Title
Using the Centroid Method to Study Feature Based Selective Attention

Author
Inverso, Matthew

Publication Date
2017

UNIVERSITY OF CALIFORNIA,
IRVINE

Using the Centroid Method to Study Feature Based Selective Attention

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Cognitive Science

by

Matthew Inverso

Dissertation Committee:
Professor Charles Chubb, Chair
Associate Professor Charles Wright
UCI Distinguished Professor George Sperling

2017

DEDICATION

This thesis is dedicated to my family—without the support of my mother, brother, and sister, I would not have been able to get this degree. Thank you so much!

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

This research was made possible through the generous support, help, and materials given by my advisors, colleagues, and friends at the University of California, Irvine.

I especially need to thank the senior member of the Chubb-Wright Lab, Charlie Chubb, Ted Wright, George Sperling. The three of them have provided so much guidance, feedback, and resources—I truly could not have done this without them.

Just as important, I need to thank all of the other members of the lab, who have given up their time and energy to make sure my work is as good as it can be! Their advice, support, and data have made this dissertation fantastic.

My friends in the department have also been invaluable to me, for helping keep me sane. I love you all.

A portion of the text of this dissertation is a reprint of the material as it appears in Attention, Perception, and Psychophysics. The co-authors listed in this publication (Peng Sun, Charlie Chubb, Ted Wright, and George Sperling) directed, supervised, and aided with research which forms the basis of the first chapter of this dissertation.

CURRICULUM VITAE

Education:
University of California, Irvine, CA Graduation Date: July 2017
Ph.D program in Cognitive Science.
Master conferred March 2016

Rutgers University, New Brunswick, NJ Graduation Date: May 2012
Double Major: Psychology (GPA: 3.96) and Statistics/Mathematics (GPA: 3.77)
Cumulative GPA: 3.84

Work Experience:
Teaching Assistant, University of California, Irvine September 2012-September 2017
Primary Responsibilities: Leading discussion sections and labs, proctoring exams, and grading papers and exams in classes ranging from Introduction to Psychology to Elementary Statistics.

Teaching Assistant, Rutgers University September 2011-May 2012
Primary Responsibilities: Grading papers and exams for Dr. Jack Aeillo's class on Organizational and Personnel Psychology.

Research Assistant, Rutgers University September 2010-May 2012
Primary Responsibilities: Prepare research subjects and administer study protocol for the various projects. Also research previous work related to this field of study.

Research Experience:
Independent Research, University of California, Irvine September 2012-July 2017
Advisor: Dr. Charlie Chubb
Primary Responsibilities: Design, code, run, analyze, and publish experiments, while working on a dissertation.

Psychology Honors Program, Rutgers University September 2011-May 2012
Advisor: Dr. Jacob Feldman
Primary Responsibilities: Continue with previous work from the summer, completing a thesis and discovering new knowledge.

REU Summer Researcher, Rutgers University, PSG June 2011-August 2011
Advisor: Dr. Jacob Feldman and Dr. Manish Singh
Primary Responsibilities: Design and run a study on perception and present findings to various people involved in the program.

Volunteer Internship, Rutgers University Psychology Department June 2009-September 2010
Supervisor: Dr. Richard Contrada

Primary Responsibilities: Helped design and run the Emotions 2 study, an investigation of the relationship between stressful life events and physical health.

Honors and Awards:

Received John I. Yellott Scholar Award	June 2014
Received Bachelor's Degree with High Honors	May 2012
Awarded Class of 1969 Honors Scholarship	February 2012
Accepted into Psi Chi, Psychology Honors Society	November 2010
Awarded School of Arts and Sciences Excellence Award	June 2010
Awarded Rutgers 2010 Academic Excellence Award*	May 2010
*Awarded to top 10% of sophomore class	
Accepted into Honors Society, Delta Epsilon Iota	September 2010
Acceptance into Honors Program of Rutgers University	January 2008

Service to Community:

May, 2009-May 2010: Peer Mentor Rutgers University

Primary Responsibilities: Helped incoming freshman in the Rutgers University Honors program to get acquainted with the school and provide them peer support.

April 2017: Reviewer for Attention, Perception, and Psychophysics

Publications:

Inverso, M., Sun, P., Chubb, C., Wright, C. E., & Sperling, G. (2016). Evidence against global attention filters selective for absolute bar-orientation in human vision. *Attention, Perception, & Psychophysics*, *78*(1), 293-308.

Poster Presentations:

First Author at Vision Science Society's Annual Conference	May 2017
First Author at Psychonomics Society's Annual Meeting	November 2014
First Author at Vision Science Society's Annual Conference	May 2012
Honors Thesis Presentation, Rutgers University	May 2012
First Author at Perceptual Science Seminar, Rutgers University	August 2011

Conference Talks:

First Author at Vision Science Society's Annual Conference	May 2016
First Author at Vision Science Society's Annual Conference	May 2015

ABSTRACT OF THE DISSERTATION

Using the Centroid Method to Study Feature Based Selective Attention

By

Matthew Inverso

Doctor of Philosophy in Cognitive Science

University of California, Irvine, 2017

Professor Charles Chubb, Chair

Feature-based selective attention is a very important ability people use to explore the world around them. Being able to suppress distractions and focus on one feature of items is instrumental in a person being able to find what they are looking for. Historically, this has been studied with a search task—one that asks participants to find a target item amongst distracters. This method, while useful, has some flaws that leave open gaps in our knowledge of how feature-based selective attention works and what features people can use it on. This dissertation explores a relatively new method of exploring this ability—the centroid method—that is capable of covering these gaps. Using this paradigm, Chapter 1 shows that a feature-dimension the search literature implied was useful—orientation—is not. Subjects cannot ignore one orientation and attend only to another. The search literature conflated orientation with local orientation contrast, and the centroid task was able to separate these features. Chapter 2 compares the centroid task to a numerosity estimation task in order to show that computing centroids does not rely on computing numerosity, eliminating some otherwise reasonable theories on how these computations are done. Chapter 3 uses the centroid method to explore a feature dimension that heretofore has not been extensively studied—internal angles. Subjects were able to use the internal angle of items as a useful feature in the centroid task, to a limited degree. As long as the angles are discriminable enough, which seems to be determined by the ratio between them, subjects could selectively attend to just a target angle (or range of angles) while ignoring distracters. Overall, the centroid method proved to be a useful tool for studying how feature-based selective attention worked in three different experiments, and when subjects could and could not use it.

General Introduction

Imagine you are in a supermarket. It is a brand new market, so you are not quite sure where anything is, but you know you need tomatoes. So, heading to the produce isle, you start looking for red, rounded things. Before finding the tomatoes, you notice the red bell peppers, but you did not notice the broccoli, though you looked right at it. This is feature-based selective attention—attending to only specific features (in this case the color "red" and the shape "rounded") based on a goal. This is a very important ability for people to have! Without being able to attend to features selectively, only things that are intrinsically salient will be noticed, and many useful items in the world will be missed. As such, the ability for people to do this has been studied and discussed heavily for decades (Bergan & Julesz, 1983; Triesman & Gelade, 1980; Wolfe, 1994).

The earliest work on feature based selective attention was focused on finding what feature dimensions people were able to attend to, and the limits of each of those. A major task often employed for this purpose became known as "visual search" (figure 1). In a search task, participants see a group of items (for example, dots), and have to say if an item with a particular target feature (for example, red) is included with the group. Sometimes, this is a really easy task, and people respond very quickly and accurately. The target, if present, pops out in an obvious way, regardless of how many distracters there are. This was taken as a sign that people could attend to the feature of interest, and let everything out fade into the background. In other cases, the task is very difficult, requiring the participant to scan each item one by one until they either find the target or decide it is not present, so responses take longer. This is taken as a sign that the feature dimension being studied is not something people can attend to.

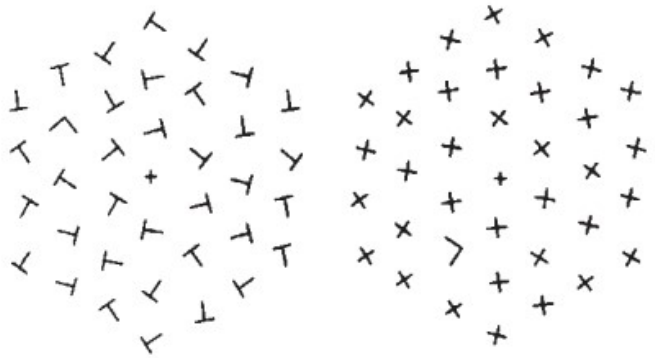

Figure 1: Two examples of visual search displays, adapted from Bergan & Julesz, 1983. The left figure shows the difficult task of finding a target L among distracting Ts. The right figure shows the easy task of finding a target L among pluses.

Still other cases had more ambiguous results. Some features tested lead to asymmetries. For example, it is very easy to find a Q shape among O shapes, but very difficult to find an O shape among Q shapes. Anne Triesman argued (Triesman & Gromican, 1988) that these asymmetries emerge from the fact that it is easier to search for something that HAS a basic feature than something that does not. In other words, in the O and Q task, the bar at the end of the Q is the feature people are attending to. This is a contentious claim (Rosenholtz, 2001), but many experimenters have used these search asymmetries to argue for or against certain features being selectable.

The search task has been very useful in determining many properties of feature based attention, but it does have flaws. First of all, there being only one target among many distracters means there is a high amount of contrast around the target. It is the thing that is different from everything else. This can be mitigated by having heterogeneous distracters, but that then adds discriminability as a confound. If we are finding a vertical bar amongst homogeneous distracters is easy but among heterogeneous distracters is hard, is that a sign that we are using orientation contrast rather than absolute orientation, or that some of the variety of oriented bars are too close to vertical, making discriminating the target very hard? The search task is not well equipped to handle questions of this sort.

Another flaw is that it is not efficient. Since the measure of interest is reaction time, and differences in mean reaction time tend to be fairly small, participants must do a lot of trials for conclusions to be reached. This can lead to search tasks being costly in time investment, which can lead to participants getting fatigued, which can add unneeded noise to the data.

Finally, and perhaps most critically, the search task is not a true attention task. To truly test attention, the only thing that should change between conditions of an experiment should be the attention instructions. Otherwise, it is hard to determine if any change in results between conditions is due to the instructions or the change in stimuli. In a search task, when the instructions are changed, the stimuli are as well. The target item only appears once, whereas distracters appear multiple times, so when the target identity changes, so does the content of the display.

One solution to these challenges is the centroid paradigm (Sun et. al., 2015). In this task, participants see a very briefly displayed cloud of items, and must indicate the center of mass of either all items, or some subset defined by a feature. For example, the cloud might consist of dots of various luminances, and participants could be instructed to attend only to those that are lighter than the background (Drew, Chubb, & Sperling, 2010). Experimenters can then use the participants response to determine how much weight they gave to each type of item. If the feature being studied is an easy one, participants will give nearly 0 weight to distracters. If, on the other hand, the feature is a hard one, participants will not be able to attend exclusively to it, and will give much more weight to the distracters.

This paradigm neatly solves all of the problems with the search task. Because there are (generally) equal numbers of targets and distracters, local contrast is, on average, the same everywhere. The centroid method eliminates the possibility of that being a useful cue. Further, it is a very efficient method—filter weights can be estimated pretty well with only a couple hundred trials, each of which is only about 3 seconds long.

The centroid task is also easily made into a true attention task. In a display with 4 bright dots and 4 dark dots, it is trivial to change the task by changing ONLY the instructions, leaving the display exactly the same between conditions. Asking the participants to attend to the bright dots should give a different result than asking them to attend to the dark dots, and that cannot be a relic of the changing display, which eliminates a potential confound.

In this document, I will explore several applications of the centroid method in the pursuit of knowledge on feature-based selective attention. Chapter 1 explores a case where we suspect the search task gave misleading results due to the local contrast issue. The centroid method is able to show that orientation is not the guiding feature people once thought. Chapter 2 compares the centroid method to a numerosity estimation task in order to eliminate one theory on how the centroid of a dot cloud is computed in the brain. Chapter 3 uses the centroid method to examine if angles are a guiding feature that can be used in feature based selective attention.

CHAPTER 1:

Evidence against global attention filters selective for absolute bar-orientation in human vision

Abstract

The finding that an item of type *A* pops out from an array of distractors of type *B* typically is taken to support the inference that human vision contains a neural mechanism that is activated by items of type *A* but not by items of type *B*. Such a mechanism might be expected to yield a neural image in which items of type *A* produce high activation and items of type *B* low (or zero) activation. Access to such a neural image might further be expected to enable accurate estimation of the centroid of an ensemble of items of type *A* intermixed with to-be-ignored items of type *B*. Here it is shown that as the number of items in stimulus displays is increased, performance in estimating the centroids of horizontal (vertical) items amid vertical (horizontal) distractors degrades much more quickly and dramatically than does performance in estimating the centroids of white (black) items among black (white) distractors. Together with previous findings, these results suggest that, although human vision does possess bottom-up neural mechanisms sensitive to abrupt local changes in bar-orientation, and although human vision does possess and utilize top-down global attention filters capable of selecting multiple items of one brightness or of one color from among others, it cannot use a top-down global attention filter capable of selecting multiple bars of a given absolute orientation and filtering out bars of the opposite orientation in a centroid task.

Introduction

A target bar of a given orientation pops out from a field of distractor bars of some other fixed orientation provided the difference in orientation between target and distractor bars is large enough (Beck & Ambler, 1973; Foster & Ward, 1991a; Foster & Westland, 1995, 1998; Nothdurft, 1991, 1992, 1993, 1994; Sagi & Julesz, 1985; Treisman, 1985; Wolfe et al., 1992). This frequently replicated result might be taken to suggest that human vision possesses a number (perhaps a continuum) of mechanisms selective for different orientations. Since Hubel & Wiesel (1959, 1962), it has been known that area V1 of occipital cortex contains neurons selective for orientation. Subsequently, Philips and Wilson (1984) measured the size and properties of the receptive fields of such neurons. Using the orientation information available in V1, a mechanism to enable orientation popout might be realized in the brain by a retinotopically organized array of neurons, each monitoring a different location of the visual field for energy in the same fixed, oriented spatial frequency band. Under this theory, a target bar would pop out from a given field of distractor bars if (at least) one of these hypothetical mechanisms were sensitive to the orientation of the target bar but not to the orientation of the distractor bars.

Evidence supporting the existence of mechanisms selective for absolute orientation. Treisman & Gormican (1988) noted that it is easier to detect an oblique bar amid vertical distractors than a vertical bar amid oblique distractors. Foster & Ward (1991a) extended this finding by making careful measurements of the threshold orientation difference required to detect a target bar amid a field of distractors as function of the orientation of the distractors. The threshold orientation difference between target vs distractor bars was found to vary strongly as a function of the orientation of the background bars. In particular (in accordance with the results of Treisman & Gormican (1988)), the orientation increment threshold was minimal when the distractor bars were either vertical or horizontal and maximal when the target bar was either vertical or horizontal.

Foster & Ward (1991a) were able to model their results in terms of two orientation-selective mechanisms that we will denote *V* and *H*, both with broad, symmetric tuning curves, *V* with peak sensitivity to vertical and *H* to horizontal orientations. To understand how the model works, consider the condition in which the background bar segments are vertical and the target bar is slightly oblique from vertical. In this case, mechanism *V* is nearly useless for the classification task because it produces high, nearly equal responses to all bars in the display. However, mechanism *H* is very useful in this case; although the response of *H* to all bars in the display is much smaller than *V*'s response, *H*'s response to a vertical bar is substantially smaller than its response to a slightly oblique bar; thus the ratio of *H*'s response to the target bar vs its response the background bars is high. Under the model, (elaborated in detail by Westland & Foster, 1995), it is the relative activation produced by the target *versus* distractor bars rather than the difference in activation that is the crucial factor controlling performance in the classification task.

Foster & Westland (1998) investigated whether human vision possesses orientation-selective mechanisms more finely tuned than those discovered by Foster & Ward (1991a). Foster & Westland (1998) used displays similar to those used by Foster & Ward (1991a) again measuring threshold values of the angle between a target and background bar elements as a function of

background-element orientation. This time, however, they measured the angular increment threshold across a more finely sampled domain of angles: θ = 0,5, … 175 deg. (rather than the set θ = 0, 22.5, … 167.5 deg. used by Foster & Ward (1991a)). In addition to the two "coarse" mechanisms *V* and *H* discovered by Foster & Ward (1991a), this more fine-grained analysis revealed "intermediate" mechanisms with preferred orientations spaced at angles of approximately 35-50 deg. and "fine" mechanisms spaced at angles of approximately 10-25 deg. Although the two coarse mechanisms have clear peak sensitivities to the cardinal orientations (vertical and horizontal), the tuning of the intermediate and fine mechanisms tends to be more idiosyncratic across different participants.

Evidence supporting the existence of mechanisms selective for orientation contrast. The work of Foster and colleagues (Foster & Ward, 1991a, 1991b; Foster & Westland, 1995; Westland & Foster, 1995; Foster & Westland, 1998) suggests that human vision possesses an ensemble of mechanisms selective for different absolute orientations. Other research, however, both psychophysical (Nothdurft, 1991, 1992, 1993,1994; Koene & Li, 2007; Sagi & Julesz, 1995) and neurophysiological (Kastner, Nothdurft & Pigarev, 1999; Nothdurft, Gallant & Van Essen, 1999, 2000; Schofield & Foster, 1995; Westland & Foster, 1995b; Van Essen et al., 1989) suggests that human vision possesses at least one mechanism sensitive to "orientation contrast" (as opposed to absolute orientation). Such mechanisms are hypothesized to respond not to bars of a specific, absolute orientation but rather to visual input patterns in which the predominant direction of bar orientation undergoes an abrupt change. Indeed, Nothdurft (1992, 1993, 1994) presents evidence to support the claim that all preattentive sensitivity to orientation is conferred by mechanisms sensitive to orientation contrast (implying that human vision is devoid of mechanisms sensitive to absolute orientation).

The existence of mechanisms sensitive to orientation contrast is suggested by Fig. 1 (after Fig. 2 from Nothdurft (1994)). This image is densely populated with bars of different orientations. Throughout most of the display, the orientations of nearby bars change very gradually. However, the orientation gradient steepens abruptly at a locus of points in the image that defines a square. All bar orientations are represented in equal proportions (i) within the square, (ii) outside the square and (iii) at the boundary of the square. Thus, any mechanism tuned to a particular absolute orientation would be activated by only a few bars in the image; nonetheless, the entire square emerges very clearly as a figure against the background suggesting the existence of a mechanism that is activated by all points at the boundary of the square, i.e., by points of high orientation contrast.

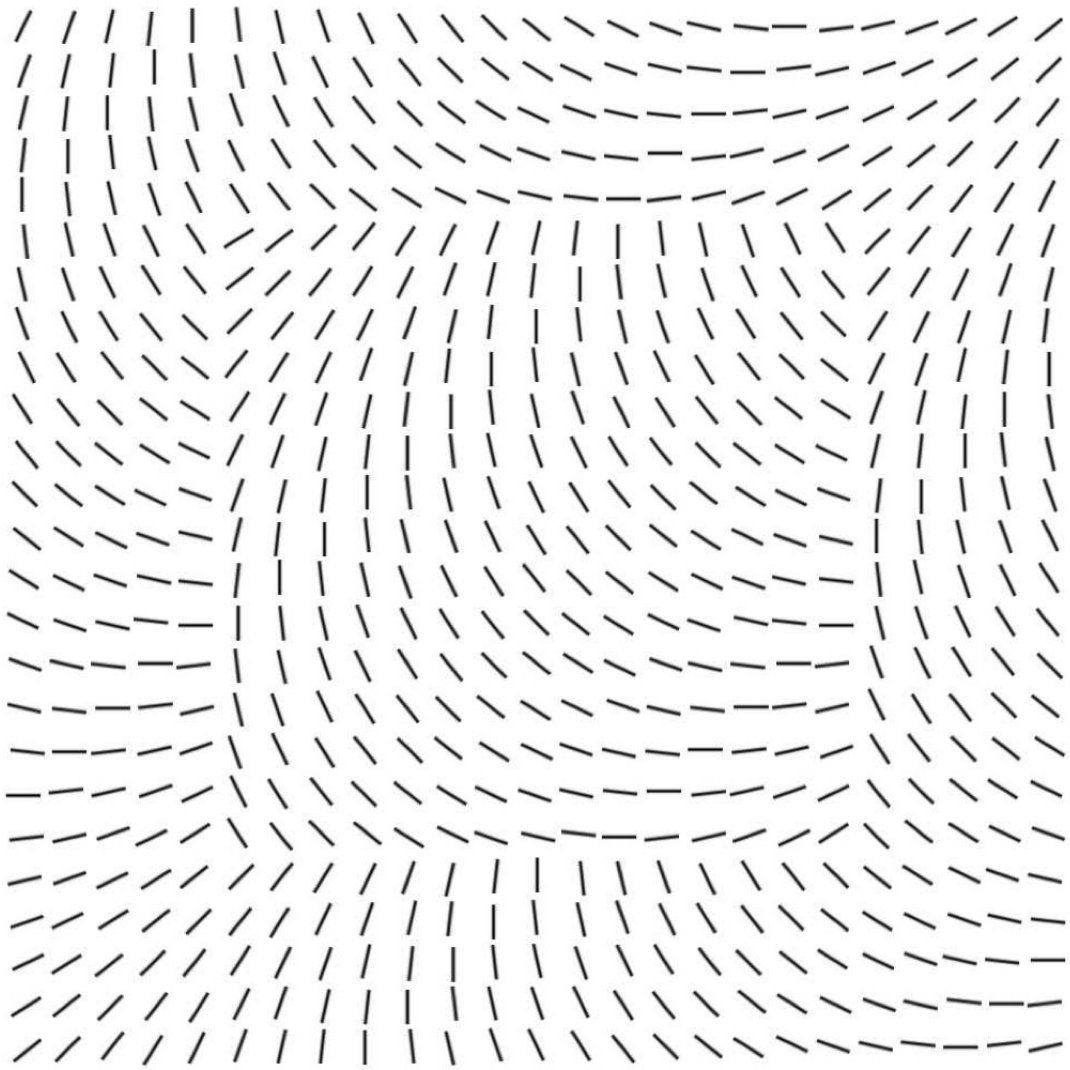

Figure 1 (after Nothdurft, 1994). Abrupt changes in orientation suffice to produce segmentation.

