irvine, 2019

Professor Charles E. Wright, Chair

When participants perform a selective centroid task, they are instructed to attend to a specific stimulus type, and therefore are assumed to deploy top-down attention in order to find the target items in the display. Bottom-up mechanisms that may be characterized as grouping may also be influencing performance. In this thesis, we explore the role that this bottom-up mechanism may have for our ability to compute statistical summaries of a cluster of target stimuli. In the first study, targets and distractors were either homogeneous or heterogeneous. Results showed that homogeneity in both target and distractor groups improves selectivity for target stimuli and that the effect of target homogeneity was greater than the effect of distractor homogeneity. This strong effect of target homogeneity is further supported by the second study, in which participants are not told which stimuli would be the target, but that they needed to find the centroid of the “most numerous color.” To explore whether top-down and bottom-up attention are separate mechanisms or whether they compete for resources within a single mechanism, participants were told to attend either to a specified color or to the most numerous color. One condition of interest

in this experiment asked participants to attend to a specific color that was not the most numerous, which caused top-down and bottom-up attention to conflict with one another. Results indicated that participants were able to ignore the most numerous color (which drove bottom-up attention) and attend to the instructed target color (which drove top-down attention). However, salient features prevailed in the third study, in which participants were asked to attend to the color and ignore the size of the stimuli. They were unable to completely disregard the size of the stimuli when making their centroid judgement, which suggests that bottom-up cues may not be obligatory, but they do influence which stimuli within the target group is given more weight. A fourth study was conducted to find when, in relation to the grouping process, luminance constancy occurred. Results hinted that constancy may be determined before grouping occurs but were overall inconclusive.

CHAPTER 1

INTRODUCTION

Imagine going to a concert with a group of friends. You get lost during all of the excitement, and you need to find your group before the night is over. As you scan the crowd for your friends, you notice your attention being drawn towards large groups of people who decided to wear matching t-shirts. A blue group is standing close to the food stands, and a green group is standing next to the stage. Despite searching for individual people, you find it difficult to ignore the large groups of colors, even though they aren't relevant to your search. You suddenly remember that most people in your group decided to wear yellow that day, so instead of searching for individuals, you look around for a large group of people wearing yellow. Eventually, you find a yellow group, and rush over to join them.

In your search for your friends, you had to employ a type of selection called feature-based attention, which allows you to focus on certain features (here, t-shirt color) of interest. Although selective attention is a top-down process that falls under our control, it can still be influenced by the stimulus properties. Salient objects tend to draw our attention in (other distracting t-shirt colors), regardless of whether or not we had intended to search for them Jonides' (1981) study showed that while it is possible to ignore top-down cues, it is impossible to ignore bottom-up cues. He also found that bottom-up attention is less affected by cognitive load than top-down attention, as evidenced by the increase in reaction time when memory load increases. Participants had more difficulty suppressing a peripheral cue than a central one, which led to the conclusions that bottom-up cues have bigger effects than top-down cues do.

One method that measures feature-based attention asks participants to find the center of mass of a set of target items. 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, Chubb, Wright, & Sperling, 2016b). The paradigm expands on the traditional search task – in which participants only give binary responses – by having participants indicate the position of the center of targets. This allows us to not only estimate guessing, but also to infer the influence that the distractors had on the response. One drawback to the paradigm is that finding the centroid seems to be an automatic response driven by bottom-up attention, rather than solely top-down attention. One explanation for why this may be is grouping, which also seems to be an automatic process. If so, this is a confound, since we might not actually be measuring top-down attention, but rather bottom-up.

Grouping principles are described in Gestalt psychology as automatic rules that are used to group items in an image together (Wagemans, J., Elder, J. H., Kubovy, M., Palmer, S. E., Peterson, M. A., Singh, M., & von der Heydt, R. (2012). The rule of similarity states that people tend to group together items that share similar features with one another (ex. All red items are grouped, or all circular items are grouped). There is the question of how strong this effect is, and where in the perceptual process it occurs. Can it be ignored? Kahneman and Henik (1977) hypothesize that attention is allocated first to groups, then to the individual members of that group. We have used our centroid task to study the automaticity of grouping and how it affects the way in which attention is allocated to items in a display.

Some examples of how we have utilized our method are described in this paper. In the first study, we varied target and distractor heterogeneity (either one or three types of stimuli) to test the filter analogy of attention. We had expected participants to create broad attention filters in order to locate all targets in the display, regardless of heterogeneity, but we found that performance is much worse when participants are asked to attend to multiple colors than when they are asked to attend to only one color. When distractors only consisted of one color, performance also improved, which may be caused by participants being able to group distractors together into clusters to be ignored. Grouping seemed to be involved, and so the second part of the study expands on the first to test the granularity of this grouping mechanism. Target colors that are more similar to each other were expected to be grouped together more easily than target colors that lie further away from one another in color space. Results showed that there is not much evidence that categorical grouping occurred in the experiment.

The second study aimed to determine whether or not this grouping process is obligatory and draws attention in, even when it conflicts with instructions to attend to a different color. Participants were asked to calculate the centroid of a certain color, ignoring numerosity, which previous studies have suggested draws attention in because our attention wants to automatically focus on, and group, similar items. Grouping seems to be a bottom-up processes, so we pitted it against top-down instructions to see if attention is a single process driven by both mechanisms, or if top-down and bottom-up attention are two separate processes that interact with each other. Results suggest that attention is one process influenced by these two mechanisms.

The third study evaluated the effects of irrelevant features on the centroid task. Normally in a centroid task, participants are asked to attend to one feature, and to use that one feature to select the targets out of the stimulus cloud. The stimuli only vary in that one feature. This third study varied stimuli on two features – color and size – and participants were asked to only attend to color. Although size was not informative of whether or not the item was a target, there were three distinct sizes that the participants may have grouped the stimuli into. If it was the case that grouping is an automatic process that cannot be ignored, then we expected to have seen effects of size on the centroid judgement, which we did.

Lastly, the fourth study aimed to determine when in visual processing grouping occurred. From the other three studies conducted, it was clear that if grouping was occurring, it was occurring before the centroid judgement is made. The fourth study compared the quickness of grouping to that of another quick process that occurs in visual processing: luminance constancy. If grouping were a relatively fast process, we would have expected to see participants group stimuli based on their physical luminance. If it were a relatively slow process compared to luminance constancy, then we could have expected participants to group based on the reflectance of the dots, which takes both the physical luminance and the light sources in to account. Our results were inconclusive, although they do seem to suggest that grouping may be a slower process than luminance constancy.

The methods that we used to collect and analyze are described below, followed by a discussion of each of the four experiments conducted.

CHAPTER 2

METHODOLOGY

I. Procedure

In all experiments, 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 is presented in Figure 1.

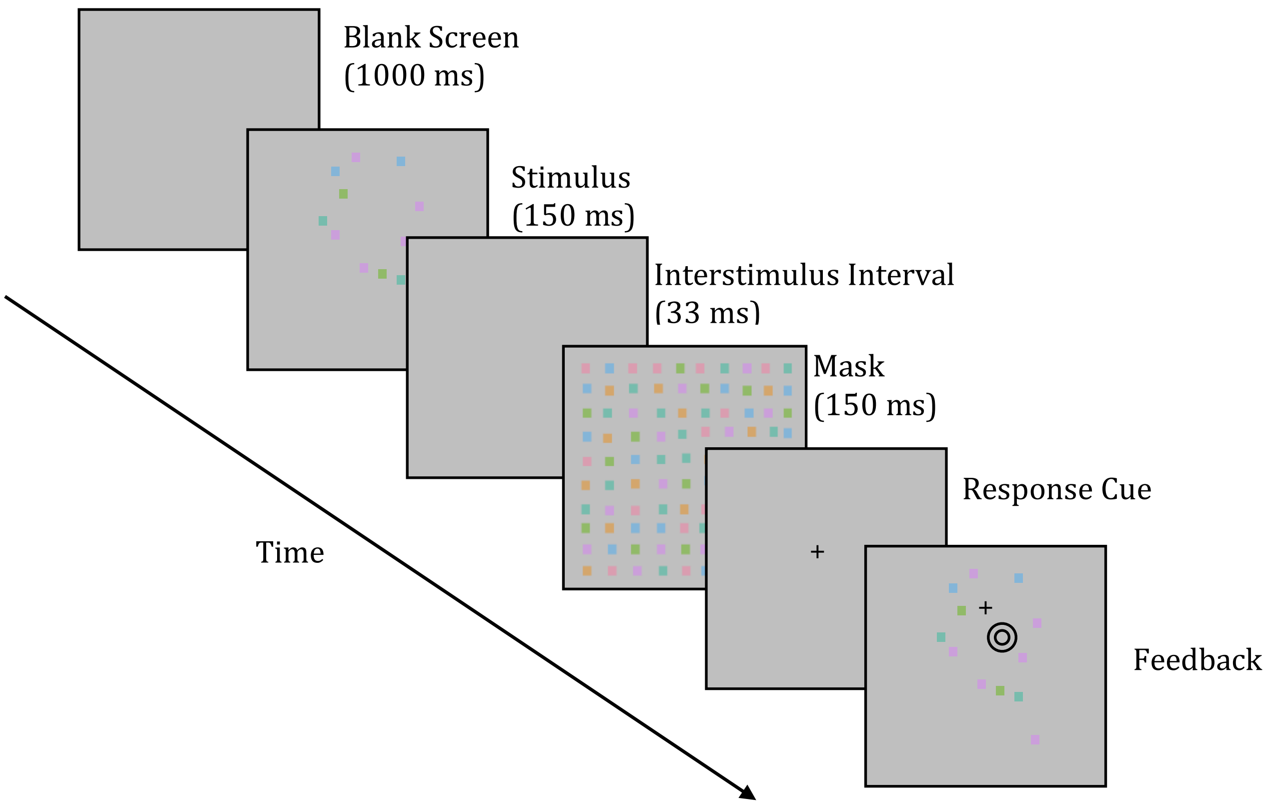


Figure 1: Sequence of display events on a trial. Stimulus displays were presented for 150 milliseconds, followed by a blank screen for 33 milliseconds, and a mask for another 150 milliseconds. Once the cross appeared, participants were able to move the mouse to adjust its position to the perceived centroid 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.

To decompose the response errors, we followed the procedures described in Sun et al. (2016b), to derive an influence function, f, from each participant's data in each condition. The first step in these analyses generates estimates of the observer's attention filter, $f\varphi$. An observer's attention filter is the vector of weights (one for each of the 8 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 value 1, for the target hue, and zero for the distractor hue(s). For the tasks with three target hues, the target filter takes the value 1/3 for each of those hues, and zero for the distractor hue(s).

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

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

where the sum is over all squares © in the display, h_i is the hue of square ©, 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

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

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_φ(h),

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

In Eq. (3) $h_{t,i}$, $x_{t,i}$, and $y_{t,i}$ are the hue and x- and y-coordinates of the i$^{\text{th}}$ 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 participants' 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 © 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 ten times more weight on the participant's response than 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 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. For example, if there are a total of 12 dots present, 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.

II. Participants

All observers in the studies are students at the University of California, Irvine, between the ages of 18 and 30. All observers had normal or corrected-to-normal vision. None of the participants were affected by color blindness. The studies were conducted in accordance with the regulations of the Institutional Review Board of the University of California, Irvine.

III. Apparatus and Stimuli

Participants ran the experiment on a Mac computer (iMac10,1) running MATLAB's Psychtoolbox package (Brainard, 1997). The stimuli were displayed on an LED monitor — integrally part of the computer — with resolution of 1920 by 1080. Stimuli were viewed from a distance of 70 cm.

The stimulus region was 640 by 640 pixels (visual angle 12.85 degrees), centered on the 1920 by 1080 pixel display. Dots were 17 by 17 pixel squares (visual angle 0.34 degrees) 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 degrees). If any 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 x,y color space are listed in Table 1. Colors 6, 7, and 8 were 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 (Derrington, Lennie, & Krauskopf, 1983) model of color space, commonly used in color research, and during pilot sessions, these were the colors for which participants exhibited the highest selectivity.

Colors	x-coordinate	y-coordinate
Color 1	0.2755	0.2264
Color 2	0.2586	0.3000
Color 3	0.2787	0.3765
Color 4	0.3493	0.4757
Color 5	0.4307	0.5026
Color 6	0.4308	0.4154
Color 7	0.3762	0.3218
Color 8	0.3173	0.2707
Gray background	0.3257	0.3478

Table 1: An example for a set of CIE 1931 x,y color space values for the 8 colors used for one participant.

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/m2) than the gray background (47.87 cd/m2) 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, Sun, Groulx, Wright, Chubb, & Sperling, 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 8 hues in the color circle, i.e., the 8 hues functioned equally well as targets.

CHAPTER 3

STUDY 1 - VARIATION IN TARGET AND DISTRACTOR HETEROGENEITY

Study 1a: Introduction

Increasing the similarity between targets and distractors, increasing target heterogeneity or 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 vs. 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 problems may emerge: some types of distractor items may pass through with sufficient salience to alter subsequent computations, and/or 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 stopbands: 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 stopbands. This explanation arrives at the same conclusion as studies cited above: 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 stopbands and sharp transitions.

The purpose of the current experiment was to find 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 Chubb, Wright, & Sperling, 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 purpose of 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 only been able to study target heterogeneity by varying the target type across trials, since those tasks rely on the presence of only one or no target present 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 stopbands mentioned above. 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 were 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 because we know that participants are good at discriminating stimuli based on hue. In fact, they are better at finding the centroid of the 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 they 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 then improves as the distractor hues move further away from the target hues. Because our target and distractor hues lie on opposite 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 between target hues and distractor hues in color space increases, participants will be able to distinguish them better and assign more weight to the targets, and perform better at the task.

Study 1a: Design

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, 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 homogenous. 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 task and the T3D3 task. The key assumption required for this to be true is that in a given task condition, (1) the participant 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 (2) this filter must remain fixed across the brief duration of a stimulus display in each of the experimental conditions tested in the current study. 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 and/or the distractors.

All trials consisted of twelve dots: six targets and six distractors. The experiment was divided into four tasks, summarized in Table 1. The tasks are labelled with how many targets are present ("T1" for one target type and "T3" for three target types) and how many distractors ("D1" for one distractor type and "D3" for three distractor types) are present in each trial. In the T1D3 and T1D1 tasks (in both of which only one target color was presented on a given trial), one of the three target colors (Colors 6-8) was randomly chosen for all six target dots on a trial. Likewise, in the T3D1 and T1D1 tasks, one of the three distractor colors (Colors 2-4) was randomly chosen for all six distractor dots on a trial. In tasks with "T3," all three target colors were present on each trial, with two dots for each color to maintain a total of six target dots. Tasks with "D3" contained all three distractor colors in each trial, with two dots for each color (Figure 2).

All blocks in each task consisted of 90 trials. Task 1 had six blocks, Tasks 2 and 3 had eight blocks each, and Task 4 had ten blocks. The number of blocks in each task varied due to the number of colors that appeared in each trial in each task. Because all six colors were presented in each trial in the T3D3, participants were able to provide enough data to calculate influence functions for each color in only six blocks. In the other tasks, not all colors 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 have sufficient precision in the influence functions. Eight participants were recruited for the study; each one completed each of the four tasks twice in a randomized order determined by a Latin square. For this study, we collected data from four experienced participants (those who have participated in other centroid experiments prior to this one) and four inexperienced participants (those who had never participated in a centroid experiment) to see if skill level had an effect on performance.

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

Table 2: Four tasks used in the experiment. Dot-color heterogeneity vs homogeneity was manipulated by a 2x2 factorial design with dot type: targets vs distractors.

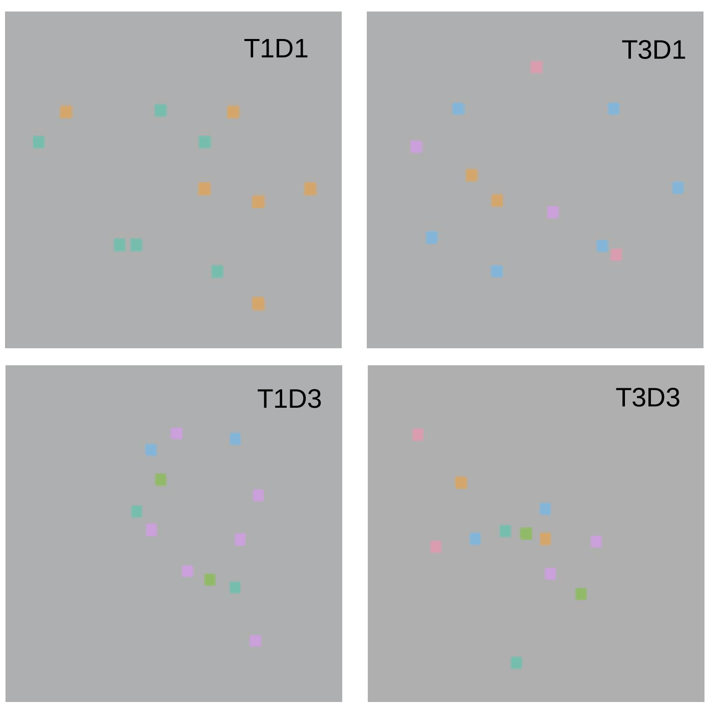


Figure 2: Sample stimuli for each of the four tasks.

Colors 6, 7, and 8 were 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 (Derrington, Krauskopf, and Lennie, 1983) model of color space, commonly used in color research, and during pilot sessions, these were the colors for which participants exhibited the highest selectivity.

Study 1a: Results

Mean errors (distance between response and centroid, measured in pixels) for full set trials 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 homogenous targets, the increase in error was over 20%; for the heterogeneous targets, the increase was 34% with homogenous 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 heterogenous instead of homogenous targets. Even without distractors present, the error in the target-only trials in the heterogenous targets tasks was greater than the error in the same trials in the homogenous target tasks ($\Delta = 1.3 = 17.2 - 15.9$, $t(7) = 5.379$, $p = 0.001$, BF[1] = 40.199).

[1]BF is the Bayes Factor, 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).

Task		Full Set Error	Target-Only Error	Significance
Homogenous targets	**Homogenous Distractors**	19.1 [16.0 22.1]	Δ = 3.1	t(7) = 6.015 *p* = 0.0005* BF = 68.8
			15.9 [13.2 18.7]	
	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	t(7) = 4.654 *p* = 0.002* BF = 20.8
			17.2 [13.8 20.2]	
	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. Values in brackets show the confidence interval for each mean. 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 were significantly greater than those 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 8 participants, each panel displaying data for one of the four tasks. Each function reflects the attention filters for the 6 hues, shown on the x-axis of the plots. Each line in one of the panels in Fig. 3. Connects the attention filter estimates for one of the conditions for that task displayed in that panel. The points on each line represent the hues that was 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 between the data points. The y-axes in Fig. 3

represent the weight of each hue in the task. Influence functions are only defined 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 zero. Error bars on each point represent the 95% confidence intervals consistent with a repeated measures analysis, with the main effect of participants removed (Franz & Loftus, 2012, Morey, 2008). These were calculated separately for the targets and the distractors, as the number of the number of target and distractor types varied between tasks.

Influenced functions averaged across all 8 participants are shown in Figure 3. These functions reflect the attention filters for the 6 colors, shown on the x-axis of the plots. Each colored line in Fig. 3. Represents a condition in the task, and the points on the line represent the color that was presented in the condition. The y-axes in Fig. 3 represent the weight of each color in the task. Influence functions are only defined up to an arbitrary multiplicative constant. As a matter of convention, we force the weights in any given influence function to sum to 1. Thus, in an ideal filter (represented by the solid black line in each plot), the weight of each target color in the T3D3 and T3D1 conditions should be 1/3, whereas the ideal weight of the single target color in the T1D3 and T1D1 conditions should be 1. In all tasks, the ideal weight of distractors was zero. Error bars on each point represent the 95% confidence intervals consistent with repeated measures analysis, with the main effect of participants removed (Franz & Loftus, 2012). These were calculated separately for the targets and the distractors, as the number of the number of target and distractor type varied between tasks.

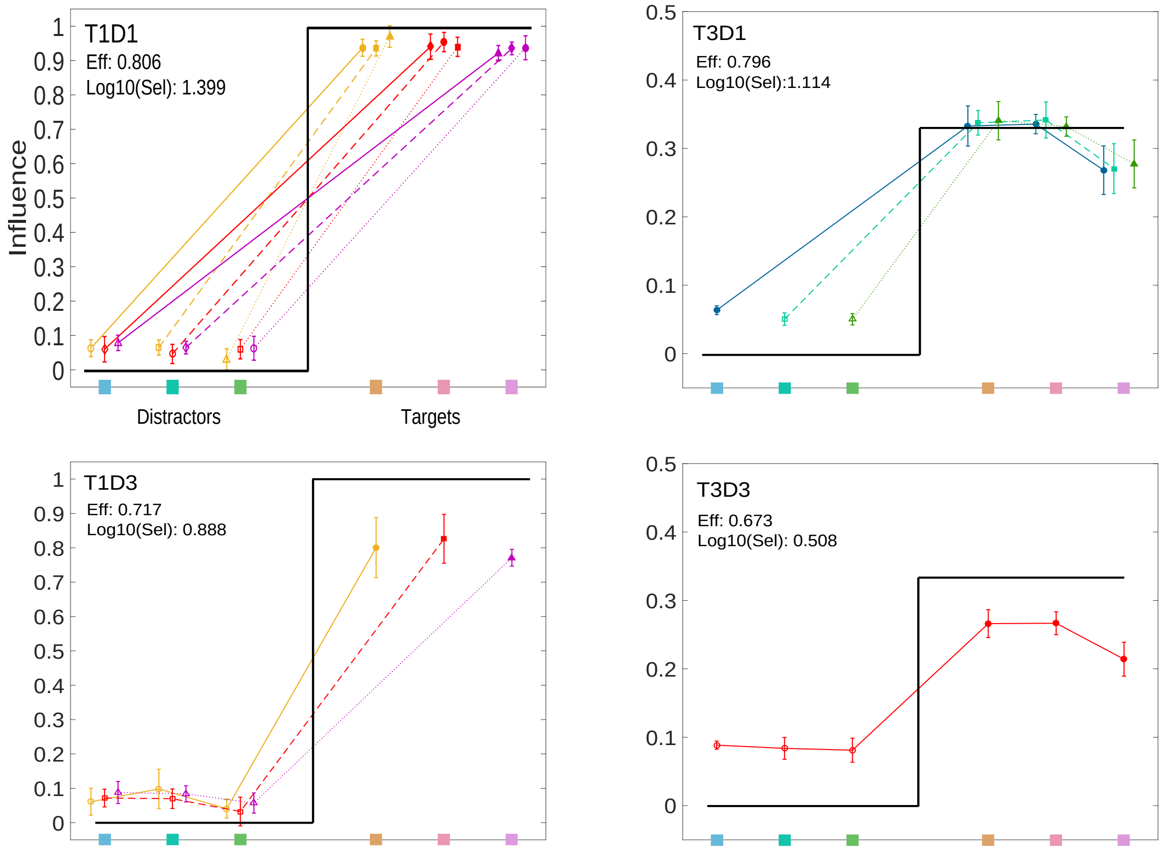


Figure 3: Influence functions, averaged across participants, for each of the four tasks. Task T1D1 consists of 1 target hue and 1 distractor hue on each trial, Task T3D1 consists of 1 distractor hue and 3 target colors on each trial, Task T1D3 consists of 1 target hue and 3 distractor hues on each trial, and Task T3D3 consists of 3 target hues and 3 distractor hues are displayed 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 participants removed and then adjusted to eliminate bias as described by Franz & Loftus (2012) and Morey (2008).

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 will we do below. Here, we note that, in the T3D3 and T3D1 tasks, the purple target (rightmost on the horizontal axis in each panel of Fig. 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 homogenous. 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 hues 3 (yellow-green) and 4 (orange) were paired or when hues 1 (blue-green) and 6 (purple) were paired. Likewise, we would have expected the best performance to emerge when hues 3 (green) and 7 (red) were paired, but none of these conditions were significantly different from one another. This non-effect of hue pairs confirms that performance in the centroid task is not driven 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 were more salient than any other.

As noted above, the important differences between tasks are best captured using log10 selectivity and efficiency. These measures, averaged across participants in each task, are summarized in Tables 4 and 5.

Log$_{10}$(selectivity):

		Target Hues			
		T1 (1 hue per trial)	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 log$_{10}$(selectivity) for four tasks, together with the confidence intervals of the values averaged across all 8 participants. 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 log$_{10}$ transformation.

Efficiency:

		Target Hues			
		T1 (1 hue per trial)	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 8 participants. 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*Efficiency, which give 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.

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 effects of 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 effects of 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 vs. 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' $\log_{10}$ 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 in the T3D3 task (for selectivity, Δ = 0.318, t(7) = 1.861, *p* = 0.112, BF = 0.91), but what difference there were did show that experts still performed better than novices on average on this task. There were 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 vs. 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 heterogenous conditions worsened as much as the novices'. This suggests that the effects are not task-specific.

Study 1a: Discussion

Results revealed that target and distractor heterogeneity both degrade performance in the centroid task, but that target heterogeneity has 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 participants and hues is 25.1 and selectivity for T3D3 is 3.22 (Table 4); this is a large difference, a factor of 6.8. The other two conditions are of intermediate selectivity, with selectivity reduced by almost twice as much due to increasing target heterogeneity (i.e. T1D1 versus T3D1: 3.6) than to increasing distracter heterogeneity (i.e. T1D1 versus T1D3: 1.9). In the logarithmic analysis, the interaction is small and not reliable, suggesting that separate effects of target and distracter heterogeneity combine multiplicatively (additively in the log domain) to produce the overall selectivity. This 2-factor log-linear model accounts for 99.5% of the variance in $\log_{10}$ Selectivity. For efficiency, the 2-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 are 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 only found when target heterogeneity was high.

Although we have strong evidence that target heterogeneity impacts 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 and/or the distractors were homogenous. 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 optimizing in some way 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 their performance in the separately blocked tasks reported above. 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, then 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 $\log_{10}$ 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 homogenous, even though participants could not predict the target to which they would have to attend on each trial. Similarly, when distractors are homogeneous, they were easier to ignore. This is consistent with findings from previous research that demonstrated how 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 to 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 (Danelljan, Hager, Shahbaz Khan, & Felsberg, 2015; Foley, 1994; Lee, Itti, Koch, & Braun, 1999; Ren & Malik, 2003; Zenger & Sagi, 1996). 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 token appearing on that trial.

A different way to understand these results would be grouping, a bottom-up mechanism that can segregate, pre-categorically, 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 three 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, Jolicoeur, and Cowan (1996) showed 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, the paper by Sun et al. (2016a) has 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 whether they reflect 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, Jolicoeur, & Cowan (1996), Nagy & Thomas (2003), Nagy, Neriani, & Young (2005). The results from this study do 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 visual(-motor) system appears to do naturally (Baud-Bovy & Soechting, 2001, McGowan et al., 1998). This tendency to locate the center of mass plays a role in grasping, as shown by a study conducted by Goodale, Meenan, Bülthoff, Nicolle, Murphy, & Racicot, (1994). Participants in the 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, & Lu (2010) or, more recently, Buetti, Cronin, Madison, Wang, & Lleras (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.

Study 1b: Introduction

Study 1a showed that, when they are identical attentional expectations across trials, both selectivity and efficiency were higher when the targets were homogenous on the dimensions of interest. This result is inconsistent with models in which selective, feature-based attention operates as a filter constructed based on attentional expectations. This result can be explained, however, by a model in which grouping of perceptually similar stimuli precedes selection. The purpose of study 1b was to introduce more closely spaced colors into the target set in order to determine how sensitive this grouping mechanism is to target variation; in other words, how far apart can the targets be spaced in feature space but target tokens of different types are still grouped together? To test this, we kept the same three target colors used in study 1 and added seven evenly spaced colors in between the central color and the colors to each side of it, for a total of 17 possible target colors. Each trial only displayed three of the possible seventeen colors, but participants did not know ahead of time which set to expect.

We refer to distances between the colors in the color circle as spacing 0-8, and the conditions are named as such (i.e. Spacing 0, Spacing 1, Spacing 2, etc…). Because performance in study 1a showed that participants were significantly better in the T1D1 task than the T3D1 task, we expect participants in this study to have higher selectivity and efficiency in conditions with smaller spacing between colors. As the space between target colors increased (i.e. the colors become more noticeably different from one another) we expected grouping to be less effective, as reflected in lower efficiency and selectivity. We also expected the influence functions to fall further away from ideal as the spaces increase. Under the assumption that grouping occurs more readily and cleanly when multiple target items types are indistinguishable from one another, we might expect to see a categorical change – i.e. something like a step function – as the spacing is increased. With sufficiently small spacing, the results would essentially be identical to Spacing 0; with larger spacings, the results would be essentially identical to Spacing 8.

Study 1b: Design

Study 1b was designed to be like an extended version of the T3D1 task from study 1. The colors used in this condition were the same three reddish targets and the middle green distractor from the first study, in addition to the 14 hues equally spaced between the outermost target hues (Figure 4). These extra 14 hues were chosen using the same minimum motion paradigm that was used to obtain the colors in study 1. As mentioned above, each condition was defined by how many spaces were between the target colors in color space. For example, in the Spacing1 condition, targets consisted of the middle red color (present in all conditions) and the two colors right next to it in the color space. We also included a Spacing 0 condition that consisted of only the one middle target red color, which mimics the T1D1 task in the first study.