The current study. Although the research of Foster and colleagues (Foster & Ward, 1991a, 1991b; Foster & Westland, 1995; Westland & Foster, 1995; Foster & Westland, 1998) seems to suggest the existence in human vision of mechanisms selective for different absolute orientations, it is important to note that Foster's experiments involve participants looking for the existence of a unique bar—one that differs from everything around it. Participants are not made aware before each trial of the orientation of the target bar, so they are not attending to any one specific orientation. As such, the Foster experiments do not demonstrate whether or not orientation-selective mechanisms are accessible in an attention task. At issue is the question: Does (top-down) selective attention in human vision have access to mechanisms sensitive to absolute bar-orientation?

The current study provides evidence that the answer to this question is no. The strategy in brief is as follows: If participants do have access to a neural image (Robson, 1980) in which, for example, vertical bars produce strong activation but horizontal bars do not, then they should be able to use this neural image to make accurate estimates of the centroids of just the vertical bars

in displays comprising mixtures of vertical target bars and horizontal distractor bars. The experiments reported here demonstrate that participants are dramatically worse at this centroid-estimation task than they are if target vs distractor items are defined by brightness (black vs white) instead of by orientation, suggesting that although human vision possesses attention filters selective for black vs white (e.g., Drew, Chubb & Sperling, 2010), it does not possess corresponding attention filters selective for vertical vs horizontal bar-orientation.

Methods: Experiment 1

Participants

This study included 8 participants, 4 males and 4 females, one of whom is an author on this paper, and 3 of whom had no previous experience in centroid estimation tasks. All methods were approved by the UC Irvine Institutional Review Board, and each participant provided signed consent.

Stimuli

There were four different types of display (shown in Fig. 2). Each display contained equal numbers of horizontal black, horizontal white, vertical black and vertical white bars. For $n = 1,2,3,4$, "n-each" displays contained n of each of the four different types of bars. Thus, for example, every 3-each display contained 3 horizontal black, horizontal white, vertical black and vertical white bars—12 bars in all.

Stimuli were presented within a square region (640 by 640 pixels) subtending 20.63 deg. in width at the viewing distance of 53.5 cm. This stimulus region was circumscribed by a thin, black frame. Each bar was 3x30 pixels in size and subtended a visual angle of 0.967 degrees in length. Note: In the center of the display, 1 deg of visual angle equals 31.02 pixels.

The bar positions were drawn from a circular, bivariate Gaussian distribution subject to the following three constraints: (1) There were always at least five pixels between any bars. (2) The overall spatial extent of each stimulus cloud--its Dispersion--was fixed from trial to trial. This was accomplished as follows. Let $x = (x_1, x_2, \ldots, x_N)$ and $y = (y_1, y_2, \ldots, y_N)$ be the vectors of x- and y-coordinates of the item center locations in a given cloud, and let $\bar{X}$ and $\bar{Y}$ be the means of the vectors x and y. Then

$$Dispersion(x, y) = \left[\frac{1}{2N\text{-}2}\sum_{i=1}^{N}(x_i\text{-}\bar{X})^2 + (y_i\text{-}\bar{Y})^2\right]^{\frac{1}{2}} = 108 \text{ pixels} \approx 3.5°. \qquad (1)$$

(3) The centroid of (all of the bars in) each stimulus cloud (regardless of the number of items in the display) was drawn from a circular, bivariate normal density whose mean was the center of the stimulus field and whose standard deviation subtended 41 pixels (1.28 deg of visual angle).

Attention conditions

There were four different attention conditions. In the Attend-to-black condition, the participant attempted on each trial to mouse-click the centroid of all black items (all horizontal and vertical black bars) in the display, ignoring all white items. The Attend-to-white, Attend-to-horizontal and Attend-to-vertical conditions were defined analogously.

Each participant participated in 16 experimental conditions, performing each of the four different attention tasks using each of the four different types of *n*-each display.

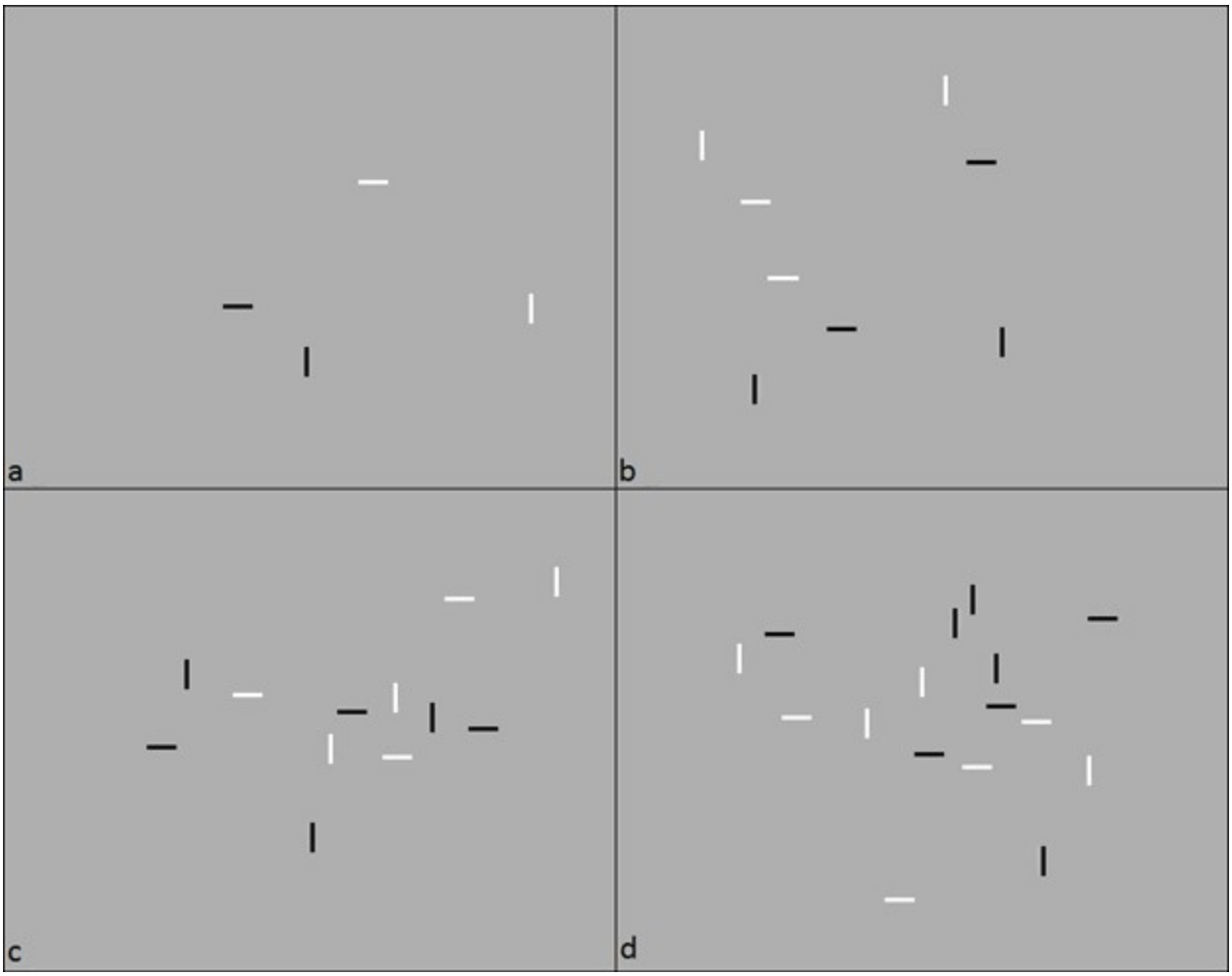

Figure 2. *Examples of stimuli from Experiment 1.* (a) 1 token of each bar type, (b) 2 of each bar type, (c) 3 of each bar type, and (d) 4 of each bar type.

The time-course of an experimental trial

The time-course of an experimental trial is illustrated in Fig. 3. The participant first viewed a blank, gray region circumscribed by a thin black square within which each stimulus cloud would be presented. The participant initiated the first trial with a button-press; each subsequent trial followed automatically after the previous trial. After the button press, the black square disappeared and never reappeared during the trial. The remaining blank gray field was presented

for one second at the outset of this and each subsequent trial. It was followed by the stimulus cloud (300ms), a blank, gray screen (33ms), and a random mask (300ms). The random mask consisted of 100 of each of the four types of bars, arranged in a regular 20x20 grid. Pilot experiments indicated that the 333ms SOA (Stimulus Onset Asynchrony, i.e., stimulus onset to mask onset) used here was sufficient to insure that performance in all attention conditions had reached asymptotic levels. After the mask, the screen again became blank. Once the participant moved the mouse, a crosshair cursor appeared in the center of the screen and tracked the mouse movements. The participant then moved the crosshair to the remembered center of the dot cloud and entered his/her response with a mouse click. A feedback display was then presented. This display contained the original stimulus, a crosshair indicating response location entered by the participant, and a bulls-eye indicating the correct response location. This feedback remained visible until the participant pressed the space bar, at which point the next trial began.

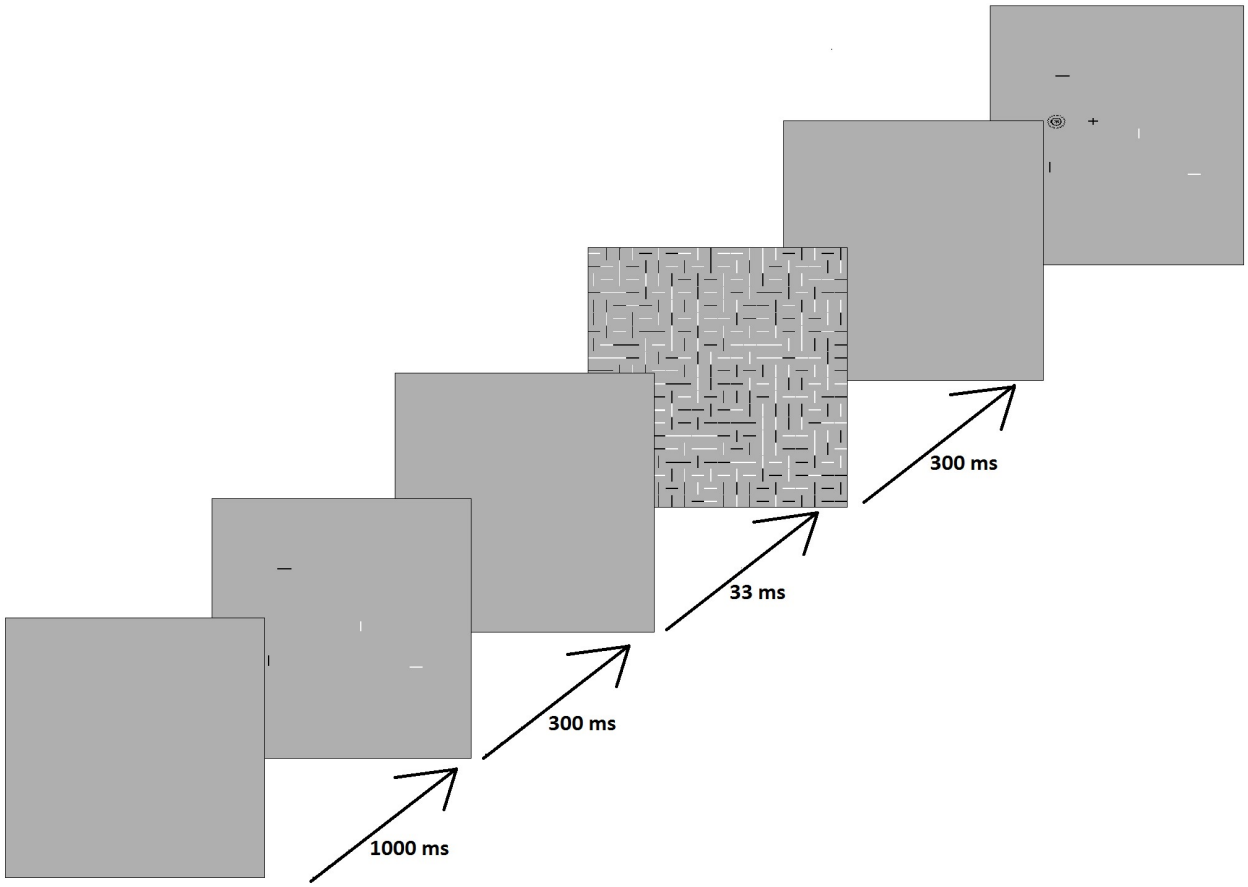

Figure 3. *The time-course of an experimental trial*. Each display is presented for the full time indicated. Feedback in this trial is based on the attend-to-black condition.

Design and Procedure

Initial Training – 1-2 sessions

Each participant first received general training in the centroid task. This was done to (1) minimize differences in the centroid computations used by different participants and (2) decrease the noise in the responses of individual participants. Subjects without any prior centroid task

experience received 800 training trials in a basic centroid task (see Sun, Chubb, Wright, & Sperling, In preparation). Each display in this training task comprised 8, square black dots, each subtending 0.3 deg. in width.

In addition, each participant received 800 trials of training with the specific stimuli used for this experiment, performing the task for 50 trials in each attention condition in each of the 1-each, 2-each, 3-each and 4-each displays in order to become familiar with the experimental set up.

Testing – 4 sessions

Upon completing his/her training, each participant ran 16 blocks per experimental session. The attention condition was fixed across all 16 blocks in a given session. The display type (1-each, 2-each, 3-each or 4-each) was fixed within a given block. Participants did these *n*-each blocks in order *n*=1,2,3,4,4,3,2,1,1,2,3,4,4,3,2,1. Each block consisted of 50 trials (for participants 1, 2, 3, and 4) or 53 trials (for participants 5, 6, 7, and 8). 45 of these trials were "full-set" trials of the sort described above; 5 other "target-only" trials contained only the target bars that would occur in a full-set trial without any of the distractor bars; and (for participants 5,6, 7, and 8 only) each block contained 3 additional "singleton" trials in which the display contained only a single bar which was randomly selected to be one of the two target types (e.g., either a horizontal white or a horizontal black bar in the Attend-to-horizontal condition) whose location was distributed identically to the location of the correct response on a full-set trial.

Subjects completed 4 such sessions, one for each of the four attention tasks. The order of these attention tasks was counterbalanced across subjects. Prior to each block, the participant was shown a gray screen containing each of the two types of target bars as well as each of the two types of distractor bars. Trials typically took about 2.5 sec, there was rest between blocks, and a session consisting of 800 trials took about 45 min.

Participant 1 later ran an additional session, in which each stimulus cloud contained 4 target bars and 16 distractor bars. This session included 16 blocks, 4 for each attention condition, and within each block there were 45 full-set trials, and 5 target-only trials.

Data analysis. We will write "*vw*", "*vb*", "*hw*" and "*hb*" for the vertical-white, vertical-black, horizontal-white and horizontal-black bar types. In the Attend-to-*Property* task (where *Property* is one of "black", "white", "vertical" or "horizontal"), the participant strives to click on the centroid of two *target* bar types and ignore the other two *distractor* types. (In the Attend-to-white task, for example, the target types are *vw* and *hw*, and the distractor types are *vb* and *hb*.) Thus, on a trial in which $type_i$ is the bar type of the ith item in the display and x_i and y_i are the *x*- and *y*-coordinates of its location, the target location ($Targ_x$, $Targ_y$) is

$$Targ_X = \frac{1}{W}\sum_{i=1}^{4n} f_{targ}(type_i)x_i \quad \text{and} \quad Targ_Y = \frac{1}{W}\sum_{i=1}^{4n} f_{targ}(type_i)y_i \qquad (2)$$

where f_{targ} is a real-valued function of bar type that assigns equal weight to two target types and weight 0 to the two distractor types, and W is the sum of $f_{targ}(type_i)$ taken over all $4n$ items i in the display.

Typically, however, the participant cannot achieve this goal; instead, the x- and y-coordinates of his/her response (R_X and R_Y) can be well-approximated by

$$R_X = \frac{D}{W}\sum_{i=1}^{4n} f(type_i)x_i + (1-D)x_{default} + Noise_X \qquad (3a) \text{ and}$$

$$R_Y = \frac{D}{W}\sum_{i=1}^{4n} f(type_i)y_i + (1-D)y_{default} + Noise_Y \qquad (3b)$$

where $x_{default}$ and $y_{default}$ are the x- and y-coordinates of the location toward which the participant's response is assumed to revert if he/she extracts only partial information on a given trial, D is a real number between 0 and 1, f is a real-valued function of bar type, W is the sum of $f(type_i)$ over all items i in the display, and $Noise_X$ and $Noise_Y$ are normally distributed random variables with mean 0 and standard deviation σ.

The function f is called the *attention filter* achieved by the participant in the given condition. The values $f(vw)$, $f(vb)$, $f(hw)$, and $f(hb)$ are called the filter weights of the four bar types. It should be noted that f is only defined up to an arbitrary multiplicative constant. For plotting purposes, we impose the constraint that $f(vw) + f(vb) + f(hw) + f(hb) = 1$.

The parameter D is called the *Data-drivenness* of participant's response. If $D = 1$, then the participant's response on a given trial is determined exclusively by the items in the stimulus cloud. At the other extreme, if $D = 0$, then the participant's response on a given trial is influenced not at all by the stimulus but instead by the combined effects of random noise plus a tendency to click on the fixed location ($x_{default}$, $y_{default}$).

Estimating model parameters. To estimate the model parameters $x_{default}$, $y_{default}$, f, D and σ, we proceed as follows:

1. Let L_X (L_Y) be the $N_{trials} \times 6$ matrix whose $(i,j)^{th}$ entry is the sum of the x-locations (y-locations) of all items of $type_j$ presented on trial i for $j<5$, is 1 (0) for $j=5$, and is 0 (1) for $j=6$.
2. Let R_X (R_Y) be the column vector of length N_{trials} whose i^{th} entry is the x-coordinate (y-coordinate) of the participant's response on trial i.

Then

1. Form the $2N_{trials} \times 6$ matrix M by appending the matrix L_Y to the bottom of L_X.
2. Form the vector R of length $2N_{trials}$ by appending R_Y to R_X.
3. Perform linear regression to derive the weights W minimizing $SS_{Residual} = ||MW-R||^2$.

Then, writing n_j, $j = 1,2,3,4$, for the number of items of type j in each stimulus cloud in a given condition, D (Data-drivenness) is estimated by

$$D = \sum_{j=1}^{4} W_j\, n_j \qquad (4)$$

For $j=1,2,3,4$, $f(j)$ is estimated by

$$f(j) = \frac{W_j}{\sum_{i=1}^{4} W_i} \qquad (5)$$

σ^2 is estimated by taking

$$\sigma^2 = \frac{SS_{residual}}{df}, \quad df = 2N - 6 \quad (6)$$

(where model parameters $x_{default}$, $y_{default}$, and D absorb 3 degrees of freedom and the attention filter f absorbs only 3 additional degrees of freedom because it is constrained to sum to 1). Finally, $x_{default}$ *and* $y_{default}$ are estimated by taking

$$x_{default} = \frac{W(5)}{1-D} \quad \text{and} \quad y_{default} = \frac{W(6)}{1-D} \quad (7)$$

A formal justification of this modeling method and a description of the methods used to estimate confidence intervals for model parameters is provided in Sun et al. (In preparation).

Quantifying performance. Note that actual performance (as described by Eq. (2)) can deviate from target performance (Eq. (1)) for several reasons. We quantify these deviations using:

1. *Imperfect Data-drivenness*. The Data-drivenness D of the participant's responses can be less than 1.
2. *Filter mismatch*. The attention filter f achieved by the participant can deviate from the target filter f_{targ}. This will cause the responses of the participant to deviate systematically from the correct responses. To quantify the degree to which the responses of the participant are immune from this sort of error, we use two descriptors of the participant's attention filter f. The first, called the *Selectivity*, is a ratio--the participant's attention weight for targets divided by the attention weight for distracters. Selectivity is given by

   $$Selectivity = \frac{f(t_1)+f(t_2)}{|f(d_1)|+|f(d_2)|} \quad (4)$$

   where t_1 and t_2 (d_1 and d_2) are the two item-types designated as targets (distractors) in the given attention condition; i.e., $f_{targ}(t_1) = f_{targ}(t_2) = 0.5$, and $f_{targ}(d_1) = f_{targ}(d_2) = 0$. When the attention filter achieved by a participant in a given condition closely approximates the target filter, Selectivity becomes very large. For this reason, it is convenient to plot $\log_{10}$(Selectivity) rather than Selectivity.
   The second, that we call *Filter-fidelity*, is a measure of how far away from the ideal filter (φ) the achieved filter (f) is. This is compared to how far away the worst possible filter a participant can get (f_{Worst}) is from the ideal filter. Specifically, we obtain Filter-fidelity with the formula

   $$Fidelity = 1 - \frac{||f\text{-}\phi||}{||f_{Worst}\text{-}\phi||} \quad (5)$$

   Notice that when the participant perfectly matches this target filter, this will be 1, and when the participant perfectly matches the worst possible filter, this will be 0.
3. *Random noise*. The standard deviation σ of the random variables $Noise_X$ and $Noise_Y$ is nonzero. This will cause the responses of the participant to deviate randomly from the correct responses. Although σ itself could be used to gauge the amount of random error corrupting the participant's responses, this model parameter is difficult to interpret because it depends on several factors (such as the size of stimulus display clouds) that are likely to vary across different experiments. To facilitate comparison

of results across experiments, we use a descriptor called *Efficiency* to quantify immunity to random error. Efficiency is the greatest lower bound on the proportion of display items that the participant must be using to compute his/her centroid estimates. If the trial-to-trial random errors $Noise_X$ and $Noise_Y$ were due solely to missing (i.e., failing to include) some of the items in the display in computing the centroid, then the participant would need to miss a proportion $p = 1 -$ Efficiency of the items on each trial in order for $Noise_X$ and $Noise_Y$ to have standard deviation σ. Thus, if the participant were to attain an Efficiency of 0.75, this would imply that he/she is including, on average, in his/her centroid computation at least three quarters of the items in the stimulus display. It should be emphasized, however, that if some of the random noise corrupting responses were due to some other source such as (i) early perceptual noise or (ii) instability in the centroid computation or (iii) motor noise, then the actual proportion of items included in computing the centroid would be higher than Efficiency.