In all conditions, 24 dots were presented: 12 targets and 12 distractors. Every condition consisted of 12 dots of the one distractor color and four dots for each of the three target colors, except for the Spacing 0 condition, which consisted of 12 dots of the one target color and 12 dots of the one distractor color. Participants completed 1650 trials divided into blocks of 50. All conditions were intermixed randomly, so participants were unable to predict which targets would be present from trial to trial. Three experienced participants completed the study.

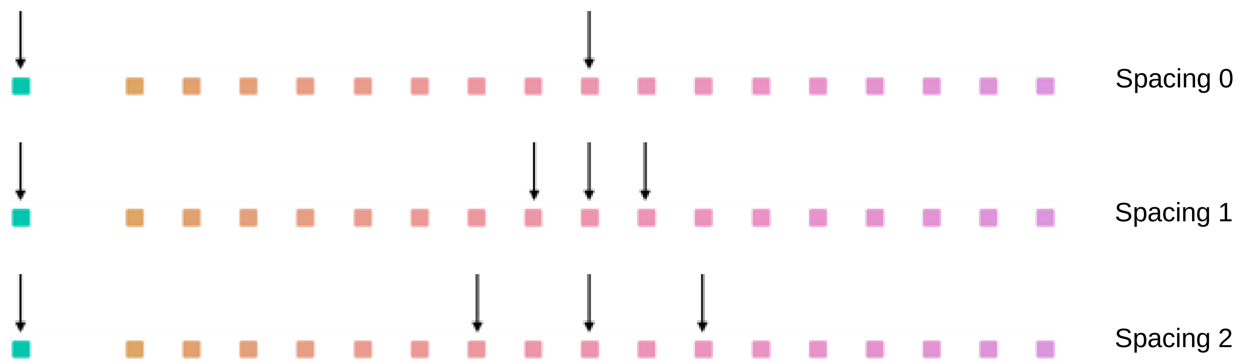


Figure 4: Examples of how colors were chosen for each trial. In all conditions, the distractor was green. In the Spacing0 condition (top), only the center target color was present. In the Spacing1 condition (center), three target colors were present: the middle red, and the two reddish hues one space to either side of it. In the Spacing2 condition (bottom), three target colors were present: the middle red, and the two reddish hues two spaces to either side of it.

Study 1b: Results

Each participants' influence function for the nine space conditions are displayed in Figure 5. By simply comparing the shape of the influence plots for the participants, it is clear that they used different strategies to try to give the most weight to the targets dots, and that none of them showed anything like a categorical change in any measure as spacing increased. Participant 1 tended to give the most weight to the middle target color and less weight to the side target colors, whereas participants 2 and 3 gave approximately equal weight to all three target types. Participants 2 and 3 also tended to give more weight to the middle target color, but not to the same degree.

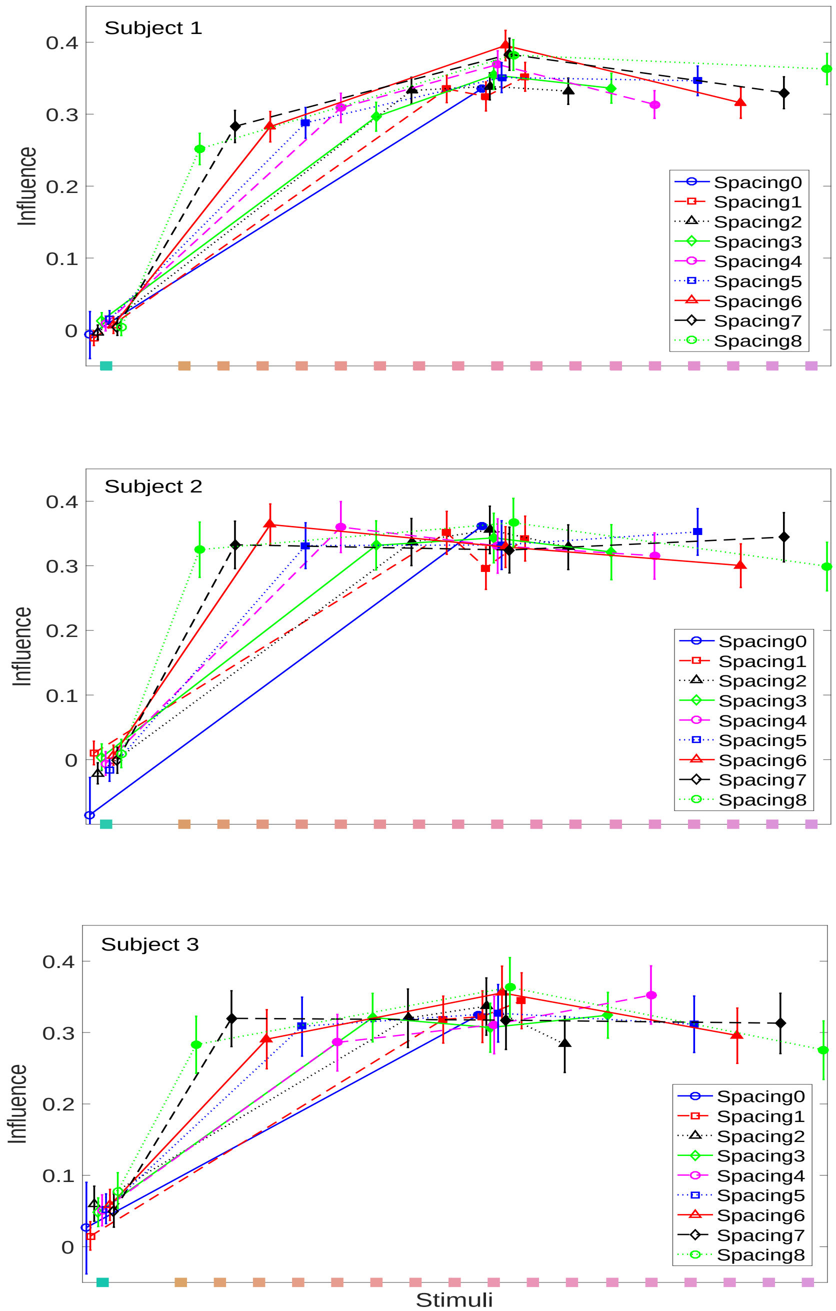
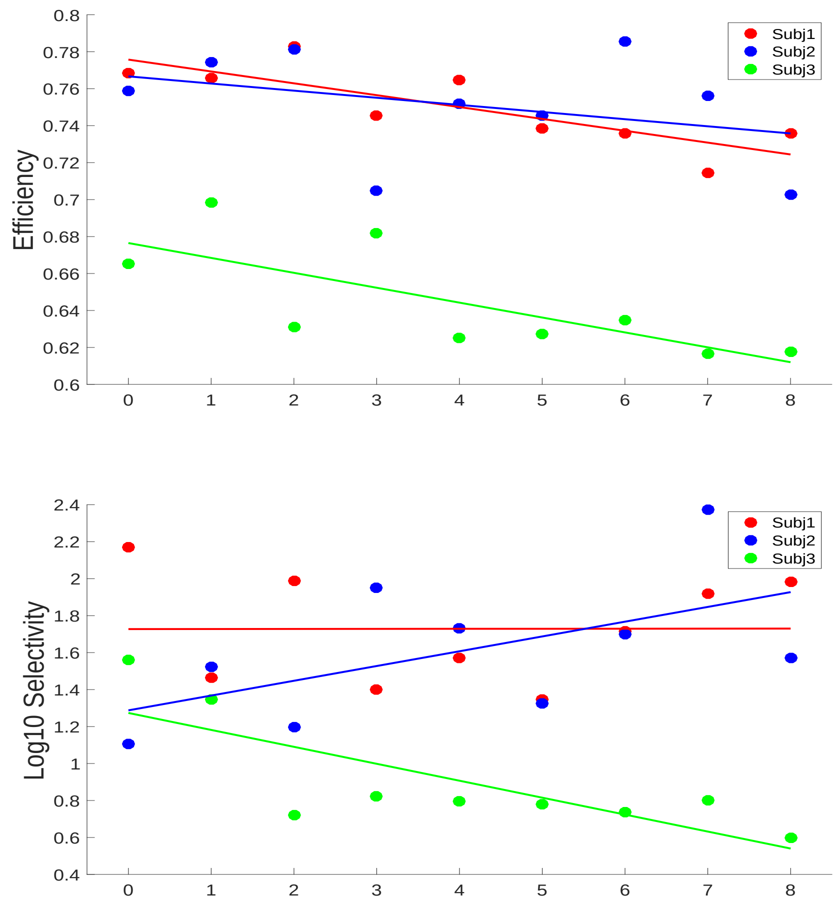


Figure 5: Influence plots for the three participants in the experiment. Condition Spacing 0 is included in the plots, but its target weight is reduced by a third to normalize it to the target weights of the rest of the conditions.

In this experiment, we looked at how spacing between the colors affected efficiency and log selectivity, but we also included a new measure: the influence ratio. This ratio is calculated by dividing the weight of the target color by the average of the weight of the colors on either side of the middle. It is similar to the selectivity ratio, except this measure only takes the targets into account, whereas the selectivity ratio divides the average weight of the targets by the sum weight of the distractors. If participants were able to group all of the target dots together effectively, we expected the influence ratio to be close to one.



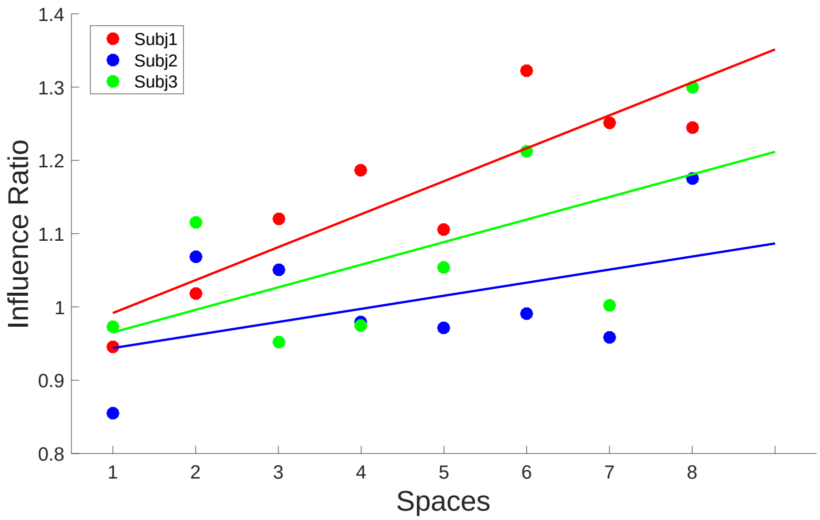


Figure 6: Plots for efficiency, selectivity ratio, and influence ratio for all three participants. A regression line is fitted to each set of data. Only data for spaces 1-8 are shown for the influence ratio because only one target color was present in the S0 condition.

The size of the space between colors tends to have a different effect on each participants' efficiency, selectivity, and influence ratios (Figure 6). We calculated a simple linear regression for efficiency, selectivity, and influence ratios based on space sizes. Slopes for each participant and measure are displayed in Table 6. A significant slope was found for efficiency for participants 1 and 3 ($t(7) = 3.750$, $p < 0.05$, BF =2.196), with an R^2 of 0.668; and $t(7) = 2.880$, $p < 0.05$, BF =1.644, with an R^2 of 0.542, respectively), but not for participant 2 ($t(7) = 0.992$, $p > 0.05$, BF =0.658 with an R^2 of 0.123). A significant slope was found for selectivity in participant 3 ($t(7) = 3.315$, $p < 0.05$, BF = 1.916, with an R^2 of 0.611), but not for participants 1 or 2 ($t(7) = 0.008$, $p > 0.05$, BF =0.467 with an R^2 of $9.134e^{-6}$; and $t(7) = 1.779$, $p > 0.05$, BF = 1.010, with an R^2 of 0.311, respectively). Lastly, a significant slope was found for influence ratios in participant 1 ($t(7) = 4.353$, $p < 0.05$, BF = 2.590 with an R^2 of 0.760), but not in participants 2 or 3 ($t(7) = 1.292$, $p > 0.05$, BF = 0.777 with an R^2 of 0.218; and $t(7) = 1.813$, $p > 0.05$, BF = 1.028 with an R^2 of 0.354, respectively).

	Efficiency	Log Selectivity	Influence Ratio
Participant 1	-0.0064 [-0.0105 -0.0024]	0.0003 [-0.0973 0.0979]	0.0450 [0.0197 0.0702]
Participant 2	-0.0039 [-0.0131 0.0053]	0.0800 [-0.0263 0.1863]	0.0178 [-0.0159 0.0516]
Participant 3	-0.0081 [-0.0147 -0.0014]	-0.0917 [-0.1570 -0.0263]	0.0308 [-0.0108 0.0724]

Table 6: Slopes for regression lines for each measure in each participant. Values in brackets represent 95% confidence intervals.

Study 1b: Discussion

From these results, we have concluded that, if participants perceive the targets in three groups, then the locations of those groups can be averaged without adding much, if any, error. Although the observed influence functions are reasonably precise, there was no consistent pattern in their variation with spacing across the three participants, which indicates that each participant employed a different strategy for the task. We found that no clear step function could describe the decline in efficiency for participants 1 and 3 as spacing increased, which is what we would have expected to find if the participants had deployed categorical grouping. We also found that distractors received more influence as spacing increased, but this only occurred for our least experienced participant. This resulted in a decrease in the participant's selectivity, which reflects how they were less able to discriminate the distractors form the targets as spacing increased. Participants may be grouping the targets into three distinct groups and averaging the positions of all three groups in order to find the centroid. The error in performance may have increased due to this step of taking the average position of the groups.

While participants do seem able to use large color differences to categorize dots into separate distinct groups, as seen in Study 1a, they do not seem able to categorize them when smaller color differences. This results may not be as surprising as we originally thought, as previous studies have proposed that linguistic color categories modulate color perception (Bornstein and Korda, 1984), and that observers who are familiar with these categories tend to distinguish colors more reliably if the color falls in the middle of the category, rather than at the boundaries (Witzel & Gegenfurtner, 2018). Because the colors used in the study blurred the boundaries between the three target hues from our color wheel (see Chapter 2), participants may have been able to categorize the colors into one group instead of three.

CHAPTER 4

STUDY 2 – PITTING TOP-DOWN AGAINST BOTTOM-UP MECHANISMS OF FEATURE-BASED ATTENTION

Study 2: Introduction

Feature-based attention has two different forms: top-down, or endogenous, attention and bottom-up, or exogenous, attention. While top-down attention relies on instructions or intention, bottom-up attention relies solely on the salience of the items. The current experiment aims to study how the two processes interact with one another – are their effects additive, as we might expect if there were two separate processes, or do they operate on a winner-take-all principle? Previous work using other paradigms has have provided evidence that attention is driven by only one mechanism at a time (Posner, 1980; Theeuwes, 1991; Reynolds, Alborzian, & Stoner, 2002; Gaspelin, Leonard, & Luck, 2015), although the question of which mechanism has more influence is up for debate. We aimed to find evidence for either an additive or winner-take-all model for attention, and if we found that the winner-take-all model was more likely, we aimed to determine whether the top-down or bottom-up mechanism was more influential in directing our attention.

Work done in our lab has shown that participants are able to find the centroid of the most numerous colors in a stimulus, despite not being told ahead of time what that color would be (Sun, Chubb, Wright, and Sperling, 2018). Based on results from Study 1, one explanation for why this may be that participants' attention was drawn towards large groups of similar-looking stimuli. Grouping large clusters of similar-looking stimuli may automatically capture attention since numerosity leads to saliency, so that participants did not need to know what the target color was ahead of time. Research conducted by Gaspelin, Leonard, and Luck (2015) studied whether or not attention captured by singletons was obligatory. Although it is commonly thought that singletons must capture attention, Gaspelin et al., found that if participants learned that a singleton never contained information about the target, they could learn to ignore it. Although numerosity is the inverse of a singleton, one could expect that participants can learn to ignore the saliency of this cue also. Using the centroid paradigm, we tested the interaction between top-down ("look for the centroid of the red or green dots") and bottom-up attention ("look for the centroid of the most numerous color"). If the additive model is correct, then the two processes may work together if they both indicate that the participant should attend to one specific type of stimuli. However, if one process indicated that the participant should attend to one type of stimuli and the other indicated that they should focus on a different type, performance should worsen. If the winner-take-all model is correct, then whether or not the two processes agree will have no significant effect on performance.

Study 2: Design

Table 6 summarizes the conditions used in the experiment, organized by task: "attend to red," "attend to green," and "attend to most numerous," Three stimulus types were used in this experiment (Figure 7), each consisting of 24 dots (Table 7). Within each task, participants were unable to predict which type of stimulus they were be shown from trial to trial. The conditions (Table 8) in which top-down and bottom up attention might reinforce each other are the ones labelled "Consistent," whereas the trials in which top-down and bottom up attention might work in opposition are the ones labelled "Inconsistent". Both conditions present in the "attend to most numerous" task should reflect the operation of bottom-up visual attention without contribution from top-down visual attention beyond “attend to the most numerous stimuli”. Although this is an instance of top-down instructions, participants could not rely on it alone to find the target color. Their response was mostly driven by the stimulus itself, as they were informed of the target color only when the stimulus was presented. Conditions in which there are equal numbers of red and green stimuli should reflect the operation of top-down visual attention without a contribution from bottom-up attention. The 1440 trials run by the participants in the experiment were divided into six sessions, each testing one of three tasks. These three tasks were run in an ABCCBA order determined separately for each of the six participants by a Latin square. One participant’s data was removed from the analysis since we were unable to create reasonable influence functions from the data.

Trial Type	Number of Red Dots	Number of Green Dots	Number of dots for other distractors (six colors)
Red Numerous	9	3	2 of each
Green Numerous	3	9	2 of each
All Equal	3	3	3 of each

Table 7: Number of each color token present in each trial type.

Although there is no neutral condition against which to assess the direction of these attention effects, it is possible to compare their relative magnitudes. Comparing the Consistent with the Bottom-up Only conditions will provide an estimate of the incremental effect of adding top-down to bottom-up attention. If there is only a single mechanism of attention and bottom-up attention can activate it maximally, then we might observe no difference between these conditions. However, if there are separate mechanisms that implement top-down and bottom-up attention, then performance in conditions that specify which colors are targets, should be substantially better than that in conditions that simply ask participants to search for the most numerous color. We could have also compared the Consistent and Top-down Only conditions to observe the effect of adding bottom-up to top-down attention, but the differences may be more complicated to interpret. While it is possible to compare the $\log_{10}$ selectivity ratio between the two conditions, it would have been difficult compare efficiency, since efficiency is affected by the number of targets present in each trial. Consistent conditions included nine target dots and Top-down Only conditions only included three target dots, which complicates how efficiency is interpreted. Finally, looking at the comparison of the Inconsistent conditions and the Bottom-up conditions should provide insight into whether the effects of bottom-up attention are obligatory, i.e. whether bottom-up attention is always triggered by an external cue even when it is inappropriate. If bottom-up attention is obligatory, performance in the Inconsistent conditions should be worse than that in the Bottom-up Only conditions, and we should see a strong influence of the most numerous color, even when instructions tell participants to attend to a less numerous color.

	Red stimuli are more numerous	Green stimuli more numerous	Equal numbers of Red and Green stimuli
Attend to red stimuli	Consistent	Inconsistent	Top-down attention to red stimuli only
Attend to green stimuli	Inconsistent	Consistent	Top-down attention to green stimuli only
Attend to most numerous stimuli	Bottom-up attention to red stimuli only	Bottom-up attention to green stimuli only	NA

Table 8: Each row represents one task (row headers display the instructions given in each task), whereas each column displays the type of stimuli present in the condition.

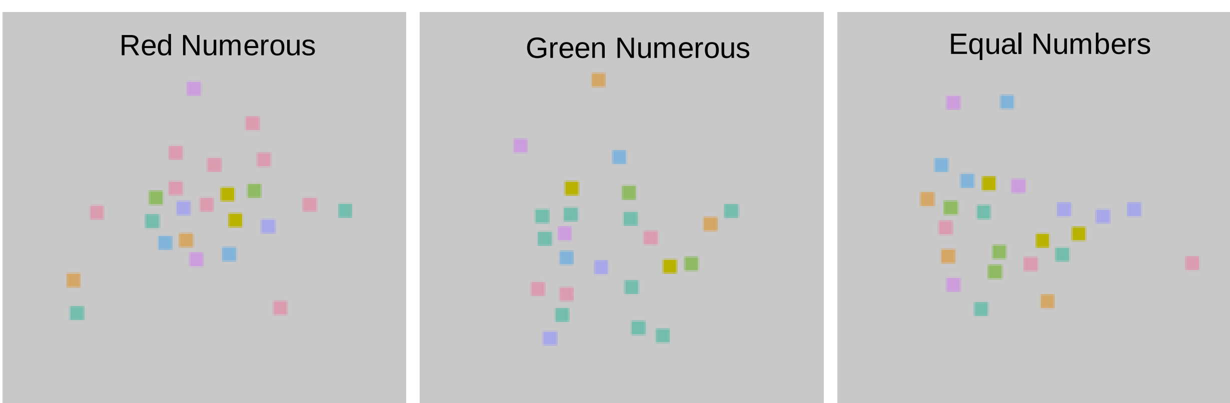
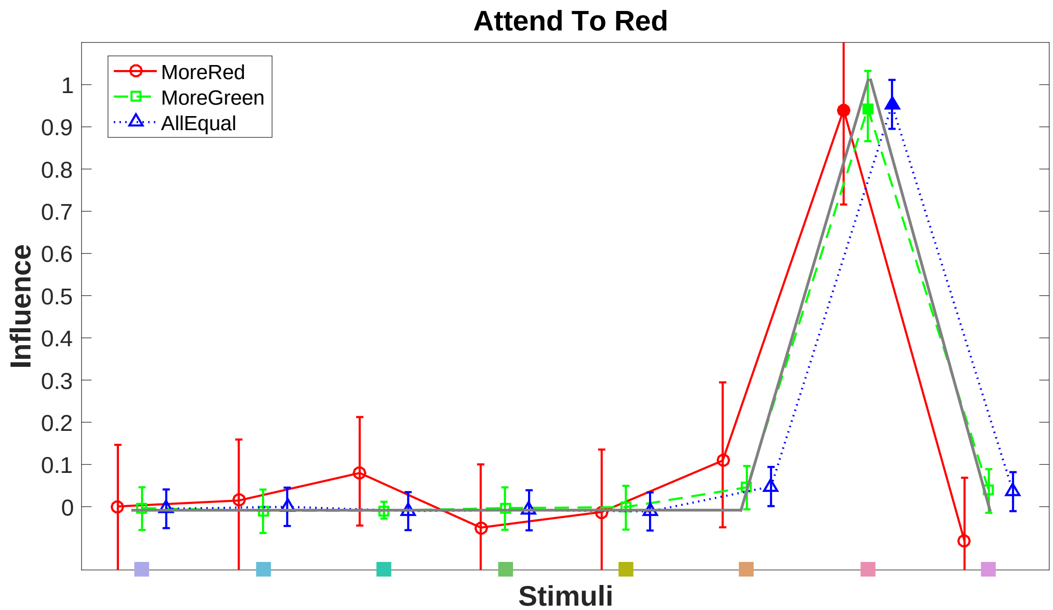
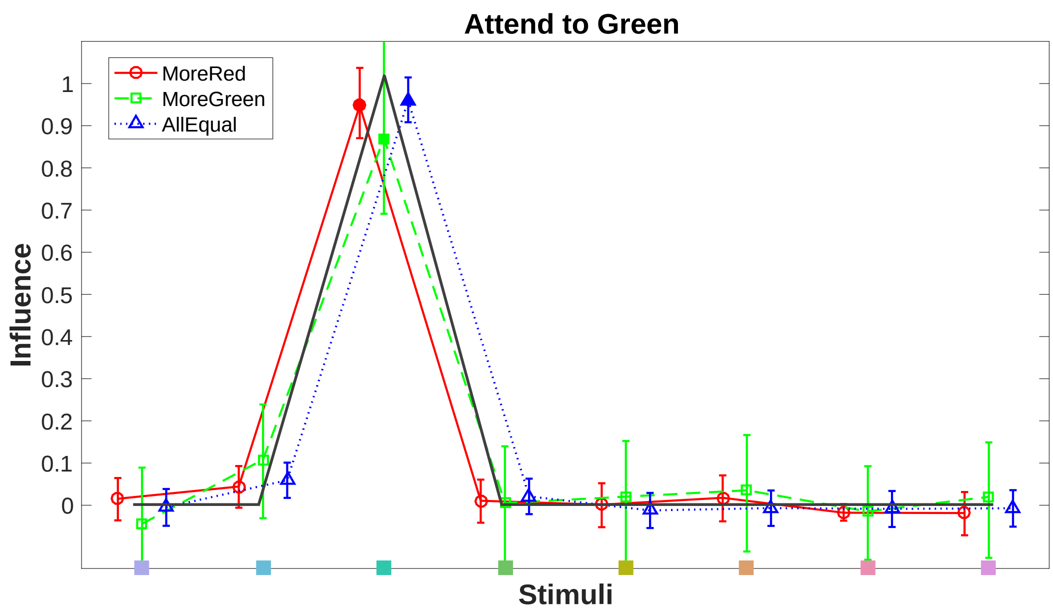


Figure 7: Sample stimuli types for the experiment.

Study 2: Results





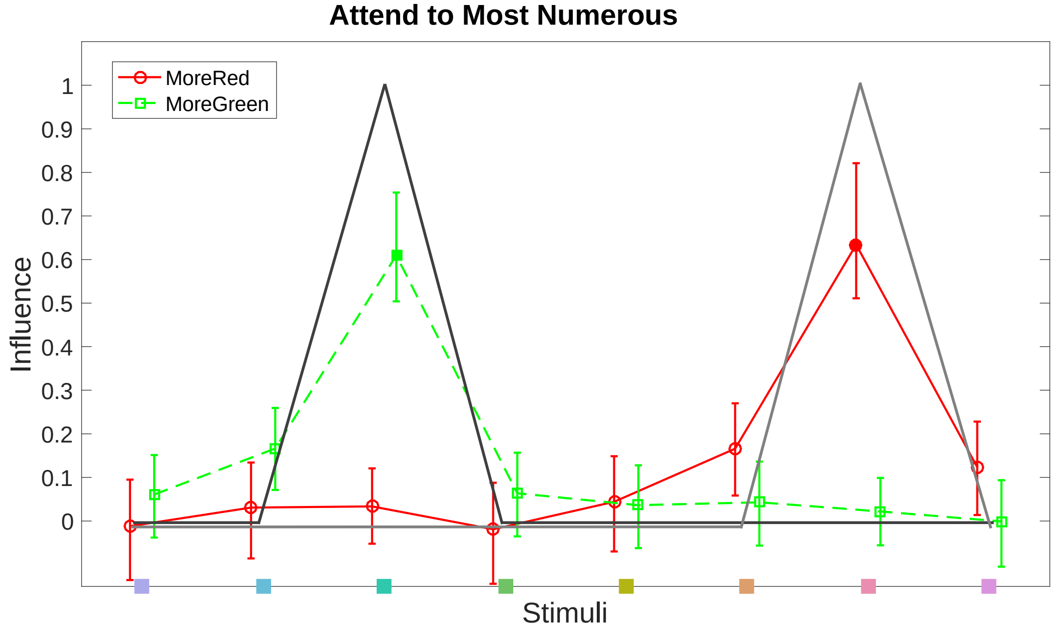


Figure 8: Influence functions for each of the tasks averaged across all participants, with gray lines indicating what the ideal influence function should look like.

Figure 8 shows that, in the Attend to Red task, the weight given to the targets was consistent across conditions, regardless of whether the number of tokens of the target type in question was the most numerous, equal to the distractor, or outnumbered by the distractor (Red Numerous versus Green Numerous : mean = -0.003 [-0.064 0.058], t(7) = -0.122, p = 0.906, BF = 0.3; Numerous versus All equal: mean = -0.010 [-0.071 0.051], t(7) = -0.384, p = 0.713, BF =0.4). The same is true in the Attend to Green task (Red Numerous versus Green Numerous : mean = 0.078 [-0.082 0.238], t(7) = 1.147, p = 0.289, BF = 0.6; Numerous versus All Equal: mean = -0.051 [-0.211 0.109], t(7) = -0.755, p = 0.475, BF =0.4) task. Additionally, the weight of the distractor color that was also the most numerous color was significantly different from the weights of the other distractors in both the Attend to Red task (mean = -0.021 [-0.041 -0.0001], t(7) = -2.381, p = 0.049, BF = 1.93) and the Attend to Green task, (mean = -0.030 [-0.047 -0.013], t(7) = -4.141, p = 0.004, BF = 12.63).

In both cases, the numerous distractor color received significantly less weight than the mean of the of the other distractors.