Results: Experiment 1

Fig. 4 plots the 16 attention filters achieved in each of the four attention tasks for each of the 1-each, 2-each, 3-each and 4-each displays, across all participants. Though some participants had more experience with centroid tasks than others, there were no substantial differences between these groups. As such, all participants were pulled together for these analyses. With the 1-each display, they achieve an attention filter very similar to the target filter as witnessed by the high values of Selectivity and Filter-fidelity for all four attention conditions (Fig. 5). In addition, they achieve very high Efficiencies in all four attention conditions as well as high Data-drivenness values.

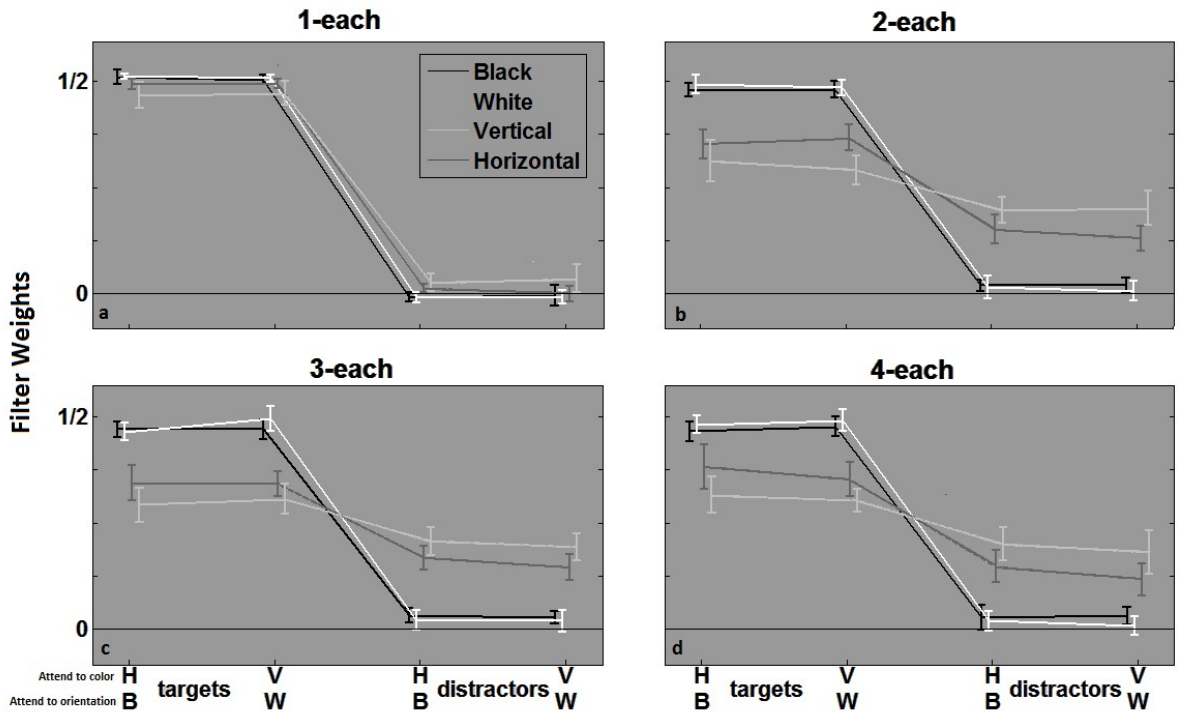

Figure 4. *Average attention filters and attention descriptors for all eight participants*. Each attention filter is plotted so that its responses to the two target items are plotted on the left and its responses to the two distractor items are plotted on the right. The H's and V's labeling abscissa tick marks indicate the orientations (Horizontal or Vertical) of the targets and distractors for the Attend-to-black attention filter (plotted in black) and Attend-to-white attention filter (plotted in white). The B's and W's indicate the color (black or white) of the targets and distractors for the Attend-to-vertical (dashed dark gray) and Attend-to-horizontal (dashed light gray) conditions. Error bars are 95% confidence intervals. Performance is excellent in all attention conditions with 1-each displays and remains excellent in Attend-to-black and -white conditions with 2-, 3- and 4-each displays, but is much worse in Attend-to-horizontal and -vertical conditions with 2-each, 3-each, and 4-each displays.

With each of the 2-, 3- and 4-each displays, however, the picture changes. Although the participants continue to perform very well in the Attend-to-white and Attend-to-black conditions, performance is dramatically impaired in the Attend-to-vertical and Attend-to-horizontal conditions. In particular, the attention filters in the Attend-to-vertical and Attend-to-horizontal conditions show very little Selectivity and Filter-fidelity for the target items vs the distractor items; in each of these two attention conditions, the participant gives roughly equal weight to all items in the stimulus cloud, yielding values of Selectivity and Filter-fidelity that are much lower for the Attend-to-vertical and Attend-to-horizontal conditions than for the Attend-to-black and Attend-to-white conditions. Note also that with the 2-, 3- and 4-each displays, Efficiency values are lower for the Attend-to-vertical and Attend-to-horizontal conditions than they are for the Attend-to-white and Attend-to-black conditions.

The difference in the effectiveness of attention filters for brightness vs for orientation is dramatized by the upper two panels in Fig. 5. The left upper panel plots mean Efficiency as a function of the number of each type of item in the stimulus display. The right upper panel plots the Selectivities in the same way, and the lower left and right panels show Filter-fidelity and Data-drivenness, respectively. Results for Attend-to-white, Attend-to-black, Attend-to-horizontal and Attend-to-vertical conditions are plotted in white, black, dashed light gray, and

dashed dark gray respectively. The values of Efficiency, Selectivity, and Filter-fidelity are similarly high in all attention conditions for the 1-each displays; however, the values of all three filter descriptors are much lower for the Attend-to-vertical and Attend-to-horizontal conditions than for the Attend-to-black and Attend-to-white conditions for all of the 2-each, 3-each and 4-each displays.

This observation is supported by Greenhous-Geisser corrected F tests. To insure homogeneity of variance, Filter-fidelities were arcsine-transformed prior to the analysis. There is a clear interaction between attention condition and display-type ($F(2.196, 15.371) = 57.132$, $p < 0.001$). However, this interaction is not seen when we restrict display types to 2-each, 3-each, and 4-each ($F(1.966, 13.759) = 1.694$, $p = 0.22$), nor is there a strongly significant difference between attention conditions for the 1-each displays ($F(1,7) = 7.168$, $p = 0.032$). The mildly significant difference seen there is due strictly to the lower Filter-fidelities achieved by participants in the Attend-to-horizontal condition vs the other three attention conditions. These results suggest that the interaction reflects mostly the difference between the 1-each vs the 2-each, 3-each and 4-each conditions, where performance in the Attend-to-horizontal and Attend-to-vertical conditions drops off suddenly in comparison the Attend-to-white and Attend-to-black conditions.

Interestingly, Data-drivenness does not show an analogous pattern; the curves for the Attend-to-vertical and Attend-to-horizontal conditions do not separate clearly from those for the Attend-to-white and Attend-to-black conditions. Rather Data-drivenness values remain high in all attention conditions across all displays.

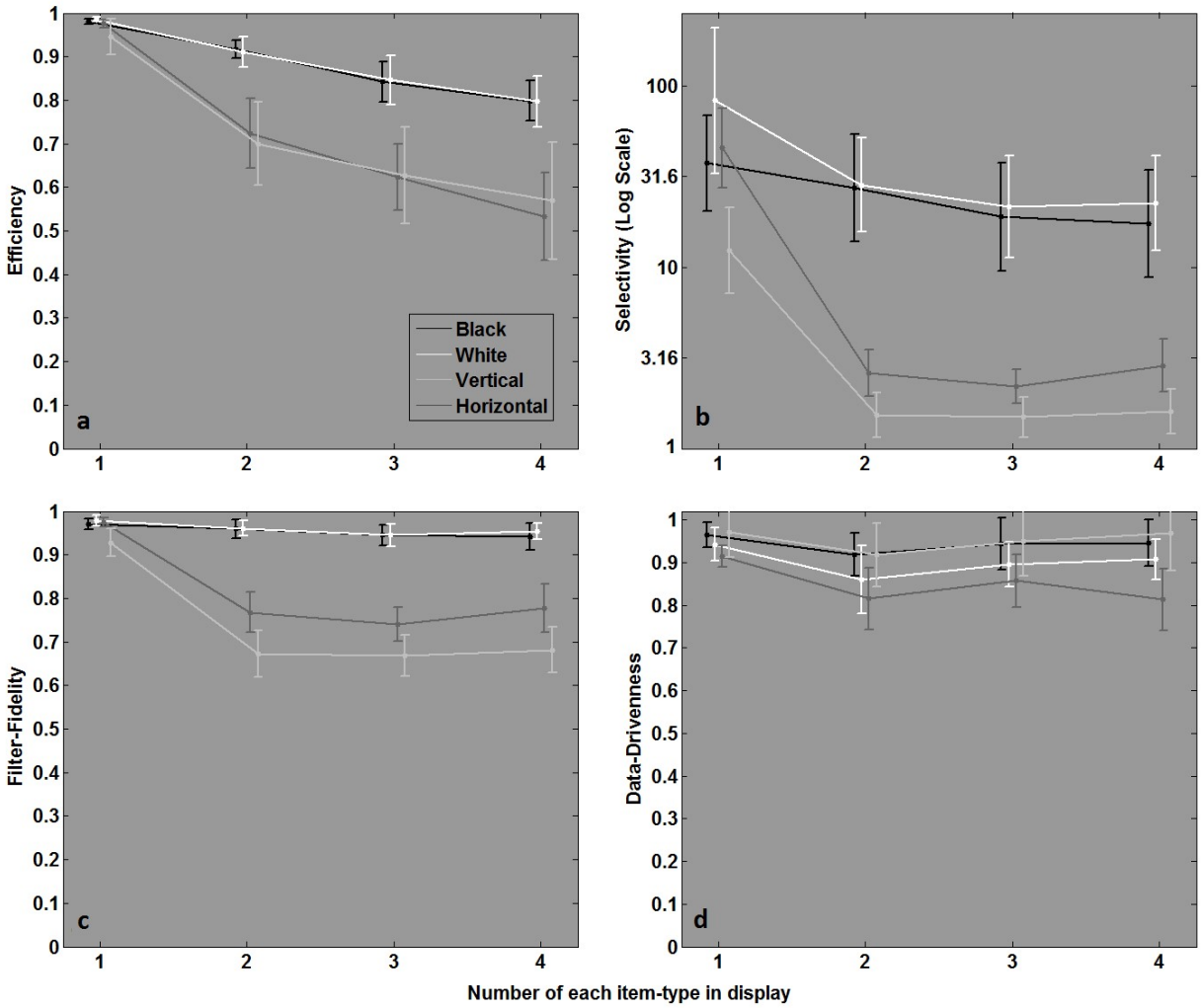

Figure 5. *Attention filter descriptors as a function of the number of items in stimulus displays for the four attention conditions*. These panels plot mean Efficiency (a), mean Selectivity (b), mean Filter-fidelity (c), and mean Data-drivenness (d) as a function of the number of each type of item in the stimulus display. Attention descriptor values for Attend-to-white, Attend-to-black, Attend-to-horizontal and Attend-to-vertical conditions are plotted in white, black, light gray, and dark gray respectively. Error bars are 95% confidence intervals for the mean over the 8 participants. Note that all three filter descriptors take high values for all attention conditions with the 1-each displays; however, Efficiency, Filter-fidelity, and Selectivity but not Data-drivenness are much lower in the Attend-to-vertical and Attend-to-horizontal conditions than in the Attend-to-black and Attend-to-white conditions for the 2-each, 3-each and 4-each displays.

Although the difference between the Filter-fidelities achieved in the Attend-to-vertical vs Attend-to-horizontal conditions seems small, it is statistically highly significant. Indeed, every participant in every *n*-each condition achieved higher Filter-fidelity in the Attend-to-vertical than in the Attend-to-horizontal condition. As this observation suggests, within any *n*-each condition, a two-tailed *t*-test confirms a significant difference between the Filter-fidelity achieved in the Attend-to-vertical vs. the Attend-to-horizontal condition. (For 1-each, $t(7)=4.0545$, $p=.0048$. For 2-each, $t(7)=5.1402$, $p=.0014$. For 3-each $t(7)=3.7373$, $p=.0072$. For 4-each, $t(7)=5.3660$, $p=.0010$.) Interestingly, the current results give us no reason to suppose that this effect depends on the number of items in the display; specifically, an *F*-test failed to reject the null hypothesis

that the mean difference between vertical and horizontal Filter-fidelities was equal across *n*-each conditions ($F(3,28)=1.8859$, $p=.1549$).

The current results provide no traction in understanding this effect. We speculate, however, that the vertical-vs-horizontal advantage may be related to the vertical gravitational frame within which the current experiments were conducted. As shown by Mikellidou et. al, 2015, the gravitational frame exerts a powerful influence on the "oblique effect" (i.e., the heightened sensitivity of participants to slight deviations in orientation from cardinal orientations (0° and 90°) versus from other orientations). If the gravitational frame operates slightly more strongly to accentuate the salience of vertical vs horizontal bars, this could produce effects of the sort observed in Experiment 1.

Discussion: Experiment 1

Attention-condition × number-of-item interaction. The results document a dramatic interaction between attention condition and the number of items in the display. Performance in all of the Attend-to-white, -black, -vertical and -horizontal conditions is at ceiling for the 1-each displays but degrades dramatically in Attend-to-vertical and Attend-to-horizontal conditions in comparison to the Attend-to-white and Attend-to-black conditions for the 2-each, 3-each and 4-each displays.

These results suggest that

1. Human vision does not possess a global attention filter selective for vertical (horizontal) vs horizontal (vertical) bar orientation. If human vision did possess such a global attention filter, then participants would be able to perform well in the centroid task in the Attend-to-vertical (Attend-to-horizontal) condition irrespective of the number of items in the display. Therefore, the fact that participants are dramatically impaired in the centroid task in the Attend-to-vertical and Attend-to-horizontal conditions compared to the Attend-to-white and Attend-to-black conditions for the 2-each, 3-each and 4-each displays suggests that, although human vision possesses a global attention filter selective for white (black) vs black (white), it does not possess a global attention filter selective for vertical (horizontal) vs horizontal (vertical) bar orientation.
2. Participants must have alternative strategies available for performing the centroid task in the Attend-to-vertical (Attend-to-horizontal) condition with 1-each displays. The current results reveal little about these strategies other than that they do not require access to a global attention mechanism; however, substantial previous research suggests that human vision has special capabilities for processing a small number of items (Cowan, 2001; Luck & Vogel, 1997; Pylyshyn & Storm, 1988; Trick & Pylyshyn, 1993, 1994).

There are, however, a couple alternative accounts of the results of Experiment 1 that need to be ruled out. Both of these accounts propose that human vision actually does possess global attention filters for vertical and also for horizontal orientation; however, Experiment 1 fails to reveal them.

The first account proposes that the operation of these global, orientation-selective, attention filters is disrupted in Experiment 1 by the very strong variations in brightness across which the filters are required to pool. Experiment 2 addresses this possibility by using stimulus displays in which all bars vary only in the task-relevant dimension.

The second account proposes that global attention filters for horizontal and vertical orientation in human vision exist but do not admit effective binding of item orientation with item location (which is required for centroid estimation). Experiment 3 addresses this possibility by using a task that requires participants to integrate information about feature identity across space but not to bind feature identities to locations. Specifically, in this task participants are asked to judge which of two item-types in a given stimulus display (e.g., vertical vs horizontal bars) is more numerous.

Methods: Experiment 2

Experiment 2 had 4 participants (2 male, 2 female), all of whom also participated in Experiment 1.

Stimuli were generated in exactly the same fashion as in Experiment 1. However, in this case, there was no variation in the unattended feature dimension (Fig. 6). As in Experiment 1, participants saw blocks consisting of 2 targets, 4 targets, 6 targets, or 8 targets, and an equal number of distractors. Participants completed 16 conditions (4 attention conditions by 4 numerosity conditions), each consisting of 2 blocks of 53 trials each, defined analogously to Experiment 1. Within any given block, each item in every display took the same fixed value on the irrelevant feature dimension. For example, in one block in the Attend-to-vertical condition, all bars (whether vertical or horizontal) in the stimulus presented on every trial were black (as shown in Fig. 6c for a 6-target display condition).

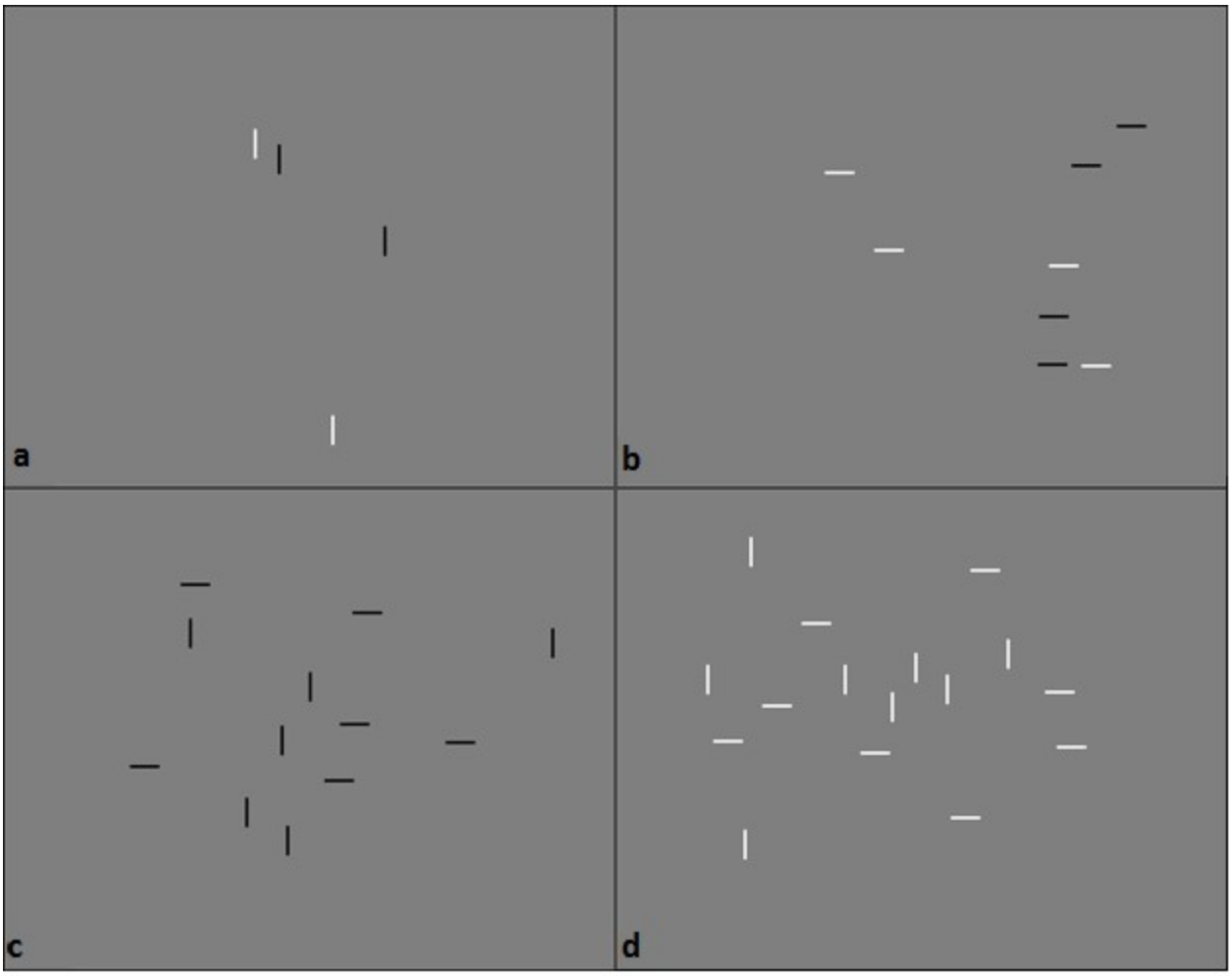

Figure 6. *Example stimuli for Experiment 2.* (a) 2 target bars (attending to a brightness), (b) 4 target bars (attending to a brightness), (c) 6 target bars (attending to an orientation), and (d) 8 target bars (attending to an orientation).

Results: Experiment 2

Fig. 7 plots the 16 attention filters achieved in the four attention tasks for all of the 2-target, 4-target, 6-target, and 8-target displays, for each participant. It should be noted that Participant 1 (plotted in black below) had much more practice with centroid tasks than the other three. As in Experiment 1, with the 2-target display, all participants achieve attention filters that are very similar to the target filter. With 2-target displays, Selectivity is high for all four attention conditions, as is Efficiency (Fig. 8).