In the two conditions of the Attend to Red task where green was not the most numerous color, the weight of green was not significantly different from the mean of the other distractor colors (red most numerous: mean = 0.069 [-0.010 0.148], t(7) = 2.060, p = 0.078, BF = 1.36; all equal: mean = -0.0212 [-0.049 0.007], t(7) = -1.787, p = 0.117, BF = 1.02) This was also true of the red distractor in the Attend to Green task (green most numerous: mean = -0.034 [-0.120 0.042], t(7) = -1.131, p = 0.295, BF = 0.55; all equal: mean = -0.017 [-0.042 0.008], t(7) = -1.594, p = 0.155, BF = 0.840). One can observe this by looking at the plots in Figure 8. When red is the target color (first plot) but green is the most numerous (green line), the influence function of the green distractor is lower than the mean of the other distractors from the same condition. This is also true in the Attend to Green (second plot), in which the most numerous distractor has a lower attention filter than the mean of the other distractors (red line). We may have only observed a significant effect of red and green distractors asserting less weight on the centroid when they were the most numerous because the increased numerosity makes these colors more easily grouped and then rejected as a group. This effect is small, and there is no significant interaction between the most numerous distractor color and the target color (Attend to Red: mean = 0.052 [- 0.009 0.113], t(7) = 2.031, p = 0.082, BF = 1.32; Attend to Green: mean = 0.002 [-0.037 0.041], t(7) = 0.111, p = 0.915, BF = 0.34). However, we were also unable to find evidence showing that numerous distractors are weighted equally with other distractors.

Results from the third plot in Figure 8 indicate that bottom-up information can be ignored when it conflicts with top-down intentions. When instructed to, participants gave greater weight to whatever stimuli was most numerous on the screen, which supports the idea that participants can use bottom-up attention to search for the centroid. We found no significant differences between the green target and red target in the Attend to Most Numerous Task (mean = 0.024 [-0.154 0.203], $t(7) = 0.325$, $p = 0.755$, BF = 0.35) .However, the weight given to the target dots in the Attend to Most Numerous Task was smaller than the weight given to the target dots in the Attend to Red and the Attend to Green tasks (Attend to Red: mean = 0.311 [0.1657 0.465], $t(7) = 4.78$, $p = 0.002$, BF = 23.4; Attend to Green: mean = 0.056 [0.216 0.417], $t(7) = 7.410$, $p = 0.0001$, BF =197.7).

	Red stimuli more numerous	**Green stimuli more numerous**	**Equal numbers of Red and Green stimuli**
Attend to red stimuli	0.655 (12) [0.607 0.702]	0.863 (3) [0.827 0.899]	0.837 (3) [1.778 1.897]
Attend to green stimuli	0.882 (3) [0.845 0.919]	0.686 (12) [0.643 0.728]	0.868 (3) [0.829 0.907]
Attend to most numerous stimuli	0.619 (12) [0.559 0.680]	0.636 (12) [0.596 0.676]	NA

Table 9: Efficiency for each condition, averaged across all 8 participants. Each row represented one of the three tasks in the experiment. Values in parentheses represent how many stimuli of the target color were present in each trial of that condition. Values in brackets show the confidence interval of the mean.

	Red stimuli more numerous	Green stimuli more numerous	Equal numbers of Red and Green stimuli
Attend to red stimuli	1.091 [1.004 1.178]	1.108 [1.003 1.213]	1.109 [0.993 1.225]
Attend to green stimuli	1.514 [1.448 1.579]	1.472 [1.410 1.534]	1.520 [1.453 1.588]
Attend to most numerous stimuli	0.830 [0.578 1.082]	0.818 [0.693 0.943]	NA

Table 10: Log_{10} selectivity ratio for each condition, averaged across all 8 participants. Values in brackets show the confidence interval of the mean.

Tables 9 and 10 display the efficiency and log_{10} selectivity ratios averaged across all participants. The results of the influence plots indicate that the participants were able to utilize both top-down and bottom-up visual attention. This replicated previous results, which indicated that selective attention could be guided by either intentions, which corresponded with top-down visual attention, or stimulus properties, which corresponded with bottom-up visual attention. We conducted t-tests on comparisons of interest for both selectivity and efficiency. Because efficiency depended primarily on the number of items to be included in the centroid and was lower in the attend to most numerous stimulus conditions (mean = 0.214 [0.156 0.271], $t(7) = 11.861$, $p = 0.001$, BF = 2582.03), we were only able to compare conditions that displayed the same number of target dots in each trial. The results from the t-tests indicated no significant differences across the critical comparisons for efficiency for both conditions from the Attend to Red task (All Equal versus Green Numerous: mean = 0.025 [-0.010 0.061], $t(7) = 1.702$, $p = 0.133$, BF = 0.9;

Most Numerous versus Red Numerous: mean = -0.036 [-0.093 0.022], $t(7) = -1.459$, $p = 0.188$, BF =0.7). When comparing the conditions from Attend to Green task, there were also no significant difference between the All Equal versus Red Numerous conditions (mean = 0.014 [-0.002 0.030], $t(7) = 2.025$, $p = 0.083$, BF = 0.8). However, when comparing the Most Numerous and the Green Numerous conditions in the same task, the difference was significant (mean = -0.050 [-0.085 -0.014], $t(7) = -3.332$, $p = 0.013$, BF = 5.5).

Looking at selectivity, we also found a significant difference for one of the comparisons. In conditions where green was the most numerous color, participants were more selective for green when they were told ahead of time that they should focus on green (mean = 0.653 [0.507 0.800], $t(7) = 10.515$, $p = 0.00002$, BF =1312.7), which shows that adding top-down attention to bottom-up attention yielded higher selectivity than either mechanism working alone. In other words, adding top-down attention to bottom-up attention substantially improved the participant's selectivity. In the red numerous condition, there was a difference in the same direction, but it was not statistically significant (mean = 0.261 [-0.046 0.569], $t(7) = 2.010$, $p = 0.084$, BF =0.8). We also found that there was no significant difference between the Most Numerous conditions and the Equal Numbers conditions (Red target: mean = 0.018 [-0.116 0.152], $t(7) = 0.318$, $p = 0.760$, BF =0.4; Green target: mean = 0.048 [-0.057 0.154], $t(7) = 1.085$, $p = 0.318$, BF =0.5), which provides no significant evidence that adding bottom-up attention to top-down attention improves selectivity. Finally, we found that there was no significant difference in selectivity between the Inconsistent and Top-down Only conditions when participants were asked to attend to green (Red Numerous vs All Equal: mean = 0.006 [-0.009 0.022], $t(7) = 0.983$, $p = 0.358$, BF = 0.5), or to red (Green Numerous vs All Equal: mean = 0.001 [-

0.029 0.032], $t(7) = 0.106$, $p = 0.919$, BF = 0.3). Comparing the average selectivity for the two conditions in the third column and the two conditions in the third row, we found that when green was the target, pure top-down conditions lead to significantly stronger selectivity than pure bottom-up conditions (mean = 0.702 [0.514 0.889], $t(7) = 8.850$, $p = 0.00004$, BF =508.7); when red was the target selectivity for the top-down condition was again larger, however, this difference was not statistically significant (mean = 0.279 [-0.083 0.641], $t(7) = 1.825$, $p = 0.111$, BF = 1.1). Although the stimulus clouds are not identical across these conditions, these comparisons do demonstrate the relative effectiveness of selective attention in these two "pure" cases.

Study 2: Discussion

While top-down attention is driven by our intentions (or in this experiment, by instructions) bottom-up attention is driven by the stimulus itself. We have evidence from previous experiments demonstrating that participants are able to use both mechanisms of attention to guide their responses to the centroid of the stimulus cloud (Sun et al., 2018). One may expect these mechanisms to be at odds with one another if the instructions of the task were to attend to a stimulus type that does not immediately grab our attention as well as another stimulus type, and for selectivity and efficiency to suffer as a result. One may also expect the effects of both mechanisms to be additive if cues from the display point attention towards one specific stimulus type. Our experiment demonstrates, however, that these results are not true in all cases.

When guidance from top-down intentions and bottom-up stimulus properties conflicted, the intention overrode the salient stimulus properties without a large reduction in either selectivity or efficiency. We also found that in both top-down tasks (Attend to Red and Attend to Green), the target color had the same influence function across all three conditions in each task (Red Numerous, Green Numerous, and All Equal). The two target colors in the one bottom-up task (Attend to Most Numerous) also had similar influence functions. However, the influence functions of the targets in the top-down attention tasks were significantly higher than the influence functions of the targets in the bottom-up task. Although the bottom-up mechanism can be used to drive attention, it does not do so as efficiently as the top-down mechanism.

Participants had highest selectivity in conditions instructed them to search for a target color that also happened to be the most numerous stimulus type. This suggests that adding top-down attention to bottom-up attention mechanisms improved their ability to attend to the target. This effect only goes one way, however. There was no significant evidence that adding bottom-up attention to top-down attention improved selectivity. Although the stimulus clouds are not identical across these conditions, these comparisons do demonstrate the relative effectiveness of selective attention in these two “pure” cases. Numerosity seems to not only affect the targets, but also the distractors. The weight of the distractor color that was also the most numerous color has significantly lower influence functions than the mean of the other distractors. This is consistent with findings from study 1. Participants achieved the highest selectivity when not only the targets were homogeneous, but also when the distractors were homogeneous. In these two cases, homogenous dots were more numerous than heterogeneous dots, which allowed participants to generate higher selectivities and efficiencies. The effect is small in the current study, but there is also no significant evidence to argue against it. From these results, it seems that grouping does help eliminate distractors.

One concern is that in the Attend to the Most Numerous task, the target color changed from trial to trial, whereas in the Attend to Red or the Attend to Green tasks, the target color remained the same across trials within a block. This may be a confound, as the expectation for each possible target color in the Attend to the Most Numerous task is less than the expectation for a specific target color in the Attend to Red or Green tasks. However, results from the T1D3 task in study 1 indicates that even when participants do not know which target color to expect from trial to trial, they were still able to achieve high selectivities and efficiencies.

These results indicate that attention does not consist of two separate mechanisms that are competing with one another to guide attention to a certain stimulus. A model that would be consistent with these findings posits that attention is a single mechanism that can be guided either by a weaker, and non-obligatory bottom-up control or a stronger, top-down command. Although participants were able to attend to the target color when it was only identified during the short duration time as the most numerous color, their selectivity for the target in this task was not as high as their selectivity for the target in the tasks that stated which color would be the target color at the beginning of the block. In fact, when participants are aware of exactly what color they should attend to, they were more selective for that color regardless of whether it was the most or least numerous. The bottom-up mechanism may be obligatory if it is the only cue that helps a participant search for a target, but it is not as strong a cue as top-down instructions.

CHAPTER 5

STUDY 3 – VARIATIONS IN IRRELEVANT FEATURES

Study 3: Introduction

From Study 2, we have concluded that there may be a mechanism grouping stimuli together based on their similarity to each other. It should be noted that the feature that they are grouped by is the feature that participants are meant to attend to in the task. A reasonable follow-up study to conduct would explore how sensitive grouping is to the distinction between task-relevant and task irrelevant features. To what degree does grouping on an irrelevant feature dimension disrupt grouping on the relevant feature dimensions? Evidence from past research in our lab indicates that distractors in irrelevant feature dimensions do have an effect on performance, but that effect has not been well-documented.

In a study conducted by Lunau and Habekost (2017), in which color was the irrelevant feature and character type (letters or digits) were the relevant feature, participants were unable to ignore color. Participants were asked to identify what letters were presented in a display, while ignoring the distractor digits. When letters and digits were each printed in a homogenous color (e.x. letters were all red and digits were all blue), then participants were able to report an average of 2.91 letters out of a total of 4. When letters and digits were printed in heterogenous colors (i.e. both letters and digits appeared in red and blue), then participants were only able to report an average of 2.78 letters. Color did not interfere when the feature of interest is size, however. The authors noted at the short display durations that they used (100 ms and 200 ms) participants may not have had enough time to group items together, so the effects they observed were small. There was a minor trend that showed that color grouping took effect at 200 ms, however, which led the researchers to predicted that at longer display durations, participants would have been able to group the stimuli by color, and performance would have been significantly worse.

In the first study, we had found that distractor heterogeneity did have a significant impact on performance, but a small one. Our study aims to determine whether obligatory grouping along an irrelevant dimension disrupts the selectivity of attention on a relevant dimension or adds error to the centroid calculation that reduces efficiency. The design of this experiment is similar to that of Study 2, but instead of having participants ignore variation on a relevant dimension among the targets, they will be asked to ignore variation on an irrelevant dimension. Taking Lunau and Habekost's (2017) results into consideration, our study will use display durations of at least 200 ms to ensure that participants are given the opportunity to group the items in the display. Previous research conducted in our lab has shown that performance is affected by features on an irrelevant dimension, but these effects were not carefully documented, as that was not the focus of those studies.

If the possibility of grouping on the irrelevant dimension is disruptive of our ability to selectively attend to target, perhaps by segregating them into separate groups from the distractors on the relevant dimension, then we would expect participants to have difficulty ignoring these irrelevant features. If we find that the irrelevant feature is not disruptive, we will explore whether participants were able to ignore grouping on an irrelevant feature, or whether these two types of groups coexist, and participants were able to select which one they wanted to attend to. They may be able to ignore the irrelevant dimension, but if there is an effect of grouping on that dimension, we may observe the largest effects in the condition in which the irrelevant dimension variation is large and target dimension variation is small.

Study 3: Design

The experiment followed the same procedure described in the methods section from chapter 2 Dots varied by both color and size, and participants were told to click on the centroid of the target feature, ignoring the irrelevant one. Each display consisted of three colors, with six dots assigned to each color. Within each of those color groups, the six dots were divided into three groups: large, medium, and small. These three sizes were represented by two dots each. There was a total of 18 dots in each trial – two copies of each of the nine unique dots (i.e. two small color 1 dots, two medium color 1 dots, two large color 1 dots, two small color 2 dots, etc…). In the conditions with only one color or one size, all 18 dots were all presented with that color or that size. The five colors used in the experiment were selected from the 64-color space used in study 1b. The target color was always a middle red (Color 3 in Table 11). In the condition with close spacing of the colors, the two distractor colors were chosen to be the colors four steps on either side of the target (Colors 2 and 4 in Table 11). In the large color spacing condition, the two distractors were chosen to be the colors four on either side of the target (Colors 1 and 5 in Table 11).

Colors	x-coordinate	y-coordinate

Color 1	0.4457	0.4661
Color 2	0.4358	0.3750
Color 3	0.3936	0.3047
Color 4	0.3463	0.2666
Color 5	0.3004	0.2543

Table 11: The five colors used in study 3.

From experiment 1b, we found that it is easier to discriminate colors that are eight steps away from each other than colors that are only four steps away from each other, but participants are still able to discriminate colors four steps apart. We hypothesized that, if grouping by size cannot be ignored, the disruption caused by size grouping would be larger in the condition with close color spacing of the targets and distractors (Table 12).

	Small Variation in Size	Large Variation in Size
Small Variation in Color	Medium	Strong
Large Variation in Color	Weak	Medium

Table 12: Predictions for strength of grouping in each condition. If irrelevant features are difficult to ignore, then grouping based on size is predicted to have the strongest effect in the condition with small color variation in large size variation, since it would create the most distinct groups. In the condition with large color variation and small size variation, the effect of grouping based on size is predicted to be the weakest, since groups made based on color would be more distinct from one another. The remaining conditions were predicted to have effects of grouping that were not as strong or as weak as the extreme cases.

The study consisted of two tasks: small variation in color (four spacing between colors) and large variation in color (eight spacing in color). In both tasks, stimuli were presented with no variation in size (one size - 0.32 degrees of visual angle), small variation in size (three sizes – 0.24, 0.32, and 0.36 degrees of visual angle) or large variation in size (three sizes – 0.16, 0.32, 0.44 degrees of visual angle). Sample display screens are presented in Figure 9. At the beginning of each session, participants were asked to locate the centroid of the target-color items while ignoring the items presented in either of the distractor colors and also to ignore the variation in size of the items. Four participants recruited for the experiment. They completed one session of the small variation in color task, one session of the large variation in color task, and two sessions wherein the two tasks were mixed together. If completing the two tasks in separate blocks had an effect on performance, the mixed sessions would eliminate these effects. Participants completed the four sessions in the order given to them, determined by a Latin square, then repeated them in reverse order. Each session consisted of one block of 140 trials.

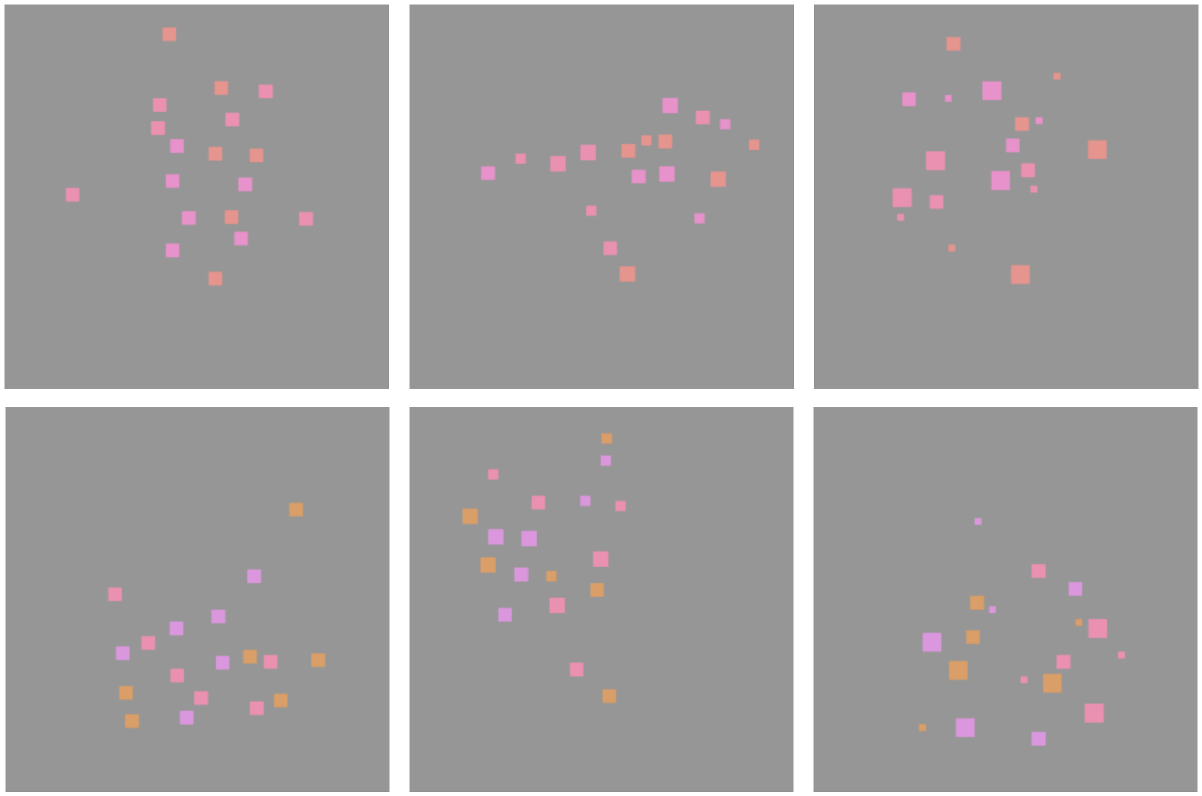


Figure 9: Sample stimuli for experiment. The rows display the two color conditions (top row shows small variation in color, bottom row shows large variation in color), whereas the columns display the three size conditions (left column shows no variation in size, middle column shows small variation in size, and right column shows large variation in size).

Study 3: Results

We fit the influence function data to a multiple regression model with five parameters, shown in Table 13. The table also includes the design matrix coefficients associated with each factor. The model was fit separately to the influence functions of the nine stimulus types in each condition for each participant. Parameter estimates from the model are displayed in Table 14. Comparing the estimates to the influence functions Figure 10 shows how well the model fit the data for each participant in each condition, with lines representing the influence functions and dots representing the predictions. The mean residual error for the model across all conditions was 0.0223.

Parameter									
Mean Distractor Influence	1	1	1	1	1	1	1	1	1
Mean Target Influence Increment	0	0	0	1	1	1	0	0	0
Distractor Color Direction (Yellow or Blue)	-1	-1	-1	0	0	0	1	1	1
Slope of the Effect of Size Variation on the Target Influence	0	0	0	-1	0	1	0	0	0
Slope of the Effect of Size Variation on the Distractor Influence	-1	0	1	0	0	0	-1	0	1

Table 13: Parameters included in the 5-parameter fit to the data from each participant.. The rows of stimuli at the top of the table are grouped to show the four conditions presented in the study. From top to bottom, the conditions are small variation in color with small variation in size, small variation in color with large variation in size, large variation in color with small variation in size, and large variation in color with large variation in size.

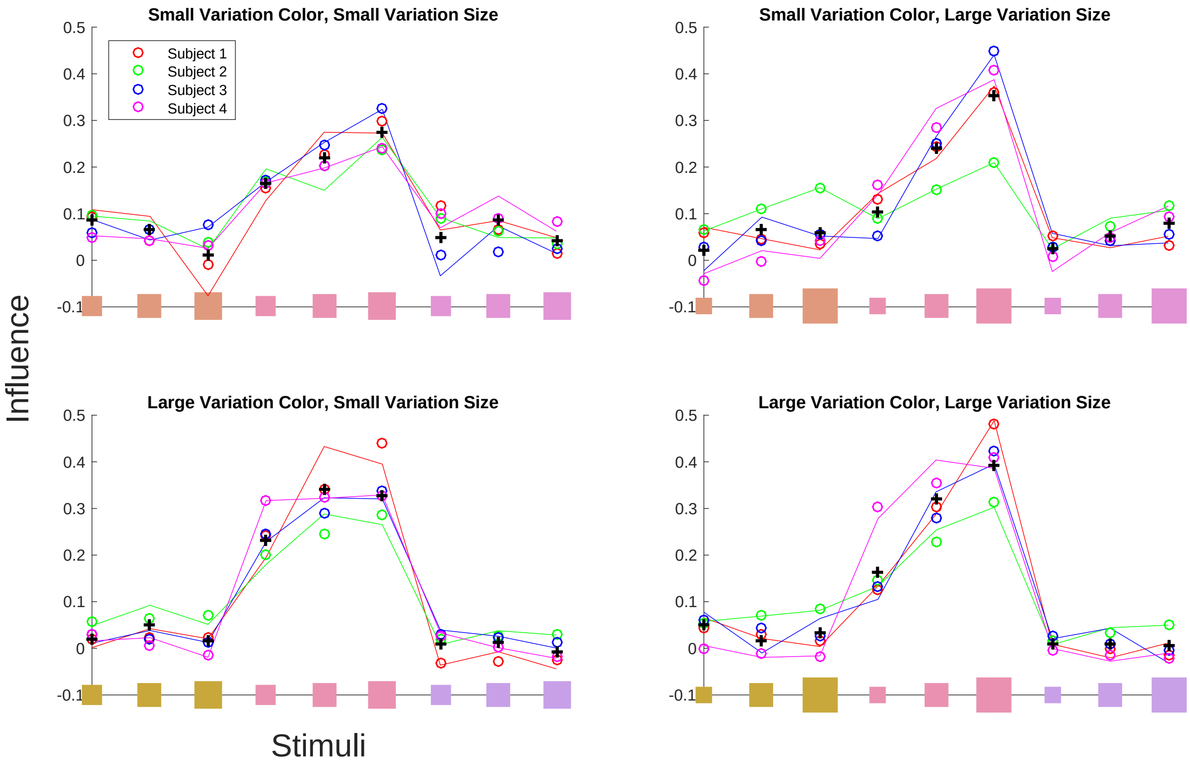


Figure 10: Influence functions (lines) and predicted data from the model (dots). Each plot displays data from one of the four tasks, with each line representing one participant. Black "+" represent the mean of the estimates for a particular stimulus, averaged across participants.

Effects of both the magnitude of the difference between the color of the targets and distractors and the magnitude of the irrelevant variation in size for each parameter are summarized in Table 14. Because there were not significant differences between mixed and blocked trials, they were combined in the analyses. The difference in size variation did not have a significant effect on either the mean distractor influence (Table 14a) or the mean target influence increment (Table 14b). The difference in color variation, however, did have a significant effect in both parameters (Tables 14a and 14b), with the mean distractor influence being smaller and the mean increment in the influence of the targets over the distractors being larger when the color difference between the targets and distractors was larger. There were no significant interaction between color and size variation in either parameter (Tables 14a and 14b). Differences in color and size variation do not have significant effects on the distractor color direction of the stimuli (Table 14c). There is also no significant interaction between color and size variation for distractor color direction (Table 14c).

Comparing the slope of the effect of size variation on the target influence and the slope of the effect of size variation on the distractor influence, we find that color and size do not have significant effects on size for the target influence, and there is no significant interaction between variation in size and color (Table 14d). There are no significant main effects of color and size variations on the effect of size for the distractor influence, but there is a significant interaction between color and size variations on distractor influence (Table 14e). When the variation in size was small, the effect of size increased when the variation in color increased. However, when the variation in size was large, the effect of size decreased when the variation in color increased. The effect of color and size variation for target influence is greater than the effect of color and size variation for distractor influence (Table 14e).

a. Mean Distractor Influence

Color Variation		Size Variation: Small	Size Variation: Large	Mean
	Small	**0.057** [0.040 0.074]	**0.050** [0.005 0.096]	**0.054**
	Large	**0.017** [-0.017 0.050]	**0.021** [-0.021 0.062]	**0.019**
	Mean	**0.037**	**0.036**	**0.036***

Color Effect (Large – Small): **-0.035*** [-0.057 -0.013]

Size Effect (Large – Small): **-0.001** [-0.032 0.030]

Interaction ($Size_S Color_S$ – $Size_S Color_L$) – ($Size_L Color_S$ – $Size_L Color_L$): **0.005** [-0.012 0.022]

b. Mean Target Influence Increment

Color Variation		Size Variation: Small	Size Variation: Large	Mean
	Small	**0.163** [0.111 0.215]	**0.182** [0.045 0.320]	**0.173**
	Large	**0.283** [0.182 0.384]	**0.271** [0.146 0.397]	**0.277**
	Mean	**0.223**	**0.227**	**0.225***

Color Effect (Large – Small): **0.105*** [0.039 0.170]

Size Effect (Large – Small): **0.004** [-0.089 0.097]

Interaction ($Size_S Color_S$ – $Size_S Color_L$) – ($Size_L Color_S$ – $Size_L Color_L$): **-0.016** [-0.067 0.036]

c. Distractor Color Direction (Yellow or Blue)

		Size Variation: Small	Size Variation: Large	Mean
Color Variation	Small	**0.002** [-0.031 0.036]	**0.002** [-0.027 0.031]	**0.002**
	Large	**-0.011** [-0.032 0.009]	**-0.013** [-0.025 -0.001]	**-0.012**
	Mean	**-0.005**	**-0.006**	**-0.005**

Color Effect (Large – Small): **-0.014** [-0.037 0.009]

Size Effect (Large – Small): **-0.001** [-0.008 0.006]

Interaction ($Size_S\,Color_S$ – $Size_S\,Color_L$) – ($Size_L\,Color_S$ – $Size_L\,Color_L$): **-0.002** [-0.024 0.023]

d. Slope of the Effect of Size Variation on the Target Influence

		Size Variation: Small	Size Variation: Large	Mean
Color Variation	Small	**0.056** [0.020 0.091]	**0.062** [0.017 0.107]	**0.059**
	Large	**0.049** [-0.013 0.110]	**0.058** [0.013 0.103]	**0.054**
	Mean	**0.053**	**0.060**	**0.057***

Color Effect (Large – Small): **-0.008** [-0.082 0.066]

Size Effect (Large – Small): **0.008** [-0.021 0.036]

Interaction ($Size_S\,Color_S$ – $Size_S\,Color_L$) – ($Size_L\,Color_S$ – $Size_L\,Color_L$): **-0.001** [-0.027 0.025]

e. Slope of the Effect of Size Variation on the Distractor Influence

		Size Variation: Small	Size Variation: Large	Mean	
Color Variation	Small	**-0.020** [-0.060 0.020]	**0.022** [-0.020 0.064]	**0.005**	Color Effect (Large – Small): **-0.007** [-0.048 0.035]
	Large	**-0.006** [-0.026 0.015]	**-0.006** [-0.030 0.019]	**-0.006**	
	Mean	**-0.013**	**0.008**	**-0.002**	
		Size Effect (Large – Small): **0.021** [-0.010 0.052]			Interaction ($Size_S\,Color_S$ – $Size_S\,Color_L$) – ($Size_L\,Color_S$ – $Size_L\,Color_L$): **-0.021*** [-0.039 -0.003]

Table 14: Coefficients for the five parameters of the model for four conditions, together with the confidence intervals of the values averaged across all four participants. Main effects of size and color and interactions are shown in the margins.

The slope of the effect of size variation on the target and distractor influence have been adjusted for the differences in size variation. Although the influence increases about the same amount for both variations in size, the effect seemed larger for the large variation in size because the larger differences in size units is not accounted for in the model. Although the units to measure the differences in size is arbitrary, it does increase on a log scale. Dividing the coefficient in half to take this larger difference in units into account shows no significant differences between the small and large variations in size.

Looking at efficiency across tasks, we found a significant main effect of color. However, there were no significant differences between the large variation in size and small variation in size within the two color variation tasks (Table 15).

		Size Variation			
		Small	Large	Mean	
Color Variation	Small	**0.581** [0.357 0.706]	**0.661** [0.493 0.829]	**0.621**	Size in Small Color Variation (Large – Small): **0.065** [-0.031 0.161]
	Large	**0.843** [0.723 0.918]	**0.850** [0.809 0.891]	**0.847**	Size in Large Color Variation (Large – Small): **0.015** [-0.030 0.060]
	Mean	**0.712**	**0.756**	**0.734**	
					Color Effect (Large – Small): **0.239*** [0.137 0.341]

Table 15: Efficiency averaged across four subjects for each condition. Values in brackets are confidence intervals. Values in the margins show the column means, the row means, and the grand mean. While there was a significant effect of color variation on efficiency, there were no significant effects of size variation within each color variation condition.