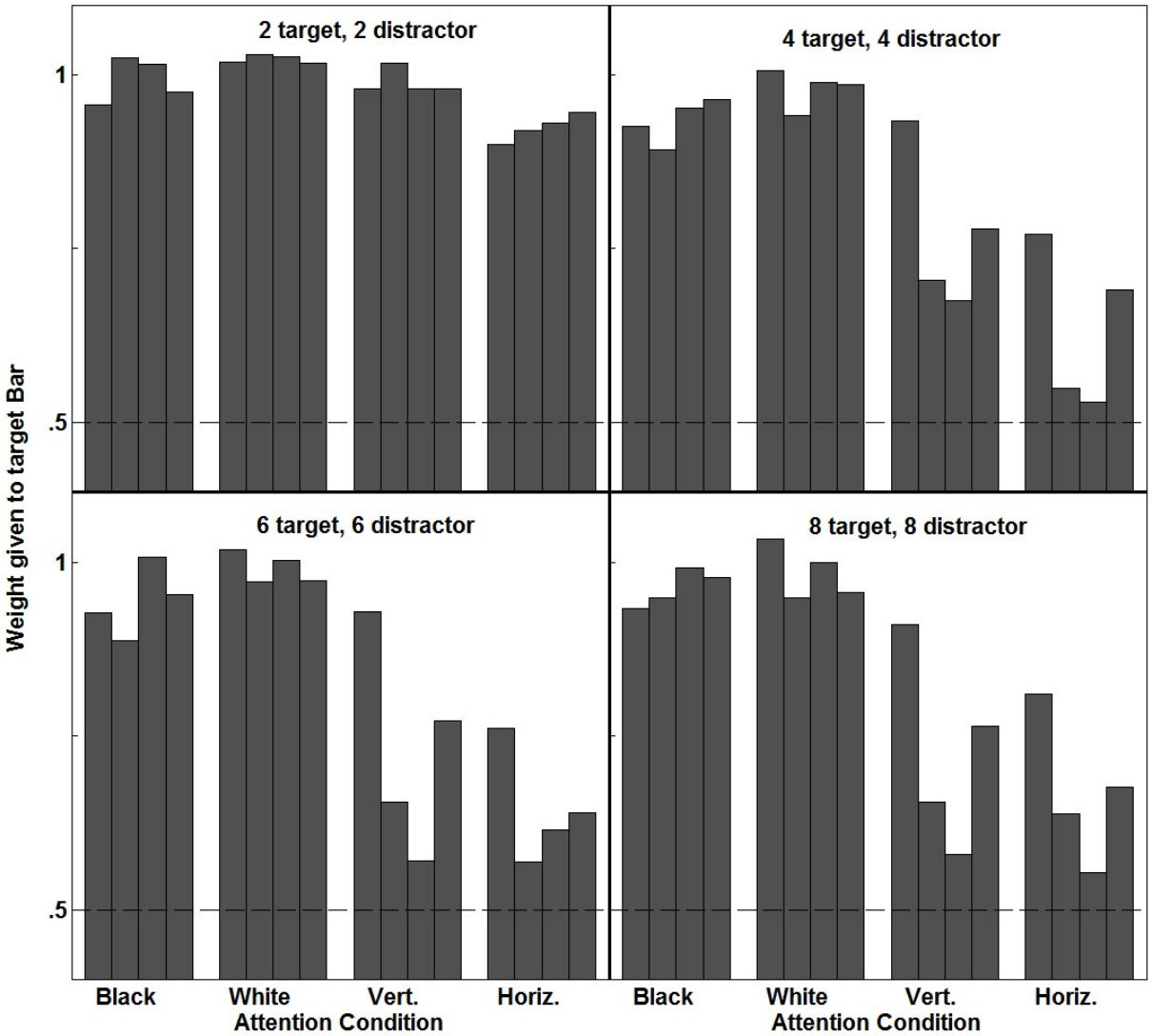

Figure 7. *Weight given to target bars for each attention and number of targets condition*. Each panel shows the proportion of weight a participant gave to the target bars (compared to distractor bars), plotted for each individual participant. Performance is excellent for all display numerosities for attention to black or white, and in the 2-target displays for attention to bar-orientation (a). In 4-target (b), 6-target (c), and 8-target (d) displays, attention filters for bar-orientation are much closer to chance performance.

Once again, however, with each of the 4-, 6- and 8-target displays, things change. The pattern seen in Experiment 1 remains clear. Participants can still perform very well in all conditions with only 2 targets, but as soon as the number of targets increases, performance in the Attend-to-vertical and Attend-to-horizontal conditions drops dramatically.

The difference in the effectiveness of attention filters for brightness versus for orientation is dramatized by Fig. 8. The top-left (top-right, bottom) panel plots mean Efficiency (Filter-fidelity, Selectivity) as a function of the number of each type of item in the stimulus display.

Results for Attend-to-white, Attend-to-black, Attend-to-horizontal and Attend-to-vertical conditions are plotted in white, black, light gray, and dark gray respectively. The values of Efficiency, Filter-fidelity, and Selectivity are similarly high in all attention conditions for the 2-target displays; however, the values of all three filter descriptors are much lower for the Attend-to-vertical and Attend-to-horizontal conditions than for the Attend-to-black and Attend-to-white conditions for all of the 4-target, 6-target and 8-target displays. It should be noted that Data-drivenness remains high in all 4 attention conditions, regardless of the number of targets. Specifically, it had a mean value of .92, and ranged from .89 to .96 with no clear difference between conditions and number of targets.

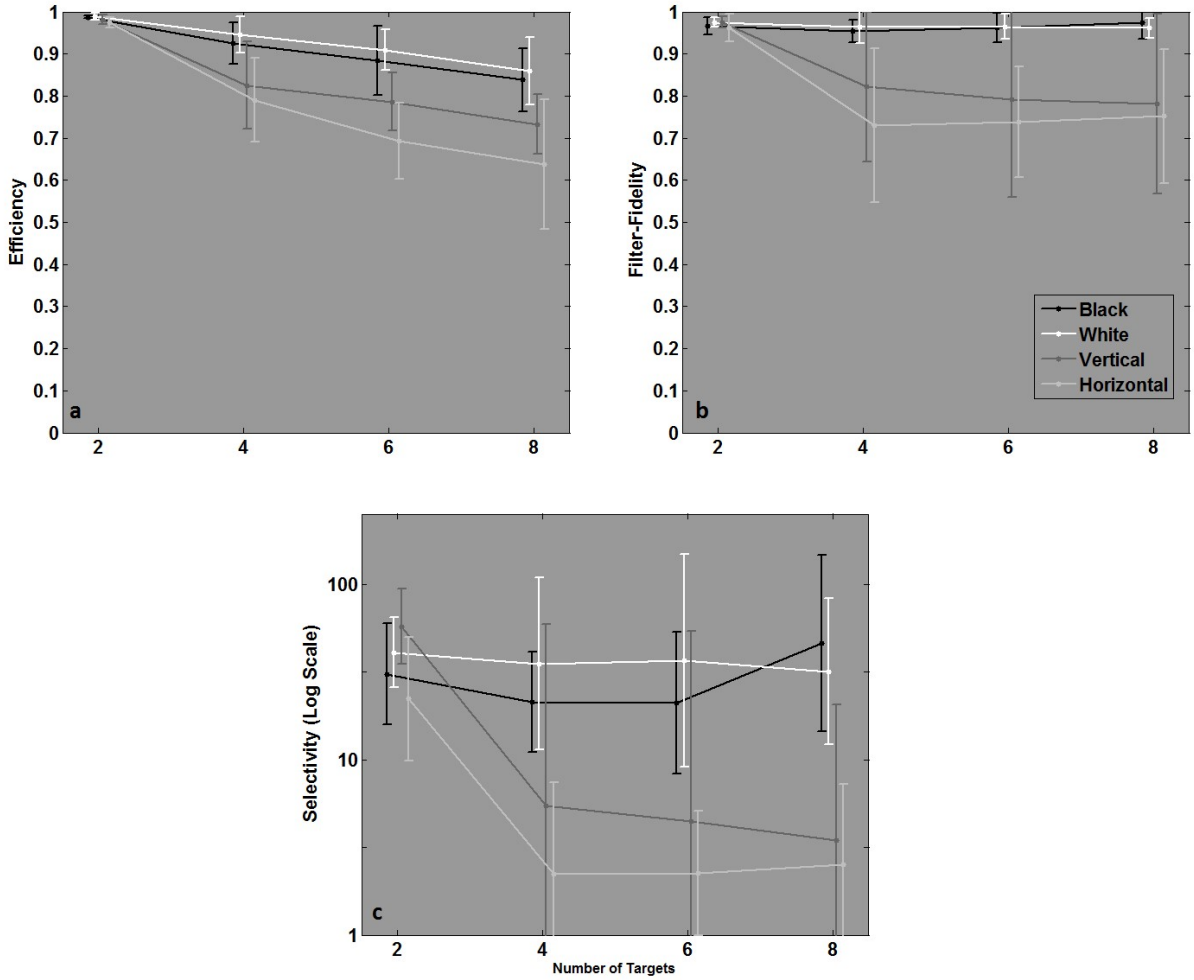

Figure 8: *Attention filter descriptors as a function of the number of targets in stimulus displays for the four attention conditions*. These panels plot mean Efficiency (a), mean Filter-fidelity (b), and mean Selectivity (c) as a function of the number of targets in the stimulus display. Attention descriptor values for Attend-to-white, Attend-to-black, Attend-to-horizontal and Attend-to-vertical conditions are plotted in white, black, dark gray, and light gray respectively. Error bars are 95% confidence intervals for the mean over the 4 participants. Note that all three filter descriptors take high values for all attention conditions with the 2-target displays; however, they all are much lower in the Attend-to-vertical and Attend-to-horizontal conditions than in the Attend-to-black and Attend-to-white conditions for the 4-target, 6-target, and 8-target displays.

F-tests (Greenhouse-Geisser corrected) confirm that the results of experiment 2 follow the same general pattern as those of experiment 1. To insure homogeneity of variance, tests were performed on arcsine-transformed Filter-Fidelities. A significant (but weak) interaction is observed between attention condition and display-type ($F(1.040,3.120) = 13.068$, $p = 0.034$). However, this interaction is not seen when display-types are restricted to 4-target, 6-target, and 8-target ($F(1.030, 3.091)=.834$, $p = 0.431$). There is a significant difference between Attention conditions in for 2-target displays ($F(1,3)=111.922$, $p = 0.002$) due primarily to the lower Filter-fidelities achieved in the Attend-to-horizontal condition vs the other conditions. The interaction again seems to reflect mostly the difference between the 2-target vs 4-, 6- and 8-target conditions, in which performance in the Attend-to-horizontal and Attend-to-vertical conditions drops off suddenly in comparison the Attend-to-white and Attend-to-black conditions.

Discussion: Experiment 2

The pattern of results here is similar to that seen in Experiment 1. Performance is strong in all of the Attend-to-black, -white, -vertical and -horizontal conditions for displays containing only 2 target items but quickly falls to a floor in Attend-to-vertical and Attend-to-horizontal conditions when the number of targets increases. Thus, in Exp. 1, the degradation in performance with 2-, 3- and 4-each displays in the Attend-to-horizontal and -vertical conditions compared to the Attend-to-white and -black conditions cannot be attributed solely to disruption of global orientation-selective filters by irrelevant variation in brightness between items.

With that said, however, it should be noted that participant 1 (black lines in Fig. 7) showed little degradation in performance with increasing display numerosity in either of the Attend-to-horizontal or Attend-to-vertical conditions. This participant (the first author) was highly practiced in many variants of the Attend-to-vertical and Attend-to-horizontal task conditions. By contrast, none of the other participants showed improvement in Experiment 2 compared to Experiment 1 (Fig. 9).

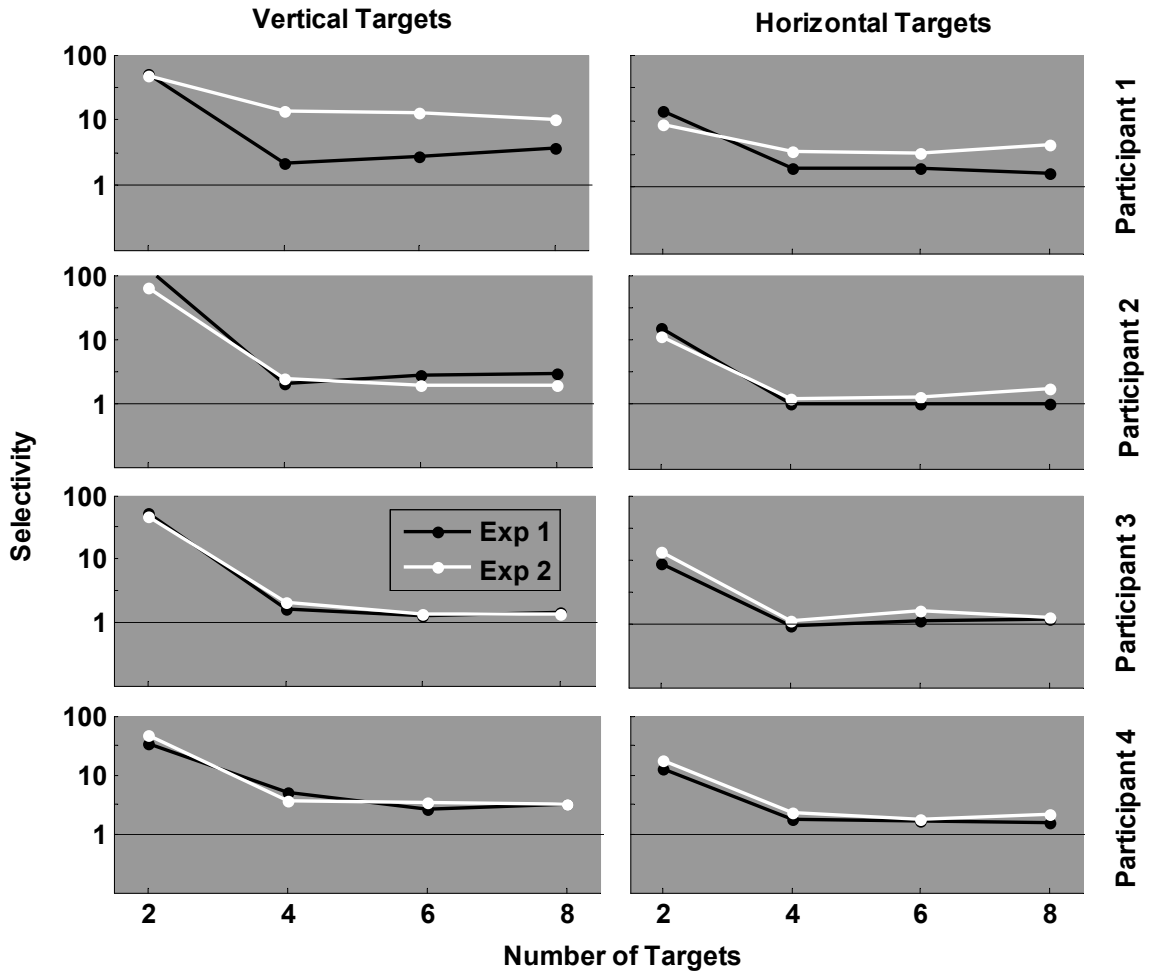

Figure 9. *Comparison of individual participants' performance (as reflected by Selectivity) in Experiment 1 vs. Experiment 2, when attending to orientation*. The black lines show participants' Selectivities in Experiment 1, and the white lines show their Selectivities in Experiment 2. The left plots show each participant's Selectivity when attending to vertical bars, and the right plots show the same participant's Selectivity when attending to horizontal bars. The trends in the two experiments are the same, and with the exception of Participant 1, everyone performed just as well in Experiment 1 as in Experiment 2.

When the number of targets is 4 or more, attending to orientation is much harder than attending to brightness, as in experiment 1. This effect is somewhat weaker than we observed in experiment 1, indicating perhaps that non-homogenous targets and/or distractors was making the task harder. However, since the effect remains regardless, this added difficulty is not what was causing the observed asymmetry between luminosity and orientation.

Methods: Experiment 3

Participants

This study used 3 subjects, 1 male and 2 females, one of whom is an author on this paper. All methods were approved by the UC Irvine Institutional Review Board, and each participant provided signed consent.

Task

The participant was instructed to attend to one feature dimension and indicate (by keypress) which feature within that dimension was more numerous. If the feature dimension was brightness, than the participant had to judge whether there were more black bars or more white bars. If the feature dimension was orientation, then he or she had to judge whether there were more vertical bars or horizontal bars.

Stimuli

The stimuli in Experiment 3 were similar to those in Experiments 1 and 2. Once again, there were 4 types of bars, black-vertical, black-horizontal, white-vertical, and white-horizontal. The locations of bars within displays were subject to the same constraints as in Experiments 1 and 2. However, in the current experiment, bars of different types were not presented in equal numbers. In the Brightness-difference condition, there were equal numbers of horizontal and vertical bars (8 of each), but unequal numbers of black and white bars (9 of one, 7 of the other). In the Orientation-difference condition, there were equal numbers of black and white bars, but unequal numbers of vertical and horizontal bars (9 of one, 7 of the other). Fig. 10 shows an example stimulus from the Brightness-difference condition.

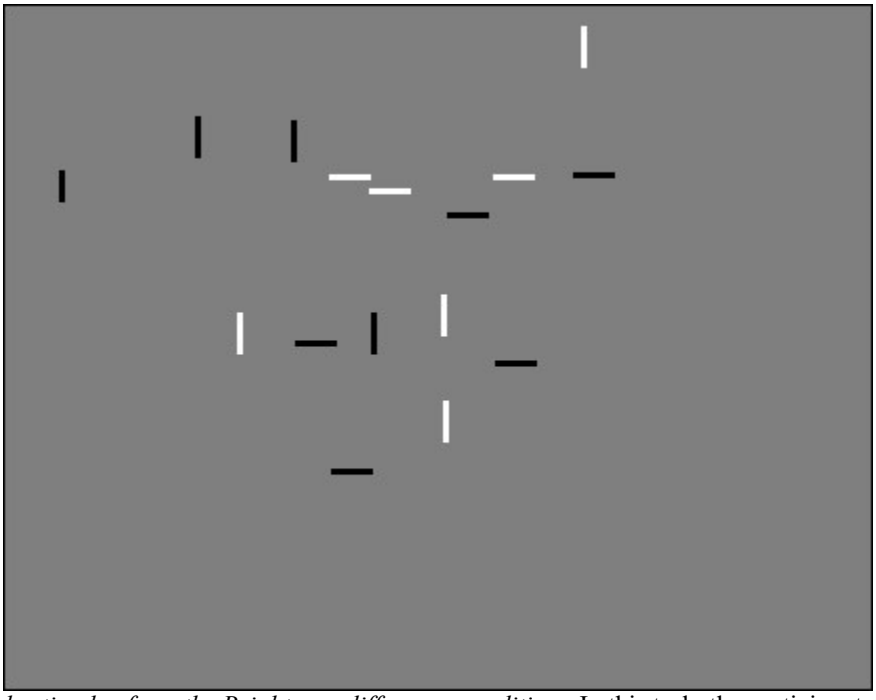

Figure 10. *Example stimulus from the Brightness-difference condition.* In this task, the participant strives to judge whether the black or the white bars are more numerous (In this example, the correct answer is BLACK.)

Design

Participants came in for 1 session of 8 blocks, each consisting of 60 trials. Task condition (Brightness-difference vs Orientation-difference) was varied by block, with Participants 1 and 3 doing Brightness-Orientation-Orientation-Brightness-Brightness-Orientation-Orientation-Brightness and Participant 2 doing the opposite. All participants were trained with 60 trials in each task prior to data collection.

Each trial began with one second of a gray screen, before the stimulus appeared for 300ms. The screen then went blank for another 33ms, before a mask appeared for 300ms. Finally, the participant was able to respond and was given feedback. The participant then initiated the next trial by pressing any key.

Results: Experiment 3

Fig. 11 shows the results for each of the three participants. Each participant achieves significantly higher D-prime in the Brightness-difference condition than in the Orientation-difference condition.

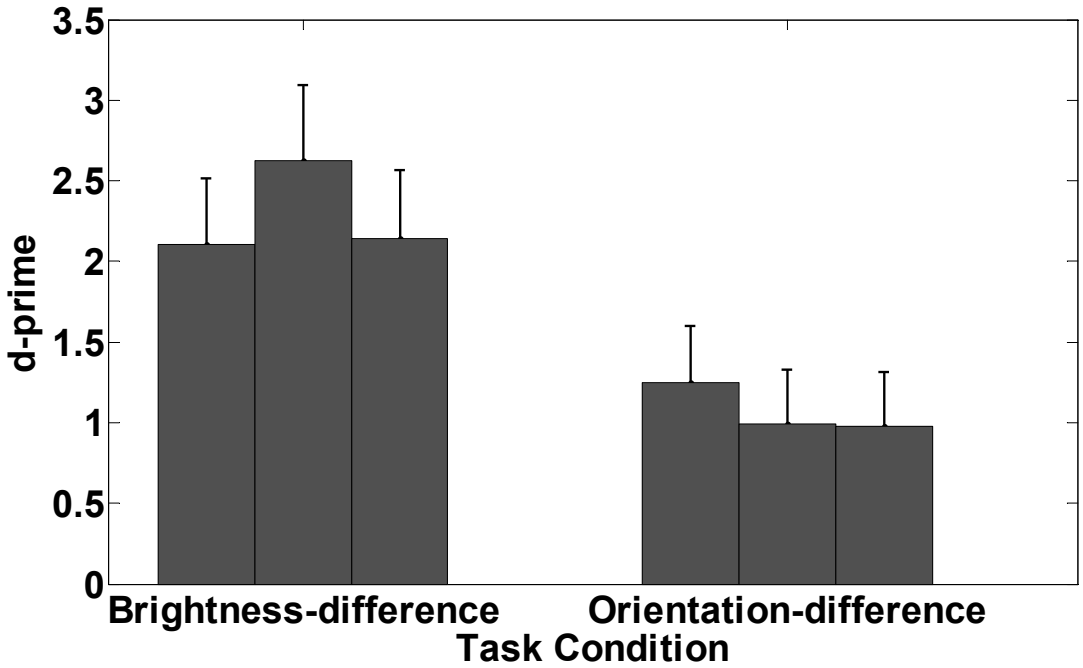

Figure 11. *Performance for each of three participants in Experiment 3 as a function of task condition*. A value of d-prime equal to 0 indicates chance performance. Error bars are 95% credible intervals derived from a Markov chain Monte Carlo simulation.