It is surprising to find that there were no significant differences in efficiency across conditions, as it suggests that grouping did not occur on the size feature. If grouping had occurred, we would have seen efficiency decrease as size variation increased, since efficiency is thought of as a measure of error, and the error would have increased as the larger groups drew attention away from the target stimuli.

Study 3: Discussion

Results from study 2 revealed that when calculating a centroid, participants seemed to automatically group stimuli together when stimuli were similar and salient (in study 2, this salient feature was numerosity). The purpose of this current study was to see if grouping occurred even when the feature that the stimuli may be grouped on was not the feature of interest for the participant. In other words, when participants were attending to the color of the stimuli, would they still group stimuli together based on the stimulus' size? Although we were originally looking for effects of the magnitude of the size variation on the mean influence of the targets and the distractors, we found none. This would seem to argue against there being any (at least any effect of) grouping based on the irrelevant variation. This in turn, argues against the kind of bottom-up, automatic grouping mechanism that we were expecting to observe. Instead, we found a large influence of size on the influence of the targets (but not the distractors).

From these results, we concluded that participants could not ignore variation in size, even though it was meant to be the irrelevant feature. In Study 2, the distracting bottom-up cue could be ignored, perhaps because it involved variation in the one feature that participants were mean to attend to (color). It seems that in this current study, because an irrelevant feature that does not conflict with the relevant feature is providing the bottom-up cue, participants were unable to ignore it. However, it seemed that they were only unable to ignore this irrelevant feature in the target group. Like in Study 2, even if there is a salient type of stimulus that can capture attention on its own, it will not do so when the instructions are to attend to a different type of stimuli. In this study, it is clear that larger dots are given more influence in both the target and distractor groups, but the effect is only significant for the targets, since participants were instructed to ignore distractors.

CHAPTER 6

STUDY 4 – RETINAL GROUPING VERSUS REFLECTANCE GROUPING

Study 4: Introduction

Based on the results from studies 1 and 2 discussed above, grouping may be an automatic process that occurs when people are presented with displays in which a subset of the stimuli are homogenous in a relevant feature. However, although the bottom-up formation of groups seems to be an automatic process, salience associated with large groups is not obligatory in capturing attention. To further characterize the grouping mechanisms that appears to play a large in the perception of spatially disorganized heterogenous stimuli, we explore the following question: at which stage does grouping occur during visual processing? Participants are able to achieve higher efficiency and selectivity values in conditions in which targets can be grouped together by similarity, despite quick display durations, which is evidence that grouping occurs early in visual processing. We wish to further determine when grouping occurs in relation to another early stage of visual processing: perceptual constancy.

Although it is generally agreed among theorists that grouping occurs early in visual processing, whether it occurs before or after constancy has been achieved is still debated (Palmer, Neff, & Beck, 1996). One example of constancy that we can easily observe is color constancy in objects covered by shadows. If grouping is a process that occurs before the perceptual computation underlying color constancy, variations in lighting would be expected to disrupt feature-based selective attention. Consider three lights: L1 and L2 are two different stimulus types viewed in an illuminated area, and L3 is a stimulus type viewed within a shadow. A critical case is the one in which L3 matches L1 in physical characteristics (i.e., before perceptual constancy) but matches L2 in terms of reflectance. What happens when, in separate conditions we ask the participant to form the centroid of all stimuli like L1 versus all stimuli like L2?

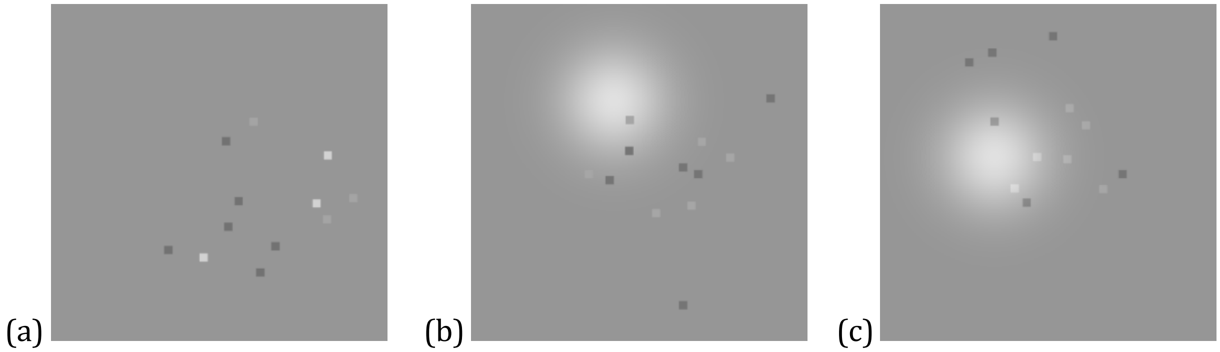
If grouping occurs after the computation producing perceptual constancy, the presence of objects in shadows would not be expected to affect performance. Schultz & Sanocki (2003) conducted an experiment in which participants were asked to group dots together based on color; the dots could either be grouped with or without taking constancy into account. In their study, they defined grouping without considering constancy as "retinal," and grouping with considering constancy as "reflectance." They found that 11.7% of participants were able to group stimuli together based on reflectance at short display times (200 ms), but the percentage increased to 88.3% at longer display times (2000ms). They tested participants with display times of 200 ms, 500 ms, 1100 ms, and 2000 ms, and found that as display duration increased, so did the proportion of participants who grouped stimuli based on reflectance spectrum. These proportions were achieved without giving participants feedback after each trial.

Because we want each trial to have short display times (on the order of hundreds of milliseconds), it would not be feasible to test participants with duration times of 1100ms or 2000ms, like Schultz and Sanocki (2003) did in their study. Instead of using the full range of display durations to increase the chances of participants perceiving the dots by reflectance, we only used one short duration (400ms; we chose this duration instead of 200ms in order to avoid floor effects) with an SOA of 50 ms and presented feedback after every trial to encourage them to rely on the reflectance spectrum when computing the centroid.

Introducing spotlights into the background for the centroid task will allow us to explore the question of whether the grouping processes we have hypothesized to underlie centroid judgements occurs before or after the computation of perceptual constancy. Participants in this study may be able to estimate centroids only based on luminance or estimate centroids only based on reflectance.

Study 4: Design

The experiment followed the same procedure described in chapter 2. The background had a luminance of 49.99 cd/m^2 in the "unilluminated" areas and a luminance of 81.83 cd/m^2 in "spotlight" areas. The spotlight was a bivariate gaussian centered around a random position within the display area with a sigma of 50 pixels. The Gaussian was cut off at 500 pixels. The center of this spotlight was constrained to only appear within 300 pixels of the center of the display to ensure that the entire spotlight would be present within the display area. The luminance of the spotlight was chosen so that the luminance of the darker dot in the spotlight was equal to the luminance of the brighter dot in the shadow. Stimuli consisted of 12 achromatic dots. Darker dots had a reflectance of 0.50 whereas brighter dots had a reflectance of 0.80. When the dots were presented on the background outside of the spotlight, they had luminances of 24.86 cd/m^2 and 40.00 cd/m^2 respectively. When the two dot types were presented in the spotlight, they had luminances of 40.00 cd/m^2 and 63.64 cd/m^2 respectively. The luminance of each dot is equal to the dot's reflectance multiplied by the illuminance of the light and shadow areas. Sample stimulus displays are shown in Figure 11.

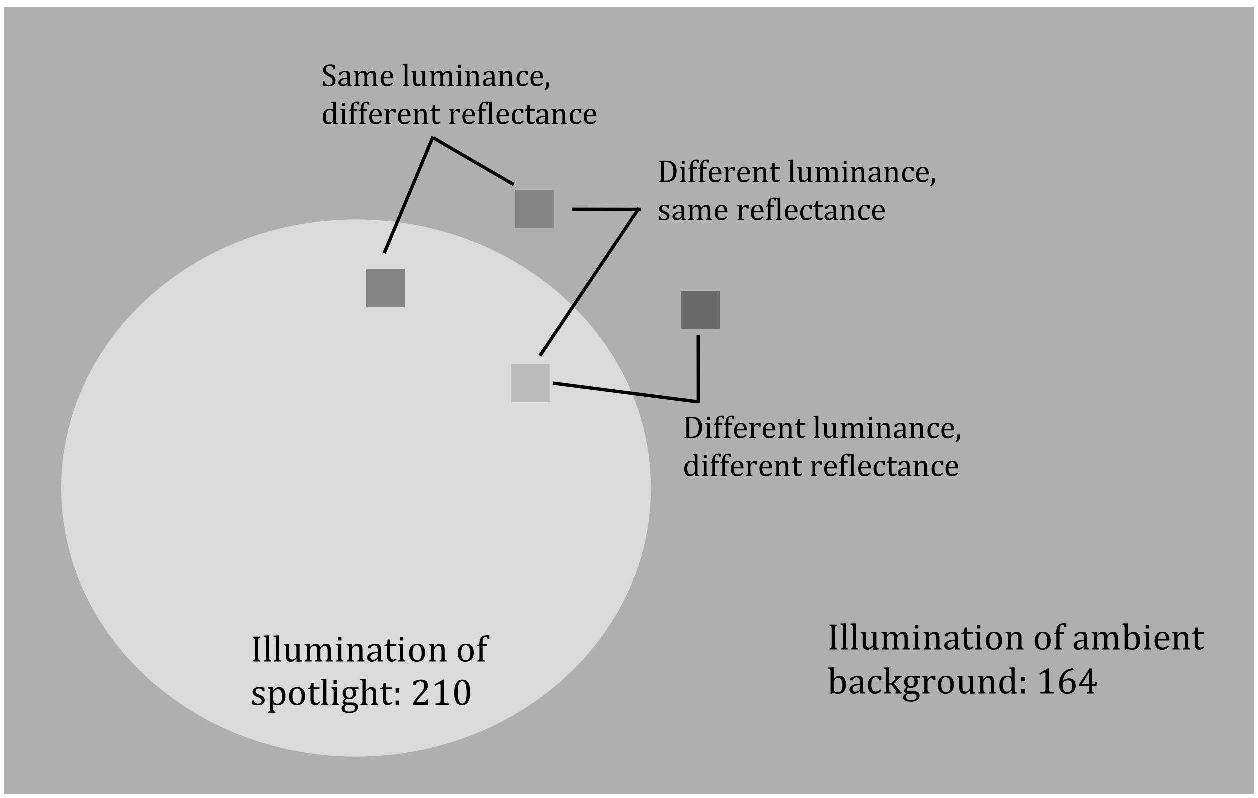


	Dark dot, outside of spotlight	Dark dot, in spotlight	Light dot, outside of spotlight	Light dot, in spotlight
Task 1: No spotlight	Yes	No	Yes	Yes
Task 2: Spotlight, no luminance constancy	Yes	No	Yes	No
Task 3: Spotlight, with luminance constancy	Yes	Yes	Yes	Yes

Figure 11: Sample stimuli for the experiment. "Spotlights" were randomly generated for each trial. (a) Task 1 has no spotlight. Luminances present are the darker dot with its luminance in the shadow area, and the brighter dot with its lumiances in both the shadow and spotlight areas. Because there is no spotlight, there is no luminance constancy to help the participants group the two target luminance levels together. (b) Task 2 has a spotlight and both reflectances. There is no luminance constancy, however, since the dots have the same luminance regardless of whether or not they are illuminated by the spotlight. (c) Task 3 has a spotlight and both reflectances. Luminance constancy is present, since the dots illuminated by the spotlight appear brighter than dots illuminated by the ambient light. The type of stimuli present in each task is shown in the table. The top row describes the type of stimulus that could be present, with the background showing the luminance of said stimulus.

The three tasks included in this experiment were 1) no spotlight, 2) spotlight, with no luminance constancy, and 3) spotlight with luminance constancy (Figure 11). Task 1 was similar to the T3D1 condition in study 1, in which targets were heterogeneous and distractors were homogenous. Six darker, distractor dots and six brighter, target dots were displayed. These are the luminance levels that will also be present in Condition 3, except without the presence of the spotlight. Therefore, luminance constancy was not present to help the participants group all of the target dots together. Participants were asked to give equal weight to both target luminance levels. In Condition 2, both reflectances were present, as well as a spotlight. The spotlight did not affect the dots' luminaces, however, so darker dots had a luminance of 24.86 cd/m^2 and brighter dots had a luminance of 40.00 cd/m^2, regardless of whether or not they were in an area illuminated with the spotlight. In Condition 3, both reflectance levels were present, the spotlight was present, and the dots that were illuminated by the spotlight did have a brighter luminance than dots that were not illuminated by the spotlight. In this condition, luminance constancy was expected to help the participants group the target dots together, despite their different luminance. However, because the luminance of the darker dot in the spotlight area were the same as that of the brighter dots in the shadow (Figure 12), the low-level grouping mechanism may have confused which dots they should group together. If participants group the dots based on physical characteristics then they are expected to group the brighter dot in the shadow area with the dark dots in the illuminated area. If they group the dots based on reflectance, then they are expected to group the brighter dots in the shadow area with the brighter dots in the illuminated area.

Six participants were recruited for this experiment and completed the tasks in ABCCB order. All participants completed three sessions of Task 1 (60 trials per session) first. Three participants then completed three sessions of Task 2 (105 trials per sessions) and then three sessions of Task 3 (105 trials per session). The other three participants completed three sessions of Task 3, then three sessions of Task 2. All subjects then completed another three sessions each of Tasks 2 and 3, but in reverse order.



	Reflectance	Ambient	Spotlight
Dark Dots	0.50	24.86 cd/m^2	40.00 cd/m^2
Bright Dots	0.80	40.00 cd/m^2	63.64 cd/m^2

Figure 12: Example that illustrates how luminance and reflectance relate to one another. Luminance values for dots and the background that were used in the experiment are included in the table. Each luminance value is calculated by multiplying the reflectance of the object with the illumination of the region area in which it lies.

Study 4: Results

Performance in the three tasks were measured by mean error, log selectivity, and efficiency, displayed in Table 16. Contrasts between the three tasks were conducted to look for main effects of spotlight and luminance constancy on performance (Table 17). P-values indicated that the only significant difference was found between the log selectivity from task 1 and task 2 (mean = 0.3470 [0.1583 0.5376], t(5) = 4.717, p = 0.005, BF = 11.22) with log selectivity being greater in task 1. There were no other significant differences between tasks, indicating that participants can perform just as well in this experiment regardless of the presence or absence of luminance constancy.

	Task 1 No spotlight present	Task 2 Spotlight present, No luminance constancy present	Task 3 Spotlight present, Luminance constancy present
Mean Error	16.7 [15.4 18.1]	17.5 [16.4 18.6]	17.0 [15.8 18.1]
Efficiency	0.735 [0.695 0.776]	0.723 [0.696 0.749]	0.737 [0.700 0.774]
Log Selectivity	1.273 [1.053 1.494]	0.926 [0.646 1.205]	1.297 [0.850 1.743]

Table 16: Means for each of the three measures of interest in the experiment taken across participants. Values in brackets are confidence intervals for each mean.

	Efficiency			Log Selectivity		
Participant	T1-T2	T2-T3	T1-T3	T1-T2	T2-T3	T1-T3
1	0.0704	-0.0665	0.0039	0.4163	-0.2977	0.1187
2	-0.0448	-0.0404	-0.0852	0.1823	-0.4785	-0.2961
3	0.0541	0.0269	0.0809	0.4949	0.0393	0.5342
4	-0.0255	-0.0120	-0.0374	0.5942	-1.3533	-0.7592
5	0.0236	-0.0286	-0.0050	0.1495	-0.3853	-0.2359
6	-0.0026	0.0324	0.0298	0.2505	0.2487	0.4992
Mean	0.0125	-0.0147	-0.0022	0.3479	-0.3711	-0.0232
SD	0.0451	0.0387	0.0568	0.1807	0.5540	0.5030
UpperB	0.0598	0.0259	0.0574	0.5376	0.2102	0.5047
LowerB	-0.0348	-0.0553	-0.0618	0.1583	-0.9525	-0.5511
T	0.681	-0.930	-0.094	4.717	-1.641	-0.113
p	0.526	0.395	0.929	**0.005**	0.162	0.915

	Mean Error		
Participant	T1-T2	T2-T3	T1-T3
1	-2.6249	-2.8194	0.1945
2	1.0216	-0.9690	1.9906
3	-1.9373	0.8999	-2.8372
4	0.9232	-0.2094	1.1326
5	-2.2640	- 0.9875	-1.2765
6	0.3388	1.0389	-0.7001
Mean	-0.7571	-0.5077	-0.2494
SD	1.6935	1.4317	1.7366
UpperB	1.0202	0.9948	1.5731
LowerB	-2.5344	-2.0102	-2.0718
T	-1.095	-0.869	-0.352
p	0.323	0.425	0.739

Table 17: Contrasts between the three tasks in the experiment for each of the measures of interest. The only significant difference is found between task 1 and task 2 in log selectivity.

Study 4: Discussion

We were unable to find systematic differences between the three tasks observed in this study. Adding a spotlight to the display seemed to only have a significant effect on log selectivity when there was no luminance constancy present. From looking at previous results, we would've expected to see a difference in performance between tasks 2 and 3, with task 2 producing higher efficiency and selectivity, since luminance constancy seems to occur later in visual processing. It appeared that in this experiment, participants were able to attend to target stimuli well, regardless of whether or not there was a spotlight present, and whether or not luminance constancy was present.

This experiment differed from other studies of luminance constancy because the displays were only presented for a short duration, and participants were forced to make luminance judgements within that time window. Other studies had not truncated the display time, allowing participants to view the stimuli for as long as they needed, which left ample time for luminance constancy to come into effect. Schultz and Sanoki (2003) did measure reaction time, and they found that shorter reaction times resulted in participants grouping based on retinal luminance while longer reaction times resulted in grouping based on perceptual luminance. The shortest reaction time that allowed most participants to group based on perceptual luminance was longer than the display time used in our experiment, yet our participants were able to give as much weight to targets in both tasks that use spotlights. We piloted a version of the experiment that used an even shorter display time (200ms) and found no difference in performance. Although this version has not been run by many participants and the results formally compared, the data collected so far indicates that performance should be similar to this current experiment's.

At these short display times, participants were expected to make quick luminance judgements to determine if the dot was a target or distractor. Our current design does not seem to challenge this ability enough. One experiment to try in the future would be to vary the SOA to try to find the threshold at which the participant can't account for luminance constancy and groups the dots based on retinal luminance. Another would be to introduce a third reflectance level to the display so that participants need to be more selective for the target dots. With this current design, it may be the case that because there are only two reflectance levels, participants are able to distinguish between targets and distractors easily because they have to only draw one boundary in feature space.

CHAPTER 7
SUMMARY AND CONCLUSIONS

In a selective centroid task, participants are instructed to attend to items with a specific feature. It is assumed that in order to find the target items, they must deploy top-down attention, but the experiments in this thesis suggest that this top-down attention is not the only mechanism that seems to be driving the centroid judgement, as bottom-up attention also seems to be influencing the judgment. While bottom-up attention is able to drive the centroid judgment on its own, it also interacts with top-down attention when participants are given multiple cues within a display. We have identified a few characteristics of bottom-up attention and suggest that these may be evidence for a grouping mechanism that facilitates search for a centroid.

Results from Study 1a revealed that one factor that may be driving bottom-up attention is the homogeneity of the target group. Target and distractor heterogeneity both degrade performance in the centroid task, but target heterogeneity had a larger effect than distractor heterogeneity. One suggestion for why participants improved their performance in the homogeneous conditions was that they were able to group all of the target items together instead of in three separate groups. With homogeneous targets, the stimuli would all be clustered into a single group because of their reddish hue, and the resulting group could pop out and be easily selected. Unlike the top-down effects of selective attention that we have been discussing, these results appeared to reflect a stimulus-driven (bottom-up) form of selective attention.

Study 1b aimed to find evidence for the categorical grouping that may have occurred in Study 1a. Using closely-spaced colors might have allowed us to find the exact point at which the colors were grouped together. Unfortunately, there was no consistent pattern in the observed influence functions from the study across the three participants as step sizes in color space increased. This indicated that each participant employed a different strategy for the task. We also found no clear step function that could describe the decline in efficiency as spacing increased, which is what we would have expected to find if the participants had deployed categorical grouping. While there does seem to be a clear boundary between large color differences that force separate groups, there does not seem to be one for smaller color differences. This could be a result of the limitations of linguistic categories. Previous studies have proposed that linguistic color categories modulate color perception, and that observers who are familiar with these categories tend to distinguish colors more reliably if the color falls in the middle of the category, rather than at the boundaries. Because the colors used in the study approached and blurred the boundaries between the three target hues from our color wheel, participants were able to treat the colors as belonging to one category instead of three.

It could have been the case that numerosity also affected performance in Study 1. With our design, when a group of dots is homogeneous, it also became the group with the most numerous color. Study 1showed that this numerosity made the stimuli pop-out, which captured bottom-up attention. The goal in study 2 was to compare performance in the centroid task when the participant is driven by instructions versus when they are driven by the stimulus itself. In other words, if there was a grouping mechanism at work, how strong was it compared to the top-down attention mechanism? One may have expected these mechanisms to be at odds with one another, especially if the instructions of the task were to attend to a stimulus type that does not immediately grab our attention as well as another stimulus type. One may also have expected the effects of both mechanisms to be additive if cues from the display drove both mechanisms to point attention towards one specific stimulus type. Our experiment demonstrates, however, that these expectations are not true in all cases.

When guidance from top-down intentions and bottom-up stimulus properties conflicted, the intention overrode the salient stimulus properties. This occurred without a large reduction in either selectivity or efficiency. We did find, however, that the influence functions of the targets in the top-down attention tasks were significantly higher than the influence functions of the targets in the bottom-up task. Although the bottom-up mechanism could be used to drive attention, it did not do so as efficiently as the top-down mechanism.

As we predicted, participants had highest selectivity in conditions that had them search for a target color that also happened to be the most numerous stimulus type. We found that top-down attention alone showed selectivity and efficiency as high as when the both top-down and bottom-up attention pointed to the same thing. This effect was asymmetrical, however. There was no significant evidence that adding bottom-up attention to top-down attention improved selectivity. These comparisons demonstrate the relative effectiveness of selective attention in these two “pure” cases. We also noted that numerosity seemed to not only affect the targets, but also the distractors. The influence function of the distractor color that was also the most numerous had significantly lower influence functions than the mean of the other distractors. This is consistent with findings from study 1, in which participants achieved the highest selectivity when not only the targets were homogeneous, but also when the distractors were homogeneous.

These results from Study 2 indicate that attention may not consist of two separate mechanisms that are competing with one another to guide attention. A model that would be consistent with these findings posits that attention is a single mechanism that can be guided either by bottom-up control or by top-down command. The bottom-up mechanism may be obligatory if it is the only cue that helps a participant search for a target, but it is not as strong a cue as top-down instructions. The top-down command is stronger, which was evident when participants were more selective for a color they were instructed to attend to, regardless of whether or not it was the most numerous.

Results from Study 2 suggested that the bottom-up mechanism that can be used to drive the centroid judgement can be ignored if it conflicts with instructions. It should be noted that the feature that was used to group the stimuli together was the feature that participants wanted to attend to (i.e. hue) to differentiate targets from distractors. The purpose of Study 3 was to determine if grouping drew participants' attention not only to stimuli with relevant features, but also to stimuli with irrelevant features. For this study, participants attended to the color of the stimuli, and tried to ignore the size. While participants were able to focus mostly on color, they were unable to ignore size completely. Participants were able to give more weight to the target color, which demonstrated their ability to focus on the relevant feature, but they also gave more weight to larger target dots than they did to smaller target dots. From these results, we concluded that participants could not ignore variation in size, even though it was meant to be the irrelevant feature. However, like in Study 2, even if there is a salient type of stimulus that can capture attention on its own, it will not do so when the instructions are to attend to a different type of stimuli. In both target and distractor groups, larger dots are given more influence, but the effect was only significant for the targets, since participants were able to ignore the group of distractor colors. These results also seem to be inconsistent with grouping as a mechanism, since attention seemed to not have been drawn to the cluster of similar-looking stimuli.

Study 4 aimed to determine how quickly grouping occurred compared to another relatively quick visual processing mechanism: luminance constancy. If grouping is a relatively early process, participants were expected to group the stimuli together based on retinal luminance, or the physical luminance of the dot. If grouping is a relatively slower process, then participants were expected to group the stimuli together based on perceptual luminance, or the perceived luminance of the dot when the light sources are taken into account. We found that adding a spotlight to the display seemed to only have an effect on selectivity of the target when there was no luminance constancy present. From looking at previous results, we would've expected to see a difference in performance between the two tasks with spotlight present. It was reasonable to expect the with no luminance constancy to produce greater selectivity and efficiency since luminance constancy seems to occur later in visual processing, according to previous research. Our experiment was unable to find evidence for whether luminance constancy or grouping occurred earlier in visual processing, since participants were able to attend to target stimuli regardless of whether or not there was a spotlight present, and whether or not luminance constancy was present in the display.

The results from these four experiments gave inconclusive evidence that a grouping mechanism may be influencing our centroid judgement. When there were multiple instances of an item type in a display, participants' responses seemed to be drawn to the items that could be grouped together, despite being in disparate locations. The proposed mechanism seemed to draw attention towards the group, although it may be ignored if instructions indicate that the group should be ignored. Results from the third study contradicted evidence of grouping, however. If such a mechanism is presumed to automatically form clusters of similar items, then the stimuli that are similar would have been expected to pop out in the display. In the study, however, participants were able to ignore the groups that would have been created based on size. Although the results from these studies do not lead to definitive conclusions about a grouping mechanism, they do reveal a few characteristics of a more general bottom-up mechanism that influences centroid judgements. We did find that if a stimulus varied on a feature that was irrelevant to search, it would still drive our centroid judgements if it was salient, although not as strongly as the instructions that tell us to attend to a specific feature. This salient irrelevant feature only influenced performance in the task if it made some target stimuli more salient than others. Top-down attention is the main driving force behind the centroid judgement, but that judgement is indeed influenced by bottom-up attention.

REFERENCES

Anstis, S. M., & Cavanagh, P. (1983). A minimum motion technique for judging equiluminance. Colour vision: Physiology and psychophysics, 155-166.

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.

Bornstein MH, Korda NO. 1984. Discrimination and matching within and between hues measured by reaction times: some implications for categorical perception and levels of information processing. *Psychol. Res.* 46: 207-22

Brainard, D. H. (1997). The Psychophysics Toolbox. Spatial Vision, 10(4), 433-436. doi: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.

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 Proceedings of the IEEE International Conference on Computer Vision Workshops (pp. 621-629).

Derrington, A., Lennie, P., & Krauskopf, J. (1983). Chromatic response properties of parvocellular neurons in the macaque LGN. Colour vision, 245-251.

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), 242-242.

Drew, S. A., Chubb, C. F., & Sperling, G. (2010). Precise attention filters for Weber contrast derived from centroid estimations. Journal of vision, 10(10), 20-20.

Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. Psychological Review, 96(3), 433-458. doi: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. JOSA 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.

Gaspelin, N., Leonard, C. J., & Luck, S. J. (2015). Direct evidence for active suppression of salient-but-irrelevant sensory inputs. Psychological Science, 26(11), 1740-1750.

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, 9, 1021.

Jonides, J. (1981). Voluntary versus automatic control over the mind's eye's movement. Attention and performance, 187-203.

Kahneman, D., & Henik, A. (1977). Effects of visual grouping on immediate recall and selective attention. Attention and performance VI, 307-332.

Kurylo, D. D., Waxman, R., Kidron, R., & Silverstein, S. M. (2017). Visual training improves perceptual grouping based on basic stimulus features. Attention, Perception, & Psychophysics, 79(7), 2098-2107.

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

Lunau, R., & Habekost, T. (2017). Effects of task-irrelevant grouping on visual selection in partial report. Attention, Perception, & Psychophysics, 79(5), 1323-1335.

Lu, Z. L., & Sperling, G. (2001). Three-systems theory of human visual motion perception: review and update. JOSA 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.

Posner, M. I. (1980). Orienting of attention. Quarterly journal of experimental psychology, 32(1), 3-25.

Ren, X., & Malik, J. (2003, October). Learning a classification model for segmentation. In null (p. 10). IEEE.

Reynolds, J. H., Alborzian, S., & Stoner, G. R. (2003). Exogenously cued attention triggers competitive selection of surfaces. Vision research, 43(1), 59-66.

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.

Schulz, M. F., & Sanocki, T. (2003). Time course of perceptual grouping by color. Psychological Science, 14(1), 26-30.

Sun, P., 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. doi:10.3758/s13414-015-0978-2

Sun, P., Chubb, C., Wright, C. E., & Sperling, G. (2018, May) 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, Florida

Sun, P., Chubb, C., Wright, C. E., & Sperling, G. (2018). High-capacity preconscious processing in concurrent groupings of colored dots. *Proceedings of the National Academy of Sciences*, *115*(52), E12153-E12162.

Theeuwes, J. (1991). Exogenous and endogenous control of attention: The effect of visual onsets and offsets. Perception & psychophysics, 49(1), 83-90.

Wagemans, J., Elder, J. H., Kubovy, M., Palmer, S. E., Peterson, M. A., Singh, M., & von der Heydt, R. (2012). A century of Gestalt psychology in visual perception: I. Perceptual grouping and figure–ground organization. Psychological bulletin, 138(6), 1172.

Witzel, C., & Gegenfurtner, K. R. (2018). Color perception: objects, constancy, and categories. *Annual Review of Vision Science*, *4*, 475-499.

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