Discussion: Experiment 3

Previous research has made it clear that people are able to estimate the numerosity of a cloud of dots of one color amid distractor dots of another color (Halberda et. al., 2006). Further, people can simultaneously estimate the numerosities of two intermixed clouds of dots differing in color (Poltoratski & Xu, 2013). Thus, it is not surprising that participants achieved high d-prime values in the brightness-difference condition. What is most striking about the current results is the relative deficit in performance in the orientation-difference condition.

Even though the task in Experiment 3 did not require the participant to bind feature-identities to locations in the visual field, every participant continued to perform better in the Brightness-difference than in the Orientation-difference condition. We conclude that the difficulty experienced by our participants in the Attend-to-horizontal and Attend-to-vertical conditions in Experiment 1 was not due to the requirement (implicit in the centroid task) that information about item orientation be bound to item locations.

It is important to note that participants did achieve d-prime values well above chance even when asked about orientation. One plausible reason for this is that participants were able to extract some small subset of the bars and make their judgments based on the numerosity of that subset. According to simulations done on that assumption, subsets would need to contain between 5 and 10 bars (out of 16 total) to achieve a d-prime in the range we saw in the orientation conditions. If a similar strategy were being used for brightness, 14 to15 bars would need to be seen. This

suggests people are more efficient at extracting brightness information than orientation information.

General Discussion

An orientation contrast mechanism. Previous research has documented an asymmetry in the sensitivity of feature-based attention for orientation versus color (Huang, 2015; Wolfe et. al, 1995). The current results support these findings and suggest a possible explanation for all of these effects. Suppose, as proposed by Nothdurft (1992, 1993, 1994), that all human preattentive sensitivity to orientation is conferred by a mechanism that is selective for orientation contrast but indifferent to absolute orientation. In Experiments 1 and 2, vertical and horizontal bars occur in equal numbers; this imbues vertical and horizontal bars with equal orientation contrast implying that performance should be very poor in the centroid task.[1]

Note, however, that if the number of distractors is increased relative to the number of targets, then performance in the centroid task should improve for the following reason: because distractor bars are more numerous, target bars will tend to differ more strongly in orientation from the bars that surround them than will distractor bars. In this case, the target bars will tend to produce higher activation than distractor bars in the orientation-contrast mechanism that we hypothesize is used for the task.

If participants do not have access to attention filters tuned to absolute orientation, how do they achieve better-than-chance performance in the Attend-to-horizontal and Attend-to-vertical conditions of Exp. 1? It is clear from Fig. 5b that participants achieve Selectivities significantly greater than 1. (None of the confidence intervals for the Attend-to-vertical or -horizontal conditions contains 1 for any of the 1-, 2-, 3- or 4-each displays.) One possible explanation is the following. Regardless of whether a participant can access an attention filter tuned to a specific bar orientation, he or she may be able to locate and identify several items of the target orientation in any given stimulus. If he/she gives these few items enhanced weight in his/her centroid computation, Selectivity will be elevated above 1. However, to the extent that he/she allows the unidentified items in the display to influence performance, his/her Selectivity will be suppressed. Such a strategy would plausibly yield the low, but above-1 Selectivities observed in Exp. 1.

The *V* and *H* mechanisms revisited. The broadly-tuned mechanisms *V* and *H* with peak sensitivities to vertical and horizontal orientations hypothesized by Foster and colleagues (Foster & Ward, 1991a, 1991b; Foster & Westland, 1995; Westland & Foster, 1995; Foster & Westland, 1998) should be ideally suited for the Attend-to-vertical and Attend-to-horizontal centroid task

[1] The horizontal and vertical bars in an *n*-each display produce indistinguishable responses in a mechanism selective for orientation contrast for the following reason. In one of the *n*-each displays used in the current experiment, every bar, whether horizontal or vertical, shares the stimulus space with *n*-1 bars of the same orientation and *n* bars of the opposite orientation. Thus, on average, the expected orientation contrast relative to its context of any bar in an *n*-each display is the same. For this reason, a mechanism selective for orientation contrast is useless for purposes of estimating the centroid of the target bars of one orientation in an *n*-each display while ignoring the distractor bars of the opposite orientation.

conditions of Experiments 1 and 2, as well as for the attend-to-orientation condition of Experiment 3. The *V* (*H*) mechanism should be strongly activated by vertical (horizontal) bars and very weakly activated by horizontal (vertical) bars. Thus, the finding that performance is so poor in the Attend-to-vertical and Attend-to-horizontal conditions with each of the *2*-each, *3*-each and *4*-each displays suggests that these mechanisms cannot be utilized in a global selective attention task.

Absolute orientation versus orientation contrast. It has been argued (Friedman-Hill & Wolfe, 1995) that people do have sensitivity to absolute orientation in a search task. However, we do not find the empirical evidence for this claim compelling.[2] Although human vision has mechanisms that enable global selection of items based on contrast, color, shape, and other attributes, we find that human vision possesses no mechanism that enables selection of multiple items based on absolute bar-orientation. A broader range of tasks is needed to further explore the inability of observers to access mechanisms sensitive to absolute orientation. We conjecture that human vision discards information about absolute orientation because in the natural world, local orientation depends on both object and head position and is therefore not a sufficiently useful indicator to prompt the evolution of attention filters for absolute orientation. On the other hand, human vision may have evolved a preattentive mechanism sensitive to orientation contrast; this reflects the fact that orientation contrast is independent of absolute orientation and often does indicate something worth noticing.

Summary

Participants were tested in a task requiring them to use top-down selective attention to estimate the centroids of target items with one feature value while ignoring distractor items with the opposite feature value. Each display contained equal numbers of white horizontal bars, white

[2] Friedman-Hill & Wolfe (1995) tested participants in a range of different search tasks, two of which are relevant for the current discussion. In their "conjunction search" task, participants judged whether a vertical-red target bar was present among horizontal-red and vertical-green distractor bars. In their "subset search" task, the target was a red bar of orientation X° among green bars of orientation X° and red bars of orientation Y°. The orientations, X° and Y°, were randomly chosen subject to the constraint that they had to differ by at least 30°. The subset search task thus forced participants to adopt a strategy of searching for a bar with an oddball orientation within the cohort of red bars. Mean reaction time in the subset search condition was significantly slower than in the conjunction search condition. The authors inferred from this result that participants have access to a different (and more effective) strategy in the conjunction task than searching within the red bars for an oddball orientation. Specifically, they suggest that in the conjunction task the participant is able to recruit spatially parallel processes selective for each of (1) bar redness and (2) bar verticality to converge on the target. Note, in particular, that this interpretation suggests the existence of a spatially parallel mechanism selective for absolute bar orientation.

We contend, however, that these findings admit other interpretations. The stimuli used in the subset task are intrinsically more complicated than those used in the conjunction task. Most importantly (in contrast to the conjunction task in which the target bar always differed from same-color distractor bars by 90°), in the subset search task the difference in orientation between the target bar and same-color distractor bars varies randomly (in both sign and magnitude) from trial to trial; moreover, this orientation difference is usually substantially less than 90°. If indeed the only orientation-sensitive mechanism available to participants is sensitive not to absolute orientation but rather to orientation contrast, then this disparity between the two tasks would be highly likely to impair performance selectively in the subset task vs the conjunction task.

vertical bars, black horizontal bars and black vertical bars. In the Attend-to-black (Attend-to-white) condition, the participant strove to mouse-click the centroid of all the white (black) target bars in the display while ignoring the black (white) distractor bars. In the Attend-to-vertical (Attend-to-horizontal) condition, the participant strove to mouse-click the centroid of all the vertical (horizontal) target bars in the display while ignoring the horizontal (vertical) distractor bars. Each participant was tested in each attention condition with displays that contained 1-each, 2-each, 3-each or 4-each of the four bar-types.

With 1-each displays participants performed very well in all attention conditions. However, with 2-each, 3-each and 4-each displays, performance was greatly impaired in the Attend-to-vertical and Attend-to-horizontal conditions compared to the Attend-to-white and Attend-to-black conditions.

The different pattern of results observed with the 1-each displays suggests that this variant of the task (which required participants to locate the two target bars and click on the midpoint between them) afforded strategies that were not available with the 2-each, 3-each and 4-each displays.

The fact that the Attend-to-white condition (analogous remarks apply to the Attend-to-black condition) yielded excellent performance with each of the 2-each, 3-each and 4-each displays suggests that in this condition, (1) participants were able to access an attention filter that produced high activation in response to white target bars and low (or zero) activation to black distractor bars, and (2) the output of this mechanism provided a neural image from which the centroid of the target bars could be extracted. The dramatic impairment in performance observed in the Attend-to-horizontal and Attend-to-vertical conditions with each of the 2-each, 3-each and 4-each displays suggests that no such strategy was available to globally select targets of one orientation from distractors of a perpendicular orientation. This suggests that participants do not have access to an attention filter that produces high activation in response to target bars of a given orientation and low (or zero) activation to distractor bars of the opposite orientation.

Additional control experiments demonstrated that the impairment of performance in the attend-to-horizontal and attend-to-vertical conditions (compared to the attend-to-black and attend-to-white conditions) persisted (1) if the stimulus displays used in the attend-to-horizontal and attend-to-vertical conditions of the centroid task were composed of items that all had the same brightness, and (2) in the context of a task (requiring the participant to judge which of two item types was more numerous in the stimulus display) in which the participant was not required to bind information about item-type to locations in the visual field.

The current results thus suggest that the high levels of performance typically achieved in tasks based on differences in orientation are mediated not by top-down mechanisms sensitive to absolute orientation but rather by a bottom-up mechanism that senses local differences in orientation.

CHAPTER 2
Comparing Efficiencies in Estimating Centroids and Judging Numerosity

Abstract

Given a brief display of a dot cloud of a few to several dozen dots, observers are good both at judging the number of dots in the cloud (its numerosity) and also at estimating its center of gravity (its centroid). To judge the numerosity of a dot cloud, the observer must merely register the presence of each dot, whereas to estimate the centroid, the observer must register both a dot's presence as well as its location. Therefore, one might expect observers to be more efficient in judging numerosity than in estimating centroids. This study tested this prediction. Stimuli were clouds comprised of 11 to 32 black dots. In the Numerosity Task, on each trial, the participant judged whether the stimulus contained fewer or more dots than the average number in that particular condition. In the Centroid Task, on each trial, the participant strove to mouse-click the centroid of the dot cloud. To compare performance across both tasks, we use a metric based on the assumption that all errors result from random decimation of the display. Specifically, for each task, we assume that an ideal observer processes a display from which dots have been removed independently with some probability 1-*p*. A participant's Efficiency for that task is the value of *p* that for which the ideal observer's performance exactly matches the subject's performance. For every subject in every condition, Efficiency was higher in the centroid task than in the numerosity task. Conclusion: Surprisingly, efficiency is equal or higher for estimating centroids than for judging numerosity. This suggests either that (1) centroids are computed without computing numerosity or else (2) the implicit estimate of numerosity utilized in the context of the centroid computation is not available to the numerosity computation.

Introduction

In the centroid task, the observer briefly views a stimulus consisting of a spatially random array of items and strives to mouse-click the centroid of the items. (Some task variants, not considered here, require the participant to try to give differential weight to different types of items.) Previous research (Drew, Chubb, & Sperling, 2010; Sun et. al., 2015) has demonstrated that performance in centroid tasks can be highly efficient. However the nature of the computations observers use in estimating centroids remains mysterious. In particular, although the centroid is the average position of the target items, it is unclear how the brain might compute such an average.

A standard way to compute a centroid is to sum the locations of all the items in the stimulus and divide the result by the number of items in the stimulus. On the other hand, recent work suggests that human vision possesses three different mechanisms for computing numerosity (Burr, Turi, & Anobile, 2010; Anobile, Cicchini, & Burr, 2016). If indeed numerosity were a visual primitive, it seems likely that the mechanism that computes it might be used as well in the centroid computation. If so, then performance achieved by observers in centroid estimation would be limited by their performance in judging numerosity.

Further evidence suggesting that the centroid computation may recruit the numerosity mechanism can be seen by looking at effects of dot density. In particular, in both numerosity judgments (Allik & Tuulmets, 1991; Dakin et. al, 2011) as well as in centroid judgments (Moreland & Boynton, 2016; Rashid & Chubb, 2017) dots in stimulus regions that are densely populated by other dots exert less influence on responses than do isolated dots. .Such a density effect would be expected if the numerosity mechanism were used in performing the centroid computation.

On the other hand, there are reasons to doubt that accuracy in the centroid task is limited by the accuracy of numerosity estimation. One consistent finding in numerosity estimation tasks is that the standard deviation of estimates increases in proportion to the number of items (Krueger, 1982; Whalen, Gallistel, & Gelman, 1999). If performance in a centroid task were limited by performance in a numerosity estimation task, then performance would also suffer in a centroid task as numerosity increases. However, this is not what is seen. Error in the centroid task seems to increase far more slowly with increasing numerosity than error in the numerosity task. For example, Drew et al. (2010) found that the errors in centroid estimates for 16-dot displays were no greater than the errors for 8-dot displays.

The purpose of the current study is to address these issues by directly comparing the efficiency in estimating centroids versus efficiency in judging numerosity. Previous research has shown that human vision possesses at least three distinct mechanisms that can be used for judging numersosity under different conditions. With very few items (i.e., 4 or fewer), people can quickly and accurately tell exactly how many items there are (Kaufman & Lord, 1949; Revkin et. al., 2008). At the other extreme, when there are many items (i.e., around 70 or more), people seem to use a statistic primarily sensitive to texture density to estimate numerosity (Anobile et. al., 2015; Anobile, Cicchini, & Burr, 2014). By contrast, substantial evidence suggests that for judging the numerosity of stimuli composed of from between 5 to 70 well-individuated items,

the statistic people use is indeed selective for the numerosity *per se* (as opposed to either the area or the density) of the stimulus cloud (Anobile, Cicchini, & Burr, 2016). The focus of the current study is on this intermediate numerosity mechanism. Accordingly, the stimuli we use are constructed specifically to engage this mechanism primarily.

Methods

All methods were approved by the UC Irvine Institutional Review Board. There were 8 adult participants (5 female). All participants provided informed consent prior to participation. Participants were tested in both the centroid and numerosity tasks using the same stimulus displays. In any given condition, on each trial, the participant was presented first with a blank gray screen that lasted 1000 ms. Then the 300 ms stimulus display appeared; this display consisted of some number of black dots within a gray stimulus field circumscribed by a thin square black frame (Figure 1). At the nominal viewing distance of 60 cm, the stimulus field subtended 17.76 deg. of visual angle, and each (5x5 pixel) dot subtended 0.14 deg. The stimulus was then replaced by a blank screen for 33 ms and finally obscured by a mask that consisted of a 16x16 grid of dots for 300 ms. The participant then had to make a response that depended on the task, after which feedback was presented. The feedback lasted until the participant pressed space bar, which initiated the next trial.

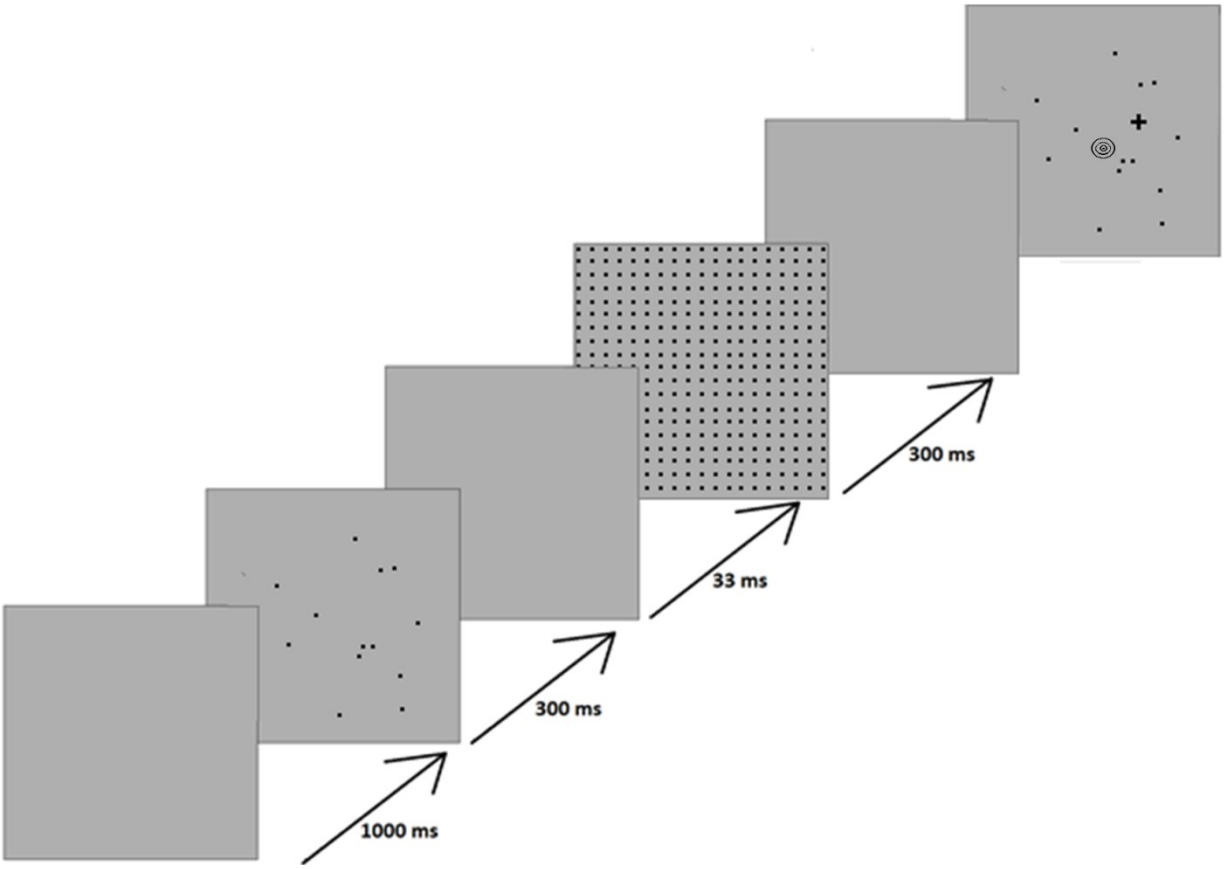

Figure 1: Time course of a trial. The feedback frame shown here is from the Centroid Task. The crosshair indicates the location of the participant's response and the bullseye indicates the true centroid. In a Numerosity Task, the feedback frame contained text (green if the response was correct, red if incorrect) reminding the participant what response he or she entered and whether it was correct or incorrect.

There were two different tasks (the centroid and numerosity tasks) and two different set-size conditions (the small-set and large-set conditions). In the small-set condition, there were randomly 11, 12, 13, 14, 15, or 16 dots on the display, whereas in the large-set condition, there were 22, 24, 26, 28, 30, or 32. In the Centroid task, the participant strove to mouse-click the centroid of all of the dots. In the context of the centroid task (but not the numerosity task), a small number of stimuli contained only one dot; the data from these "singleton" trials were used to estimate the component of the participant's response error that had nothing to do with computing a centroid. On singleton trials, the location of the single dot was distributed identically to the correct response location (centroid) of a random trial from the current condition. In the Numerosity task, the participant had to indicate by key-press if there were fewer dots than average (11, 12, or 13 in the small-set condition, and 22, 24, or 26 in the large-set condition) or more than average (14, 15, or 16 in the small-set condition, and 28, 30, and 32 in the large-set condition).

Dots were placed as described in Sun et. al (2016), with one exception (Figure 2). The dispersion[3] of the stimulus cloud, instead of being fixed (as recommended by Sun et al, 2016), was roved independently of the number of dots. This prevented participants from using density as an informative cue in numerosity tasks. Specifically, dispersion was normally distributed, with a mean of 108 pixels and a standard deviation of 15 pixels.

[3] Dispersion is given by $Dispersion(x,y) = \left[\frac{1}{2N_{stim}-2}\sum_{i=1}^{N_{Stim}}(x_i-\bar{X})^2 + (y_i-\bar{Y})^2\right]^{\frac{1}{2}}$, where N_{stim} is the number of dots in the stimulus, x_i is the x-coordinate of the ith dot, y_i is the y-coordinate of the ith dot, $\bar{X}$ is the mean x-coordinate of all dots, and $\bar{Y}$ is the mean y-coordinate of all dots.

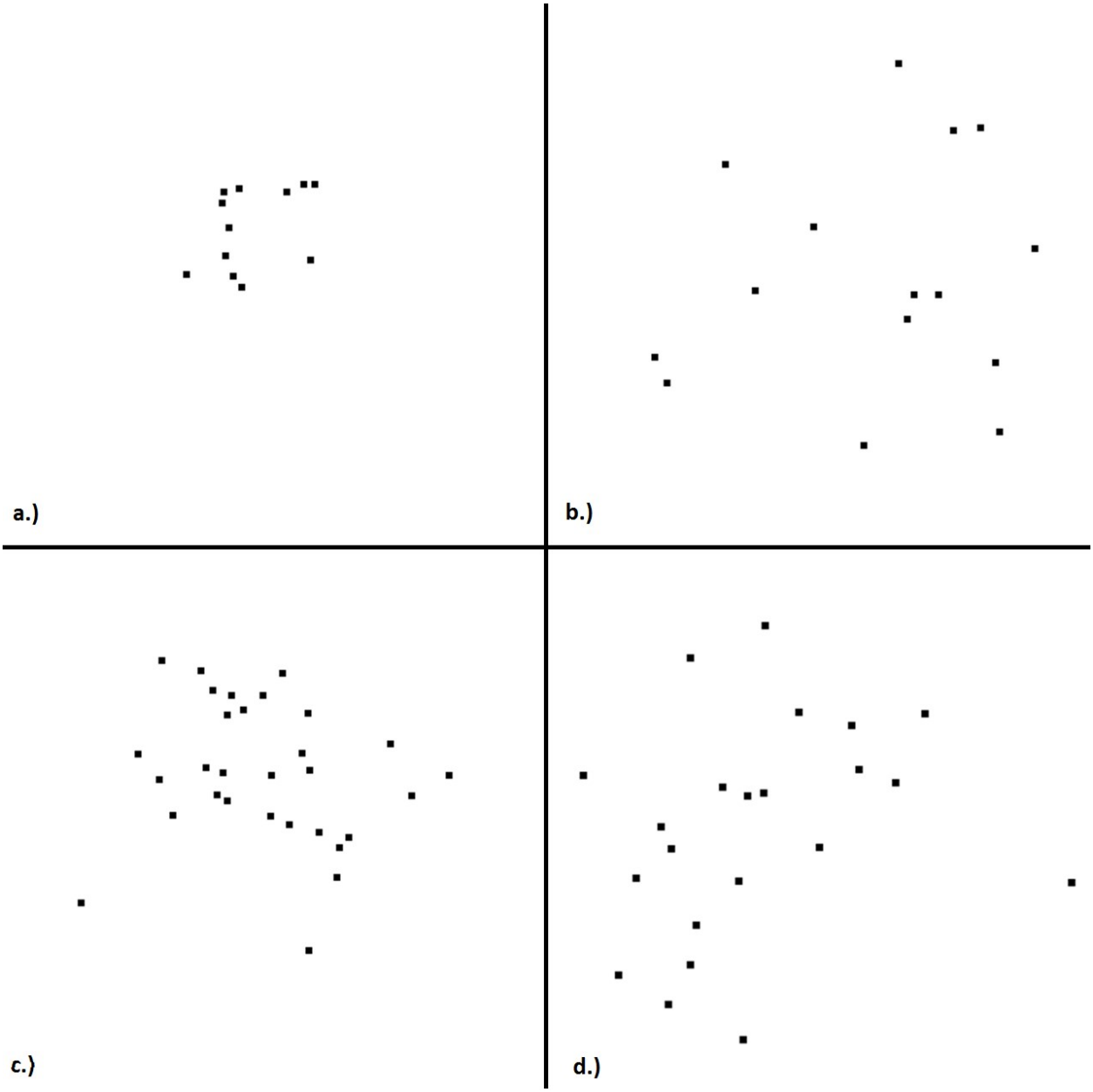

Figure 2: Examples of stimulus displays. (a) A small-set display (consisting of 12 dots) with low-dispersion (b) A small-set display (consisting of 15 dots) with high-dispersion. (c) A large-set display (consisting of 30 dots) with low-dispersion. (d) A large-set display (consisting of 22 dots) with high-dispersion.

Thus there were four task/condition pairs: centroid/small-set, centroid/large-set, numerosity/small-set and numerosity/large-set. Participants completed these task/condition pairs in a counterbalanced ABCDDCBA format, where the mapping of task/condition pairs onto A, B, C, and D varied randomly using a Latin Squares design across participants. Each participant performed 264 trials per block in the numerosity tasks, and 264 plus 8 singleton trials in the centroid tasks. The first 100 trials in each condition were discarded as training; thus, 428 trials of data were retained for a given participant in in each task/condition pair

Analysis

In order to compare performance in the Centroid and Numerosity tasks, we needed to translate participants' error in these tasks onto the same scale. For this purpose, we used a measure called "Efficiency." Let S be a stimulus that is generated using a process identical to the process used to generate the stimuli in a given condition in our experiment. For any given probability p, define the decimated stimulus $D_p(S)$ to be the stimulus that contains a random subset of the dots in which any given dot in S survives in $D_p(S)$ with probability p, independently of all the other dots in S. The Efficiency with which a participant performs in a given task is the probability p for which the responses of an ideal observer given the stimuli $D_p(S)$ deviate from the correct responses to S (the ideal observer's distance error) by an amount that matches the error of the participant.[4]

In the centroid task, the estimation of Efficiency uses a hypothetical performer (HP) whose centroid-computation errors are due exclusively to random decimation of dots from stimulus displays. It should be noted, however, that performance of actual participants in the centroid task is degraded both by centroid-computation error as well as by non-centroid-computation error. For a given participant, non-centroid-computation error is estimated in the current study from her responses on singleton trials (trials using only a single dot). Given that this source of error does not affect numerosity judgments, it is important to subtract it from the average error-per-centroid-trial achieved by the participant to estimate the centroid-computation error. We then use a Bayesian approach to derive a posterior density characterizing the dot-survival probability p of the HP given the data from the participant. Thus, if a participant has an Efficiency of 0.75, that implies the HP needs to process (on average) 75% of the dots in a given stimulus in order to match the centroid-computation error of the subject. For details, see Appendix 1.

In the numerosity task, the Efficiency calculation also uses a hypothetical performer (HP) whose performance is limited by decimation of the stimulus display. It should be noted, however, that the performance of a particular participant in the numerosity task may be degraded by non-numerosity-computation error. The task requires the participant to compute an internal statistic that gauges the numerosity of the stimulus; however, the necessity of producing a binary judgment (less than average vs. greater than average) also requires the participant to compare this internal statistic to a criterion. If the participant uses a suboptimal criterion, then her performance will be systematically degraded; however, this potential form of degradation does not afflict the centroid task. For a given dot-survival probability p, there exists a criterion C_p that optimizes performance in the numerosity task. However, if we assume that the HP uses C_p in making its judgments, then we are likely to underestimate the participant's numerosity-computation Efficiency. To avoid this, we take a conservative approach and assume only that the HP uses some unknown criterion in making its judgments. More specifically, we assume a uniform prior distribution on the criterion C used by the HP as well as on the dot-survival

[4] It should be noted that this definition deviates from the definition of "Efficiency" in Sun et al. 2015. In particular, Sun et al., 2015, defined Efficiency to be the proportion p of dots satisfying the following condition: The variance of the response of an ideal observer presented with a stimulus comprising precisely p of the original set of dots on each trial is equal to the variance of the random error compromising the subject's responses. Thus, in Sun et al., 2015, dots were not decimated independently from the stimulus display prior to the action of the ideal observer; a fixed number was decimated from the display on each trial.

probability p. We then estimate a posterior joint distribution on p and C and sum across C to derive the marginal posterior distribution characterizing p. For details, see Appendix 2.

It is worth noting that this does assume that, after training, participants fix C. This may not be true, as it is possible that the feedback given trial to trial causes participants to adjust their value of C on the fly. However, we do not believe that this is happening on a major scale that will affect the data. To explore this possibility, we analyzed the probability for a participant to answer that there are more dots than average when seeing n dots given that they (a) incorrectly said "more" the trial before, (b) incorrectly said "less" the trial before, or (c) were correct the trial before. If there was no major effect of a roving C, these three lines should be identical—if there were an effect, the lines would be significantly different from one another. We found no difference between these curves and thus conclude that variations in C are negligible.

Efficiency is a useful measure because it translates response accuracy from each task into a form that can be compared across the two tasks. As error decreases in either task, efficiency increases. This makes it an ideal measure to use for our purposes.

Results

To analyze the data, we calculated a posterior probability of a given efficiency value p given a participant's data. We then looked at the means of these distributions, and compared the centroid and numerosity task for each participant. These efficiency values for each participant and task are shown in Figure 3.

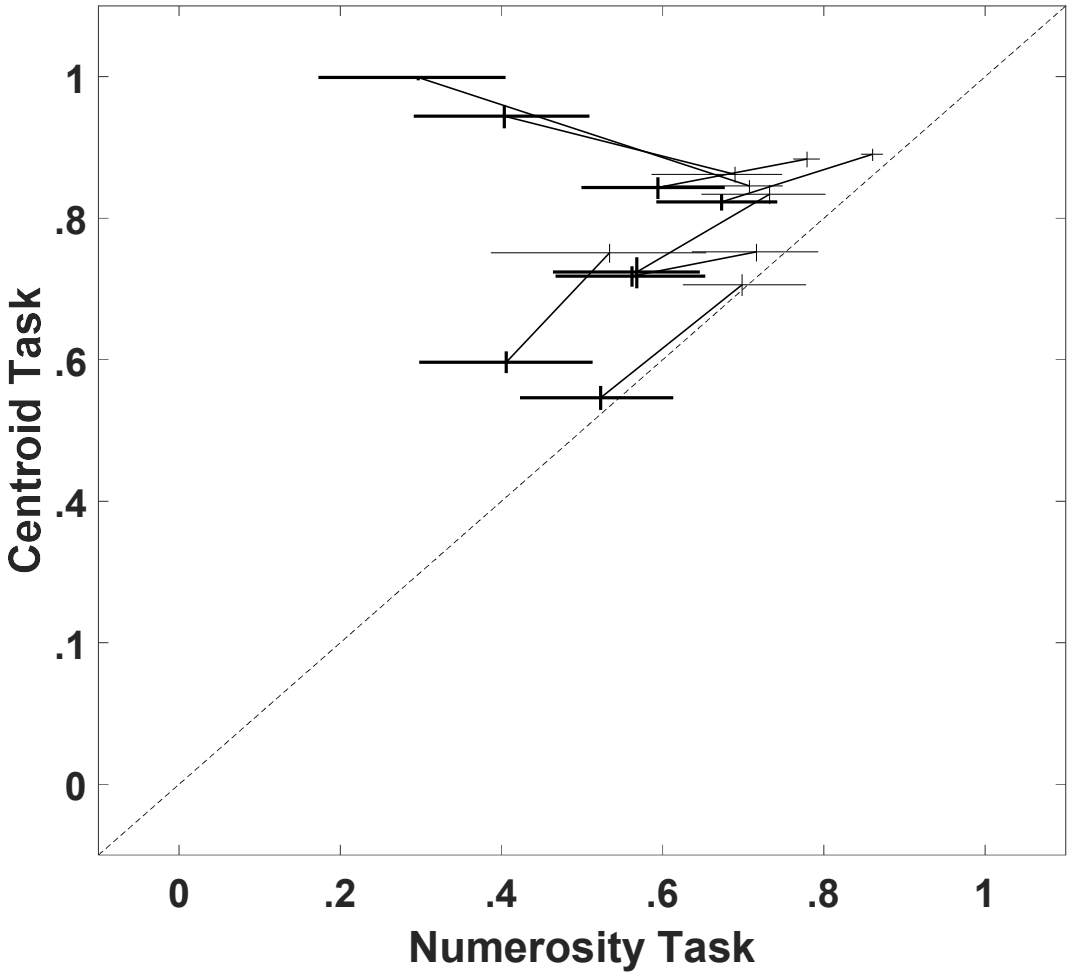

Figure 3: Mean efficiency for each participant in small- and large-numerosity stimuli for centroid and numerosity judgments. Each pair of intersecting horizontal and vertical lines represents intervals of the data from a single participant in a single stimulus condition in each of the two tasks. The diagonal line connects a participant's data from the two stimulus conditions. Thin lines with thin error bars show a participant's efficiency in the small-numerosity conditions (in both the numerosity and centroid tasks); thick lines and error bars show efficiency in the larger-numerosity conditions. Error bars are the 95% credible intervals for the distributions. The diagonal dashed line indicates where efficiency in the two tasks is equal—data above that line indicates efficiency is better in the centroid task, and data below indicates efficiency is better in the numerosity task.

One representative example of the posterior distributions on efficiency is shown in Figure 4. For 7 of the 8 participants in both numerosity conditions, the probability of a random number drawn from the centroid distribution being higher than a random number drawn from the numerosity distribution is over .85 (ranging from .88 to <.99, as determined by simulating 10000 times), and even for the eighth participant this probability is .60 in the small numerosity condition and .70 in the large numerosity condition.

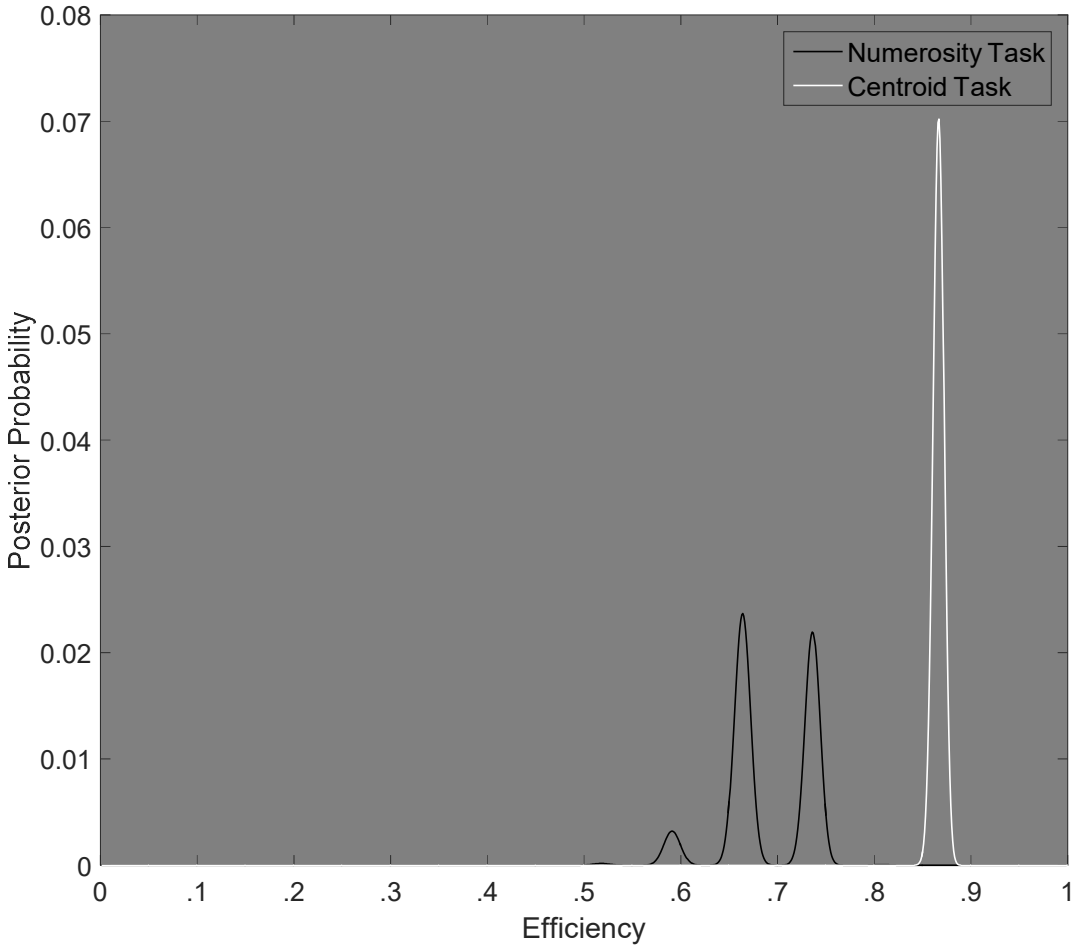

Figure 4: The probability distributions of a representative subject in the smaller numerosity condition. The plots for the other subjects and in the larger numerosity condition had curves qualitatively similar to this.

Discussion

The prediction we discussed at the outset stated that if centroids were computed by using the numerosity mechanism in computing the average of the locations of the dots in the stimulus, then the numerosity task should be strictly easier than the centroid task. This implies that the numerosity task should yield higher efficiency performance than the centroid task. However, we see the opposite result. In our data, participants all achieved higher efficiencies in the centroid task than in the numerosity task. We assume that our participants are using the intermediate numerosity mechanism (Anobile, Cicchini, & Burr, 2016) to perform the numerosity task. We conclude that this mechanism is not recruited for estimating centroids. Our results suggest moreover that either (1) the computation used to estimate centroids does not involve extracting numerosity or else (2) the numerosity estimate extracted by the centroid computation is (a) more accurate than the estimate provided by the numerosity mechanism, and (b) not available for use in the numerosity task.

There are several possible explanations why participants perform better in the centroid task than in the numerosity task. It is possible that centroids are computed by a process that uses the stimulus input to produce an output without providing access to the results of steps in between. In this case, even if the computation used within this module did produce an estimate of the numerosity of the stimulus cloud, this estimate would not be available to other processes. In particular, this estimate would not be available to the process used by the participant to produce his or her response in the numerosity task. If this were the case, then high efficiency in a centroid task would not translate to high efficiency in a numerosity task. The two tasks would be functionally separate consistent with our results.

There are also some possibilities for estimating centroids that do not require counting the dots. One rather simplistic approach assumes we have a network of neurons responding to the display of dots. Each one responds mildly to dots near the center of its receptive field, and highly to those further away (in a parabola shape). These responses are additive when multiple dots are present. If the receptive fields of such a net were wide enough to capture the entire display region, the centroid of the dot cloud would always fall exactly where the least active neuron in the net is. Of course, such a system would be far more complicated so as to not rely on a single active neuron, but a mechanism that functions in this way is possible. This would give us access to a reliable centroid calculation without the need to count the number of dots, and thus could cause results like we have shown. Computing the centroid would no longer require numerosity information.

Another possibility that doesn't use numerosity information directly could be that the centroid computation uses a low-pass filter on the stimulus, and relies only on the density of the resulting image. Areas where dots are a lot of dots are present would end up getting more weight in the computation, while much of the numerosity information is lost. This possibility can be tested in future work by doing a centroid task with blurred stimuli—if we are using only low-spatial frequency information, this should not make much difference in performance.

As of now, further work is needed to narrow down the potential solutions to the problem of how exactly the brain computes centroids. Regardless, it is clear that centroids are not computed in the simplistic way that seemed reasonable to assume. People do not simply locate their dots and average their location.

Appendix 1: Efficiency Computation in Centroid Tasks

On a given trial in the centroid task, the participant is presented with a stimulus comprising some number of dots. Let x_i and y_i be the x- and y-coordinates of dot i; then the target location ($Targ_x$, $Targ_y$) is defined by

$$Targ_X = \frac{1}{N}\sum_{i=1}^{n} x_i \text{ and } \quad Targ_Y = \frac{1}{N}\sum_{i=1}^{n} y_i \qquad (1)$$

Invariably the response of the participant deviates from this target. Some of this deviation is due to motor error. We estimate this component of the error from the singleton trials included in our design. The rest of the error (which may include both random noise as well as unmodelled, systematic error) we gauge using the Efficiency measure described in this section. This measure assumes that all non-motor error in the participant's responses is due to random, independent decimation of dots from the stimulus cloud. Specifically we define Efficiency in terms of a sub-ideal observer whose response on each trial is obtained by computing the centroid of a dot cloud derived by randomly and independently removing a given dot from the stimulus cloud with probability 1-p. The Efficiency of a given participant in the centroid task is the dot-survival probability p required for the error of sub-ideal observer to match the non-motor error of the participant.

Consider a cloud with N dots. Given a dot-survival probability p, the probability that k dots remain in the decimated cloud is given by

$$P_k(p) = \binom{N}{k} p^k (1-p)^{(N-k)} \qquad (2)$$

The difference between the x-coordinate of the centroid of a cloud in which k dots survive vs. the original N-dot cloud is given by

$$Q_k \;=\; \frac{1}{k}\sum_{j=1}^{k} X_j - \frac{1}{N}\sum_{j=1}^{N} X_j \;=\; \sum_{j=1}^{k}\left(\frac{N-k}{Nk}\right) X_j - \frac{1}{N}\sum_{j=k+1}^{N} X_j \qquad (3)$$

where X_j is the x-location of the jth dot. A similar computation can be done for the y-locations Y_j.

Suppose the X_j's and Y_j's are independent and normally distributed, each with mean 0 and standard deviation σ. In this case, some algebra applied to Eq. (3) shows that Q_k is distributed normally with mean 0 and standard deviation $\left(\frac{N-k}{Nk}\right)\sigma^2$.

We assume that the response of the sub-ideal observer is degraded by motor error identically distributed to that of the participant. This motor error is assumed to be normally distributed in each of the x and y directions with mean 0 and standard deviation σ_{one}, estimated from singleton trials, i.e., trials in which the participant had to click the

centroid of only a single dot whose location was distributed identically to the target location on a non-singleton trial. Thus, on a trial in which k dots survive in the decimated stimulus cloud, the horizontal component of the sub-ideal observer's response will be

$$D_X = Q_k + S \qquad (4)$$

where S is normal with mean 0 and standard deviation σ_{one}. Since D_X is normal, its density function is given by

$$f_k(\alpha, \sigma) \frac{1}{\sqrt{2\pi} s_{k,\sigma}} exp\left[\frac{-\alpha^2}{2s_{k,\sigma}^2}\right] \qquad (5)$$

with standard deviation $s_{k,\sigma} = \left(\frac{N-k}{Nk}\sigma^2 + \sigma_{one}^2\right)^{\frac{1}{2}}$.

Consider a single experimental trial t on which the number of dots in the stimulus was N_t, the standard deviation of the x- and y-coordinates of the dots in the cloud was σ_t , $Targ_{X,t}$ and $Targ_{Y,t}$ are given by Eq. (1), and the x- and y-coordinates of the participant's response location are $R_{X,t}$ and $R_{Y,t}$. Then, for

$$D_{X,t} = R_{X,t}\text{-}Targ_{X,t}, \quad D_{Y,t} = R_{Y,t}\text{-}Targ_{Y,t}, \qquad (6)$$

the likelihood of dot-survival probability p is given by

$$\lambda_t(p) = \sum_{k=1}^{N_t} f_k(D_{x,t}, \sigma_t) f_k(D_{y,t}, \sigma_t) P_k(p). \qquad (7)$$

Note that the sum in Eq. (7) does not include a term for $k = 0$. This term is omitted because the centroid is not defined in the case in which all dots have been decimated from the stimulus cloud. We assume (as confirmed by our results) that $P_k(p)$ will be very small for $k = 0$. Applying Eq. (7) over all N_{trials} trials yields

$$\Lambda(p) = \prod_{t=1}^{N_{trials}} \lambda_t(p) \qquad (8)$$

We assume a uniform prior density on p. Under this assumption, the posterior density of p is

$$f_{posterior}(p) = \frac{\Lambda(p)}{\int_0^1 \Lambda(q)dq} \qquad (9)$$

We use numerical methods to approximate $f_{posterior}(p)$.

Appendix 2: Efficiency Computation in Numerosity Tasks

Efficiency in the numerosity tasks has the same definition as in the centroid tasks—it is the proportion of dots a subideal observer (who only makes mistakes by missing dots) would need to see in order to match the errors of the participant. However, the calculation needs to be done in a different way. We assume that each dot on the screen will survive in the decimated cloud with probability p, independent of other dots. Once the subideal observer counts the dots that survive, it will use an internal criterion number of dots (C) to decide if the display came from one of the three display-types comprising fewer-than-average dots or one of the three display-types comprising more-than-average dots.

It seems likely that our participants may themselves be using suboptimal criteria in producing their responses. Therefore, if we assume that for any given dot-survival probability p, the ideal observer uses the criterion that would be optimal for that p, then we may end up underestimating the participant's efficiency for the following reason: the dot-survival probability p will need to be lowered to compensate for the systematic error injected into the participant's responses by using a suboptimal criterion (leading to lower estimated efficiency). The hypothesis under investigation proposes that efficiency in the numerosity task is lower than efficiency in the centroid task; for the argument supporting this claim to be convincing, it is important that we use an estimate of efficiency in the numerosity task that errs on the side of *higher* (not lower) efficiency. For this reason, instead of allowing our ideal observer to use the optimal criterion C for a given p, we assume a uniform prior distribution on C for any given p. Then we compute the posterior joint distribution of p and C and integrate across C to derive our estimate of the posterior density characterizing p.

Since C is unknown (and not assumed to be optimal given p), we examine the likelihood of a participant's efficiency being p across a range of possible values of C. Under the assumption that the sub-ideal observer always bases its judgment on a decimated cloud comprising an integer number of surviving dots, responses will be invariant with respect to changes in C between two successive integers (e.g., 7 and 8). Accordingly, we restrict our consideration to the following values of C: 0.5,1.5,…,15.5 in the 11-to-16-dot condition, and 0.5,1.5,…,31.5 in the 22-to-32-dot condition.

For any given stimulus dot-cloud size N, criterion C and dot-survival probability p we let P_{Fewer} be the probability that the sub-ideal observer with that C and p says "*Fewer*" in response to a stimulus with N dots. This is given by

$$P_{fewer}(p, N, C) = \sum_{h=0}^{floor(C)} \binom{N}{h} p^h (1-p)^{(N-h)} \qquad (10)$$

And

$$P_{more}(p, N, C) = 1 - P_{fewer}(p, N, C) \qquad (11)$$

Using these probabilities, as well as the number of trials with k dots where the participant said "fewer" (F_k) or "more" (M_k), we can determine the log likelihood of any combination of p and C, as follows (Eq. (12) is for the 11-to-16-dot condition; the likelihood function for the 22-to-32-dot condition is exactly analogous.):

$$\Lambda(p, C) = \prod_{N=11}^{16} P_{fewer}(p, N, C)^{F_N} P_{more}(p, N, C)^{M_N} \qquad (12)$$

The set of values C upon which $\Lambda(p, C)$ is defined is finite (C = 0.5, 1.5,…, 15.5). We discretize p as well, taking p = 0.001, 0.002,…, 1.0. As in the case of the centroid task, we assume a uniform joint prior distribution on p and C. Then the joint posterior distribution is

$$\hat{f}_{posterior}(p, C) = \frac{\Lambda(p, C)}{\sum_{c=0.5}^{15.5} \sum_{p=0.001}^{1.0} \Lambda(q, c)} \qquad (13)$$

Since we are only interested in the marginal distribution of p, we can then sum over C to derive

$$\hat{f}_{posterior}(p) = \sum_{C=0.5}^{15.5} f_{posterior}(p, C) \qquad (14)$$

A point estimate of efficiency can be obtained by computing the mean of $\hat{f}_{posterior}(p)$.
For concreteness, we have written all of our equations focusing on the 11-to-16-dot condition. All of the equations are similar for the 22-to-32-dot condition. However, in this case N ranges from 22 to 32 in units of 2 in Eq. (12), and C ranges across 0.5, 1.5,…, 31.5.

Chapter 3:
Examining Selective Attention to Angles Using the Centroid Task

Abstract

We have previously determined that people are unable to use absolute orientation as a selective attention cue in the centroid task. However, if we take two oriented bars and combine them to form a chevron, could people then use angle information as a cue? To investigate that, we ran a centroid task with angles as the feature of interest, and found people are able to attend to just the target angles. We further found that discriminibility between angles changes with the ratio of the angles, not the absolute distance between them. Overall, it seems that angle is a useful cue for selective attention, despite the constituent oriented line segments not being useful.

Introduction

As discussed in previous chapters, the centroid method remains a powerful tool in studying visual selective attention. It allows us to quickly and accurately determine what type of features people can attend or cannot attend to, and answer questions perceptual discriminability.

One potentially salient feature that has not yet been studied extensively is internal angles. It is plausible that people can discriminate two distinct angles and attend to only one in a very quick viewing. A pair of line segments that join with a 40 degree internal angle has a very different shape from a pair of line segments that join at a 120 degree internal angle, after all. However, as shown in chapter 1, people cannot attend selectively to oriented bars. Angles, which are essentially just two oriented bars that meet at a vertex, may also be hard to filter. This view of angles—being simply two oriented bars joined together—is supported by neural work. Ito and Goda (2011) inverstigated single neuron responses to cevron shapes in macaques, and found that these responses can be modeled very well as a simple summation of responses to the line segments that make up the chevron. If this is an indication that people are only sensitive to the oriented bars and not the angle itself, we wold predict that using chevrons in a selective attention task should be impossible. It is not clear whether or not people will be able to attend selectively to one angle over another, and yet this has not been studied much at all. It is known that people have a good memory for recently-seen angles (Hubbard & Blessum, 2001), that search is efficient for a target chevron among straight-line distracters (Fahle, 1991), and that people can use the angular relationship between target oriented bars and distracter oriented bars in a search task (Wolfe & Hill, 1992), but these results do not say whether or not people can use the angle itself as a selected feature. The centroid task is well equipped to figure out if this is a task people are capable of or not.

Further, if angles can be used as a feature in a selective attention task, it is important to know how discriminable different sets of angles are. How far must angles be from one another to be perceptually distinct enough to make the task possible? Is this distance related to the absolute distance between the angles, the ratio between them, or some other measure? Further, are the any "special" angles that are distinct in spite of other rules (such as right angles)? All of these questions are worth exploring.

It is quite possible that the discriminbility of angles relates to their absolute difference. Perhaps angles that are 30 degrees apart from one another are equally discriminable, regardless of what each specific angle is (40 vs. 70 degrees is just as distinct as 120 vs. 150 degrees). Alternatively, the discriminibility could follow Weber's law and relate to the ratio between the angles (30 vs. 60 degree angles would be as discriminable as 60 vs. 120 degrees).

Furthermore, it's quite possible that angles that cross the boundary between acute and obtuse are easier to discriminate than similarly different angles that stay on one end of the boundary or the other. Crossing a categorical boundary can help with classification, which might make discrimination easier. This is not expected, as other work using angled stimuli have not found anything special involving right angles (Snippe & Koenderink, 1994), however it is important to make sure. Another possible special case occurs when angles get large. Eventually, as an angle approaches 180 degrees, it begins to appear like a line segment. These high-degree angles may

also appear more distinct from other angles than their difference (either absolute or ratio) would suggest.

This paper aims to answer all of these questions about a person's ability to use angles in a selective attention task.

Experiment 1: Methods

Participants in our study were 7 students at the University of California, Irvine (3 males, 4 females). All had given informed consent to be in our study.

Each participant ran 3 different tasks in a counterbalanced order. In one task, participants were meant to attend to only acute chevrons. In another, obtuse chevrons were the target, and in the third, all chevrons were target chevrons. In all tasks, a trial consisted of participants seeing a brief, 200ms display consisting of 8 chevrons, randomly rotated. 4 of these were acute, and 4 were obtuse. Each acute chevron was independently and randomly one of 22.5, 45, or 67.5 degrees. Each obtuse chevron was independently and randomly one of 112.5, 135, or 157.5 degrees (Figure 1). Chevrons were placed using the same methods as in chapter 1 of this document. A single trial began with a blank screen for 1 second, followed by the stimulus for 200 ms. After the image disappeared, a mask was displayed, followed by a blank gray screen. The participants task was to identify by mouse click the centroid of the target chevrons. Feedback would be given by showing the participant a screen with the stimulus display, a bullseye to indicate the true centroid, and a crosshair to indicate where they clicked. When the participant pressed any key, the next trial would begin (Figure 2). The true centroid shown on the feedback screen was computed by averaging the location of the vertex of each chevron. A pilot study was done without feedback to try to determine what point on each chevron people naturally used to compute the average location of all chevrons. The result was that it did not make any difference—there is far more variability in random noise than in using different points along the chevron.

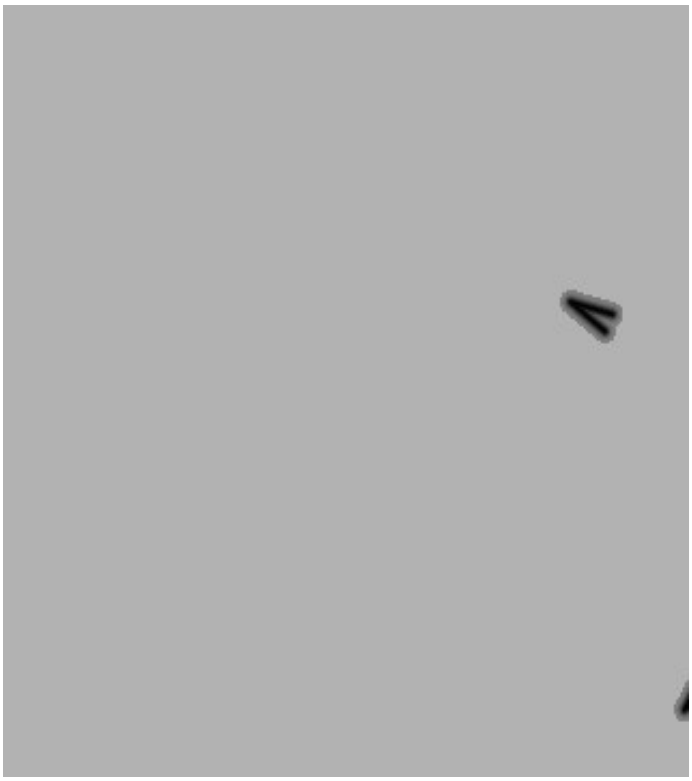

Figure 1: A sample stimulus.

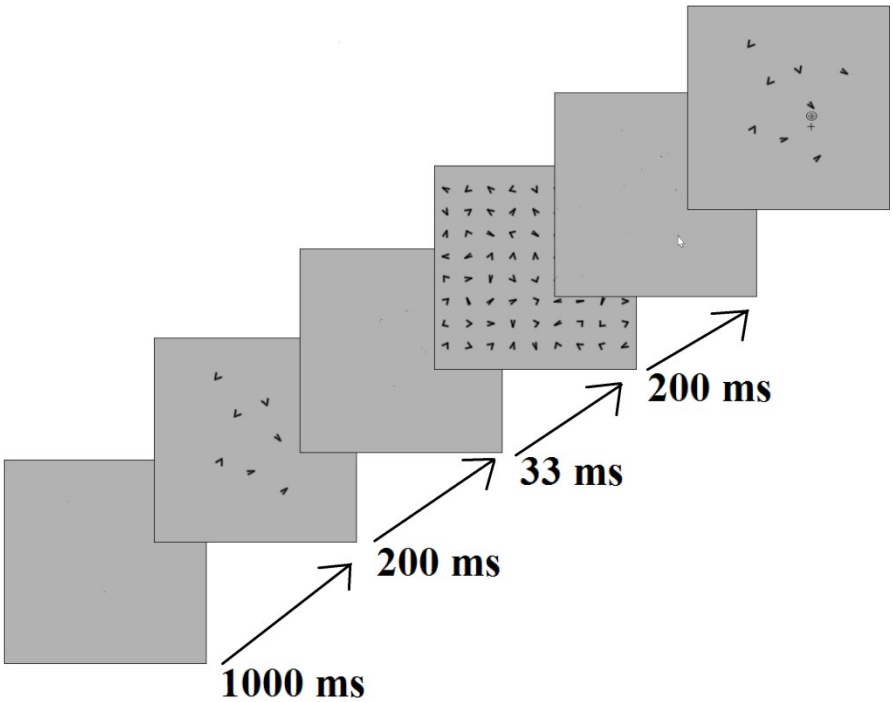

Figure 2: The time course of a trial. Feedback here is given for the “attend to acute chevrons” condition.

Experiment 1: Results

To analyze our results, we use the methods detailed in Chapter 1 to find selectivity, efficiency, and filter weights in every condition where attention was deployed. The task where every chevron was a target was analyzed to show that participants were able to do the task, but is otherwise uninteresting. Results are summarized in table 1. Selectivity is reasonable—participants could give more weight to the targets than the distracters. However, it is not as high as we have seen in centroid tasks using other feature dimensions.

Figure 3 shows the weights our 7 participants achieved for every possible target chevron in each task. Dashed lines indicate perfect performance. Participants were clearly giving more weight to target chevrons than distracter chevrons. Chevrons near the boundary between target and distractor were harder to discriminate, however, and that is reflected in these results.

Table 1: Average efficiency and selectivity per task

	Efficiency	Selectivity
Attend to Acute Chevrons	0.7501	2.9939
Attend to Obtuse Chevrons	0.8102	2.9750

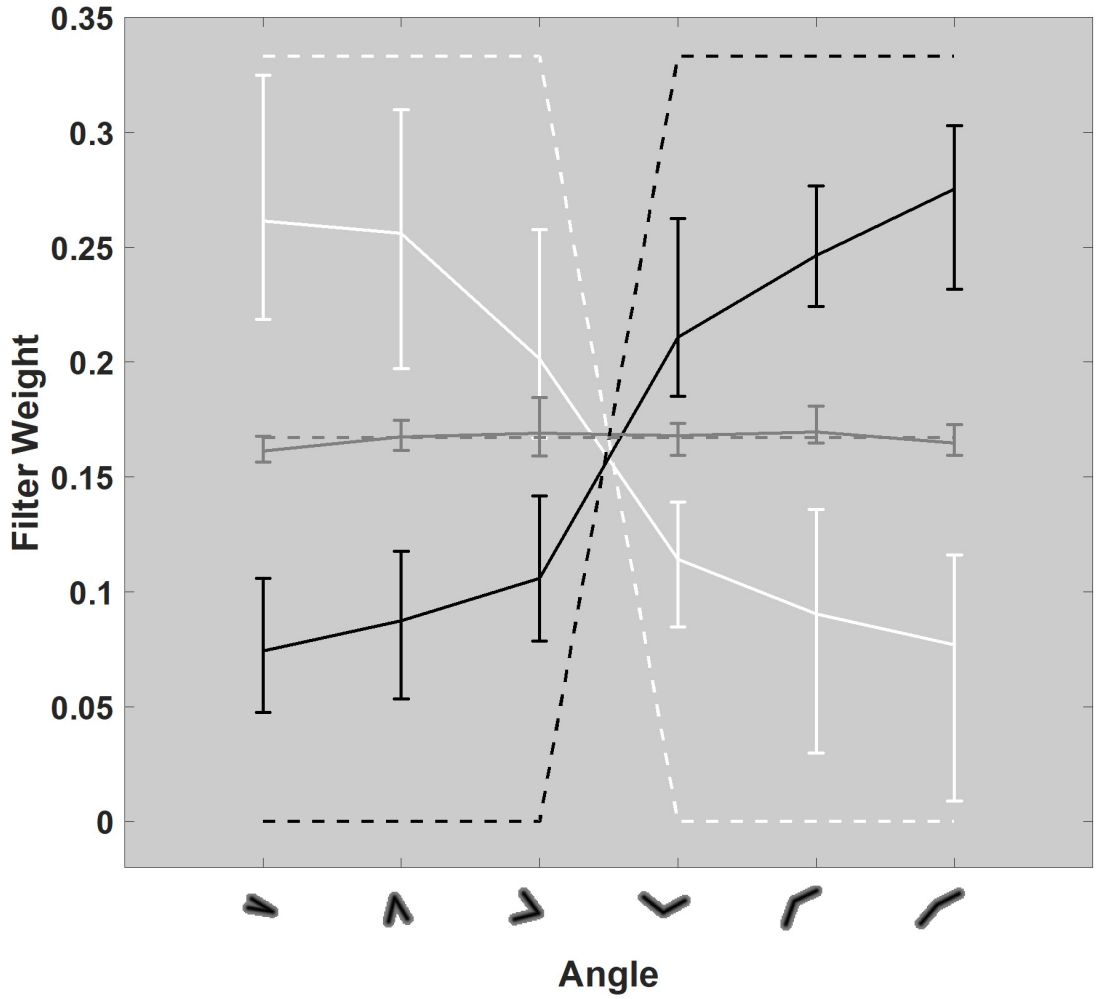

Figure 3: Filter weights given to target and distracter chevrons, averaged across participants. The white line represents the attend to acute task, the black line represents the attend to obtuse task, and the dark gray line represents the attend to all task. Dashed lines show ideal performance.

Experiment 2: Methods

After establishing that participants could, to some degree, use chevrons as targets, we decided to explore the discriminability between chevrons. To that end, each participant came for 5 experimental sessions. In all sessions, a trial consisted of participants seeing a brief 200ms display consisting of 4 each of two different chevrons, randomly rotated. Chevrons were placed using the same methods as in chapter 1 of this document. The timing of a trial, as well as the participants task, was otherwise identical to that in experiment 1.

In the "30/60" session, the two types of chevrons were 30 and 60 degrees. Participants were instructed to go through 14 blocks of 100 trials each. There were three conditions for them to do. In one condition, every chevron was a target, and participants had to indicate the centroid of every chevron on screen. In another, only the 30 degree chevrons were targets, and the 60

degree chevrons were to be ignored. The third was the opposite of the second. Participants ran 500 trials in each of the attention condition and 400 trials in the condition where every chevron was a target, in a counterbalanced, ABCCBA order. The first 100 trials in each of the attention conditions were discarded as training.

The "75/120" session ran similarly, except chevrons were 75 and 105 degrees. These chevrons were chosen to be the same absolute distance apart as those in session 1 (30 degrees), but now additionally cross the categorical boundary from "acute" to "obtuse". Similarly, The "120/150" session was designed the same way but with chevrons of 120 and 150 degrees.

In the "60/120" session, we repeated the design but using chevrons that are 60 and 120 degrees. These now maintain the Weber ratio of the chevrons in the "30/60" session, but cross the categorical boundary from "acute" to "obtuse". In this session, we did not have participants run the condition where all chevrons were targets, so there are only 10 blocks of 100 trials each.

Finally, in the "140/170" session, we repeated the design but using 140 and 170 degree chevrons. These chevrons were chosen to determine if there was a special relationship with chevrons that were very close to being a straight line (Figure 4).

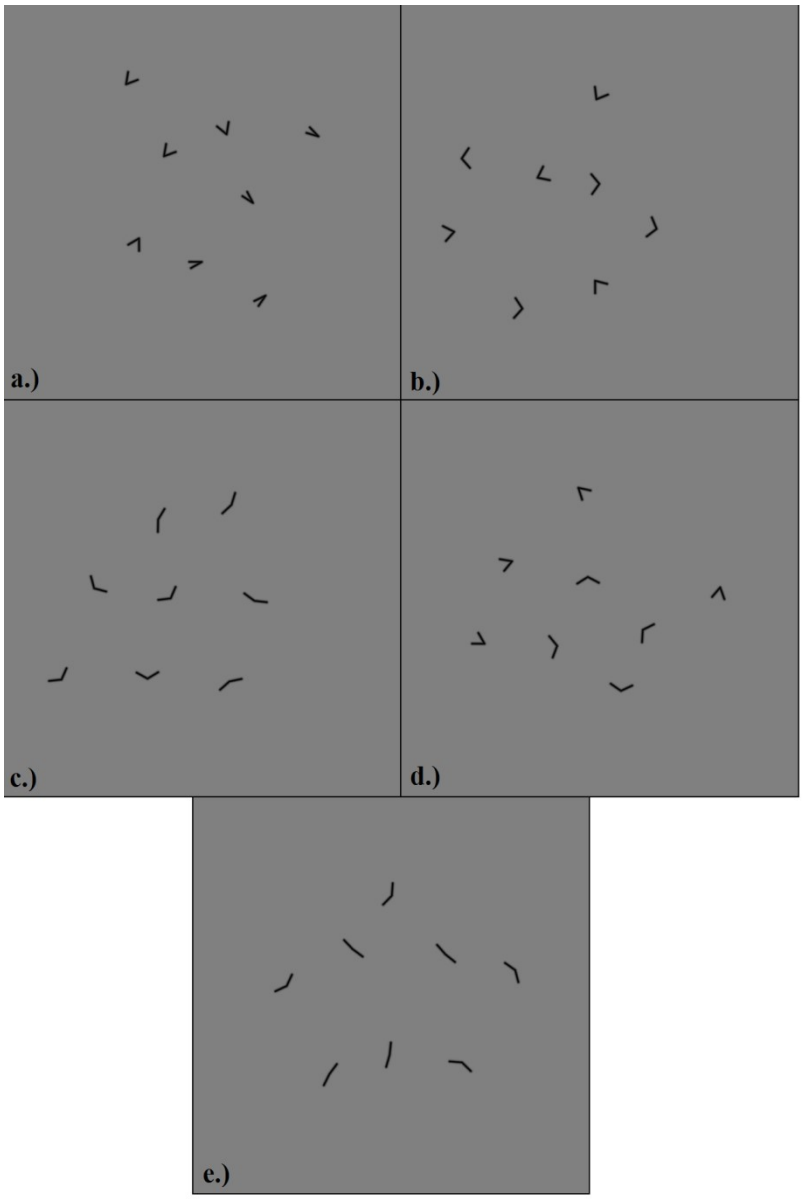

Figure 4: Sample stimuli. (a) shows what the stimulus cloud in the "30/60" session looked like. (b) depicts the "75/105" session, (c) the "120/150" session, (d) the "60/120" session, and (e) the "140/170" session.

Results: Experiment 2

To analyze our results, we use the methods detailed in Chapter 1 to find selectivity, efficiency, and filter weights in every condition where attention was deployed. Trials where every chevron was a target were analyzed to make sure the participants were able to do the task, but are otherwise uninteresting. Results are summarized in Tables 2 and 3.

In every session, efficiency was reasonably high, so participants were able to process a sizable number of the chevrons. Selectivity, however, tells different story. In the "30/60" session, performance was high, with a mean selectivity of 2.2888 across participants. The "75/105" session was significantly lower (1.4052, p=.0308). The "120/150" session also had significantly

lower selectivities than the “30/60” session (1.3341 p=.0176), but is not significantly different from the “75/105” session (p=.6994).

The “60/120” and “140/170” sessions (Average selectivities of 2.5159 and 2.5221 respectively) were not significantly different from the “30/60” session (p=.6204 and .6130, respectively). Attention filters for all 5 conditions, averaged across all participants, are plotted in figure 5. These are made such that the ideal filter gives 1.0 weight to the target chevrons and 0 weight to the distractor chevrons.

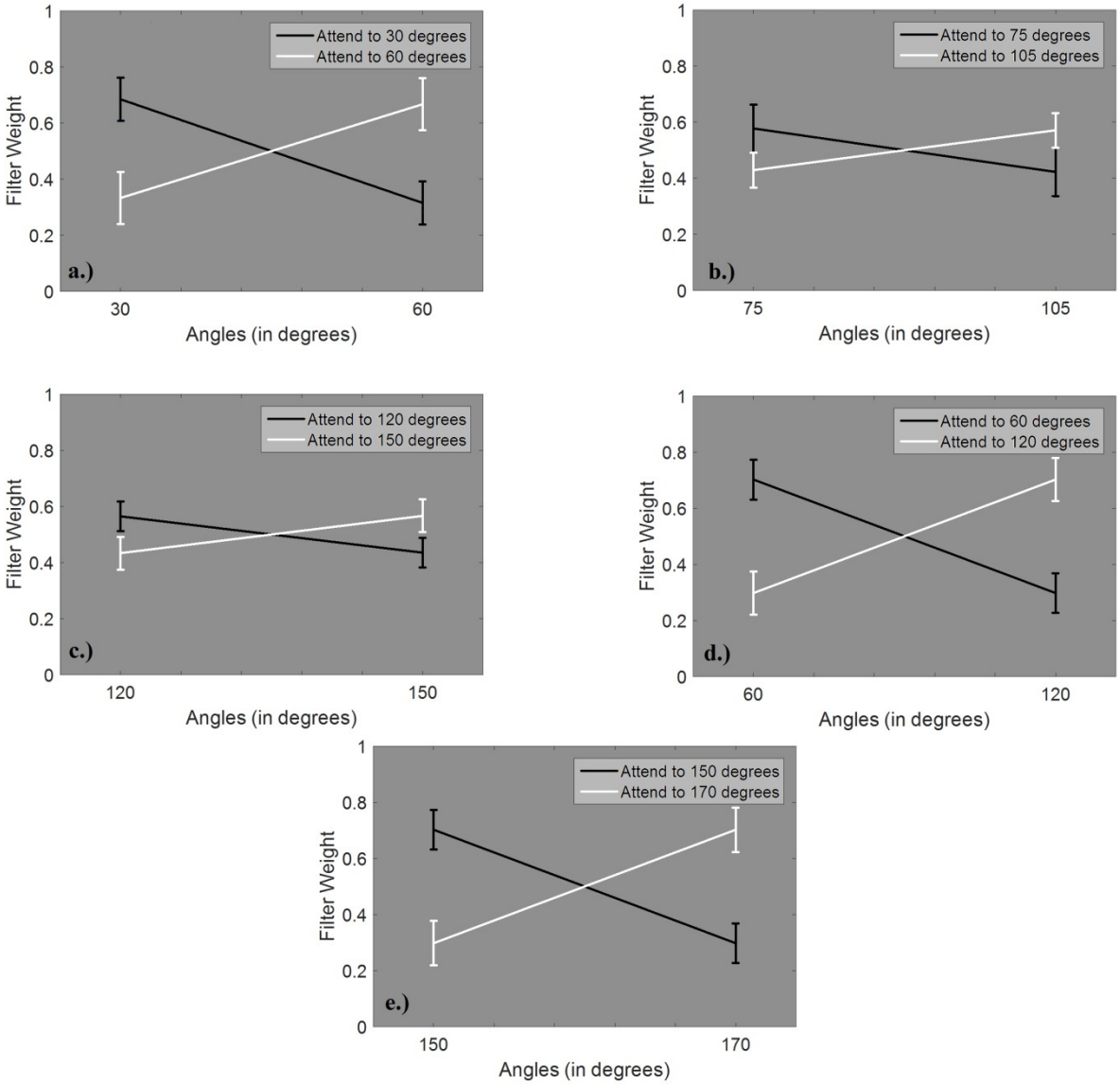

Figure 5: Filter weights given to target and distractor chevrons, averaged across participants. In all cases the black line represents the attend to smaller chevron task and the white line represents the attend to larger chevron task. (a-e) are the weights for sessions 1-5, respectively.

Table 2: Average selectivity per session and task

	”30/60”	”75/105”	”120/150”	”60/120”	”140/170”
Attend to Bigger Chevron	2.2030	1.3658	1.3461	2.4992	2.5129
Attend to Smaller Chevron	2.3746	1.4445	1.3221	2.5325	2.5313

Table 3: Average efficiency per session and task

	”30/60”	”75/105”	”120/150”	”60/120”	”140/170”
Attend to Bigger Chevron	.8178	.7436	.8065	.8084	.8069
Attend to Smaller Chevron	.7859	.7682	.8186	.7996	.8024

Discussion

The main goal of this chapter is to determine if angles were usable as a feature in a centroid task. Our results show that this does indeed seem to be the case. Though their results are not as impressive as results in other centroid tasks, participants are giving significantly more weight to target chevrons than distracter chevrons.

A secondary goal was to identify how performances changes with different angles. We discussed three competing ideas. It could have been only the absolute difference between the target and the distracter angles that made a difference. However, our current results seem to contradict this. The “30/60”, “75/105”, and “120/150” sessions all had the same absolute difference between chevrons, but performance was significantly better in the “30/60” session than the other two. So, this theory can be dismissed.

Another possibility is that there is something special about the categorical boundary between acute and obtuse. Again, this seems not to agree with our results. If it were the case, we would expect performance in the “75/105” session to be different from and better than the “30/60” and “120/150” sessions. However, the “75/105” and “120/150” sessions are very similar, and significantly worse than the “30/60” session.

The final possibility is that performance is proportional to the ratio of angles. So, in the “30/60” session where the larger chevron (60 degrees) is double the smaller chevron (30 degrees), we’d expect performance to be better than in the “75/105” session where the large chevron (105 degrees) was only 1.4 times the smaller chevron (75 degrees). Likewise, we would not expect a huge difference between the “75/105” session and the “120/150” session, where the larger chevron (150 degrees) is only 1.25 times the smaller chevron (120 degrees). This is well supported by our data. The “60/120” session furthers this, since performance in that session is not any different from performance in the “30/60” session, and the ratios are the same.

However, there is a major problem with this explanation when we look at the “140/170” session. The larger chevron here is only about 1.21 times the smaller chevron, yet performance is as good as in the “30/60” and “60/120” sessions. This, however, is to be expected. 170 degree angles are perceptually similar to straight lines (which are, in essence, just 180 degree angles), and as such

stand out as a different type of item. Instead of looking for larger or smaller chevrons, participants here could be simply looking for chevrons or line segments—a much easier comparison. So, it stands to reason that angles that approach 180 degrees are special. Efficiencies being high in every session is also interesting. This implies that the participant is able to process many of the chevrons in every case—the hard tasks are only hard because of discrimination. The "75/105" and "120/150" sessions are much harder for participants than the other sessions, yet they are able to see about as many dots.

Overall, it does seem like angles are a useful, though weak, feature for attention. People can quickly discriminate and attend to items based on their angle, though when the ratio between target and distracter angles is small, distracters are hard to completely ignore.

General Conclusion

The centroid method is an excellent tool for studying feature based selective attention, and many different conclusion can be drawn with the use of this task. Because, in its most basic form, it is a pure attention task, we can draw conclusions that would not be possible from a search task.

In chapter 1, we were able to demonstrate the centroid tasks ability to answer the question "Can people use a certain filter dimension in selective attention?" To do this, we explored using the centroid task to attend to either horizontal or vertical bars, ignoring the other orientation. We found that this task was very hard for people to do—they could not ignore the distracters. Adding more distracters (without changing the number of targets), made the task a lot easier. Adding these distracters serves to increase the local orientation contrast around the targets, so the improved observer performance implies that orientation contrast, rather than absolute orientation, is the feature people try to use in doing a feature based selective attention task with orientated bars.

Following that, in chapter 2 we looked at a way of conceptualizing how the centroid is computed in the brain. We considered a theory that the centroid was computed by finding the individual items, counting them, and averaging their locations. This theory resulted in the prediction that a numerosity estimation task should be strictly easier than a centroid task. Testing showed the opposite conclusion, which indicates that the brain has some other way of estimating centroids that does not require counting all the dots.

Finally, in chapter 3, we repeated the question of chapter 1 with a new feature dimension, but importantly added the question of "How discriminable are two item types that differ on some feature dimension?" We explored this with angles as our feature dimension of choice. Because an angle simply consists of two oriented bars put together, it was not completely clear that we would be able to ignore distracters. However, we found that people were able to do this task provided that the target and distracter angles were sufficiently discrimible. Further, we showed that discriminibility is based on the ratio of the target angle to the distracter angle, rather than their absolute distance apart, and that some special angles were more discriminable that would be otherwise predicted.

Taken together, these results speak to the utility of the centroid task in exploring feature based selective attention. Centroids provide a robust and efficient pure attention task to answer many questions about what feature dimensions we can and cannot filter on, and how discriminable these features are.

References

Allik, J., & Tuulmets, T. (1991). Occupancy model of perceived numerosity. *Perception & Psychophysics, 49*, 303–314

Anobile, G., Cicchini, G. M., & Burr, D. C. (2014). Separate mechanisms for perception of numerosity and density. *Psychological Science, 25*, 265–270.

Anobile, G., Cicchini, G. M., & Burr, D. C. (2016). Number as a primary perceptual attribute: A review. *Perception, 45*(1-2), 5-31.

Anobile, G., Turi, M., Cicchini, G. M., & Burr, D. C. (2015). Mechanisms for perception of numerosity or texture-density are governed by crowding-like effects. *Journal of Vision, 15*(5), 4-4.

Beck, J. & Ambler, B. (1973) The effects of concentrated and distributed attention on peripheral acuity, *Perception and Psychophysics, v. 14*, 225–230.

Bergen, J. R., & Julesz, B. (1983). Parallel versus serial processing in rapid pattern discrimination. *Nature, 303*(5919), 696-698.

Burr, D. C., Turi, M., & Anobile, G. (2010). Subitizing but not estimation of numerosity requires attentional resources. *Journal of Vision, 10*(6), 20-20.

Cowan, N. (2001). Metatheory of storage capacity limits. *Behavioral and brain sciences, 24*(01), 154-176.

Dakin, S. C., Tibber, M. S., Greenwood, J. A., Kingdom, F. A. A., & Morgan, M. J. (2011). A common visual metric for approximate number and density. *Proceedings of the National Academy of Sciences of the United States of America, 108*, 19552–19557.

Drew, S. A., Chubb, C. F., & Sperling, G. (2010). Precise attention filters for Weber contrast derived from centroid estimations. *Journal of Vision, 10*(20), 1-16.

Fahle, M. (1991). Parallel perception of vernier offsets, curvature, and chevrons in humans, *Vision Research, 31*(12), 2149-2184

Foster, D.H. & Ward, P.A. (1991a). Asymmetries in oriented-line detection indicate two orthogonal filters in early vision, *Proc. Roy. Soc. Lond. B*, 243, 75-81.

Foster, D.H. & Ward, P.A. (1991b). Horizontal--vertical filters in early vision predict anomalous line-orientation identification frequencies*, Proc. Roy. Soc. Lond. B*, 243, 83-86.

Foster, D.H. & Westland, S. (1995). Orientation contrast vs orientation in line-target detection. *Vision Research, 35*, 733-738.

Foster, D. H. and Westland, S. (1998). Multiple groups of orientation-selective visual mechanisms underlying rapid orientated-line detection, *Proc. Roy. Soc. Lond. B 265*, 1605– 1613.

Friedman-Hill, S. R., & Wolfe, J. M. (1995). Second-order parallel processing: Visual search for the odd item in a subset. *J. Experimental Psychology: Human Perception and Performance, 21*(3), 531-551.

Halberda J. Sires S. F. Feigenson L. (2006). Multiple spatially overlapping sets can be enumerated in parallel. *Psychological Science, 17* (7), 572–576.

Huang, L. (2015). Color is processed less efficiently than orientation in change detection but more efficiently in visual search. *Psychological Science, 26*(5) 646-652.

Hubbard, T. L., & Blessum, J. A. (2001). A structural dynamic of form: Displacements in memory for the size of an angle. *Visual Cognition, 8*(6), 725-749.

Hubel, D.H. & Wiesel, T.N. (1962). Receptive fields, binocular interaction, and functional architecture in the cat's visual cortex. *J. Physiol., 160*(1), 106-156.

Hubel, D.H. & Wiesel, T.N. (1959). Receptive fields of single neurons in the cat's striate cortex. *J. Physiol., 148*, 574-591.

Inverso, M., Sun, P., Chubb, C., Wright, C. E., &Sperling, G. (2016). Evidence against global attention filters selective for absolute bar-orientation in human vision. *Attention, Perception, & Psychophysics, 78*(1), 293-308.

Ito, M., & Goda, N. (2011). Mechanisms underlying the representation of angles embedded within contour stimuli in area V2 of macaque monkeys. European *Journal of Neuroscience, 33*(1), 130-142.

Kastner, S., Nothdurft, H.C. & Pigarev, I.N. (1999). Neuronal responses to orientation and motion contrast in cat striate cortex , *Visual Neuroscience, 16*, 587–600.

Kaufman, E. L., & Lord, M. W. (1949). The discrimination of visual number. *American Journal of Psychology, 62*, 498–525.

Koene, A.R. & Li, Z. (2007). Feature-specific interactions in salience from combined feature contrasts: Evidence for a bottom–up saliency map in V1, *Journal of Vision 7*(7):6, 1–14.

Krueger, L. E. (1982). Single judgments of numerosity. *Perception& Psychophysics, 31*(2),175- 182.

Luck, S.J. & Vogel, E.K. (1997). The capacity of visual working memory for features and conjunctions, *Nature, 390*(6657). 279-281.

Mikellidou, K., Cicchini, G.M., Thompson, P.G., & Burr, D.C. (2015). The oblique effect is both allocentric and egocentric. *Journal of vision, 15*(8), 24-24.

Moreland, J. C., & Boynton, G. M. (2016). A neurophysiological explanation for biases in visual localization. *Attention, Perception, & Psychophysics*, 1-10.

Nothdurft, H.C. (1991). Texture segmentation and pop-out from orientation contrast. *Vision Research, 31*,1073-1078.

Nothdurft, H. C. (1992). Feature analysis and the role of similarity in preattentive vision. *Perception & Psychophysics, 52*, 355-375.

Nothdurft, H.C. (1993). The role of features in preattentive vision: Comparison of orientation, motion and color cues. *Vision Res. 33*, 1937-1958.

Nothdurft, H.C., Gallant, J.L. & van Essen, D.C. (1999). Response modulation by texture surround in primate area V1: Correlates of "popout" under anesthesia, *Visual Neuroscience, 16*, 15–34.

Nothdurft, H.C., Gallant, J.L. & van Essen, D.C. (2000). Response profiles to texture border patterns in area V1, *Visual Neuroscience, 17*, 421–436.

Nothdurft, H.C. (1994) Common properties of visual segmentation. In Higher Order Processing in the Visual System. Eds. G.R. Bock and J.A. Goode, Chichester, UK: Wiley (Ciba Foundation Symposium 184), pp. 245-259.

Philips, G. & Wilson, H. (1984). Orientation bandwidths of spatial mechanisms measured by masking. *J. Opt. Soc. Am. A, 1*(2), 226-232.

Poltoratski, S., & Xu, Y. (2013). The association of color memory and the enumeration of multiple spatially overlapping sets. *Journal of vision, 13*(8), 6.

Pylyshyn, Z.W. & Storm, R.W. (1988). Tracking multiple independent targets: Evidence for a parallel tracking mechanism. *Spatial vision, 3*(3), 179-197.

Raphael, S., Dillenburger, B., & Morgan, M. (2013). Computation of relative numerosity of circular dot textures. *Journal of Vision, 13*(2), 17-17.

Revkin, S. K., Piazza, M., Izard, V., Cohen, L., & Dehaene, S. (2008). Does subitizing reflect numerical estimation? *Psychological Science, 19*, 607–614.

Robson, J. G. (1980). Neural images: The physiological basis of spatial vision. In C. S. Harris (Ed.), Visual Coding and Adaptability (pp. 177–214). Hillsdale, NJ: Erlbaum.

Rosenholtz, R. (2001). Search asymmetries? What search asymmetries?. *Attention, Perception, & Psychophysics, 63*(3), 476-489.

Ross, J. (2003). Visual discrimination of number without counting. *Perception, 32*(7), 867-870.

Sagi, D. & Julesz, B. (1985). 'Where' and 'what' in vision, *Science, 228*, 1217–1219.

Schofield, A.J., & Foster, D.H. (1995). Artificial neural networks simulating visual texture segmentation and target detection in line-element images. *Philosophical Transactions of the Royal Society of London, Series B, 350* (1334), 401-412.

Snippe, H. P., & Koenderink, J. J. (1994). Discrimination of geometric angle in the fronto-parallel plane. *Spatial Vision, 8*(3), 309-328.

Sun, P., Chubb, C., Wright, C. E., & Sperling, G. (2016). The centroid paradigm: Quantifying feature-based attention in terms of attention filters. *Attention, Perception, & Psychophysics, 78*(2), 474-515.

Sun, P., Chubb, C., Wright, C. E., & Sperling, G. (2016). Human attention filters for single colors. *Proceedings of the National Academy of Sciences, USA, 1137*(43), 1-9, E6712-E6720.

Treisman, A. (1985). Preattentive processing in vision. Computer *Vision, Graphics, and Image Processing, 31*(2), 156-177.

Treisman, A. & Souther, J. (1985). Search asymmetry: a diagnostic for preattentive processing of separable features. *J. Exp. Psychol. Gen. 114*, 285-310.

Treisman, A. & Gormican, S. (1988). Feature analysis in early vision: Evidence from search asymmetries. *Psychological Review, 95*(1), 15-48.

Trick, L.M. & Pylyshyn, Z.W. (1993). What enumeration studies can show us about spatial attention: evidence for limited capacity preattentive processing. *Journal of Experimental Psychology: Human Perception and Performance, 19*(2), 331.

Trick, L.M. & Pylyshyn, Z.W. (1994). Why are small and large numbers enumerated differently? A limited-capacity preattentive stage in vision. *Psychological review, 101*(1), 80.

Van Essen, D. C., DeYoe, E. A., Olavarria, J., Knierim, J., Fox, J., Sagi, D. & Julesz, B. (1989). Neural responses to static and moving texture patterns in visual cortex of the Macaque monkey. In Lam, D. M. K. & Gilbert, C. D. (Eds.), Neural mechanisms of visual perception (pp. 137-154). Woodlands, TX: Portfolio.

Westland, S. and Foster, D. H. (1995). Optimized model of oriented-line-target detection using vertical and horizontal filters*, J. Opt. Soc. Amer. A 12*, 1617–1622.

Whalen, J., Gallistel, C. R., & Gelman, R. (1999). Nonverbal counting in humans: The psychophysics of number representation. *Psychological Science, 10*(2), 130-137.

Wolfe, J. M. (1994). Guided search 2.0 a revised model of visual search. *Psychonomic bulletin & review, 1*(2), 202-238.

Wolfe, J. M., Chun, M. M., & Friedman-Hill, S. R. (1995). Making use of texton gradients: Visual search and perceptual grouping exploit the same parallel processes in different ways. In T. Papathomas, C. Chubb, A. Gorea & E. Kowler (Eds.), Early vision and beyond. (pp. 189-198). Cambridge, MA: MIT Press.

Wolfe, J.M., Friedman-Hill, S.R, Stuart, M.I., O'Connell, K.M. (1992). The role of categorization in visual search for orientation. J. *Exp. Psychol. Hum. Percept. Perform., 18*, 34-49.

Wolfe, J. M., & Friedman-Hill, S. R. (1992). Visual search for oriented lines: The role of angular relations between targets and distractors. *Spatial Vision, 6*(3), 199-207.