Title

Visual Search Does Not Fully Characterize Feature-Based Selective Attention: Evidence from the Centroid Paradigm

Author

Winter, Amy Nicole

Publication Date

2017

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA,
IRVINE

Visual Search Does Not Fully Characterize Feature-Based Selective Attention:
Evidence from the Centroid Paradigm

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Psychology

by

A. Nicole Winter

Dissertation Committee:
Professor Charles F. Chubb, Chair
Professor Charles E. Wright
Professor Donald D. Hoffman

2017

DEDICATION

To my fiancé, Chris Mitsch
To my parents, Carie Osburn and Bob Winter

TABLE OF CONTENTS

			Page
LIST OF FIGURES			iv
LIST OF TABLES			vi
ACKNOWLEDGMENTS			vii
CURRICULUM VITAE			viii
ABSTRACT OF THE DISSERTATION			ix
INTRODUCTION			1
CH. 1	CONJUNCTIVE TARGETS IN CENTROID AND SEARCH TASKS		6
	1.1	Introduction	6
	1.2	Experiment 1: Size-Color Conjunctions	10
		1.2.1 Methods	10
		1.2.2 Results	17
		1.2.3 Conclusions	27
	1.3	Experiment 2: Luminance-Shape Conjunctions	28
		1.3.1 Methods	28
		1.3.2 Results	33
		1.3.3 Conclusions	41
	1.4	General Discussion	41
CH. 2	EXPERIMENT 3: TARGET-DISTRACTOR SIMILARITY		50
	2.1	Introduction	50
	2.2	Methods	52
		2.2.1 Display Types	53
		2.2.2 Experimental Conditions	55
	2.3	Results	58
	2.4	Discussion and Conclusions	68
CH. 3	EXPERIMENT 4: EQUISALIENCE ANALYSIS		71
	3.1	Introduction	71
	3.2	Methods	75
		3.2.1 Participants	75
		3.2.2 Procedure	76
		3.2.3 Stimuli	80
		3.2.4 Design	82
	3.3	Results	85
	3.4	Discussion and Conclusions	92
CONCLUSION			94
REFERENCES			97

LIST OF FIGURES

		Page
Figure 1.1	Feature vs. Conjunctive Targets in Visual Search Displays	7
Figure 1.2	Example Centroid Trial	9
Figure 1.3	Exp. 1 Centroid Displays	12
Figure 1.4	Exp. 1 Visual Search Displays for Feature Targets	13
Figure 1.5	Exp. 1 Visual Search Displays for Conjunctive Targets	14
Figure 1.6	Exp. 1 Visual Search Reaction Times	18
Figure 1.7	Exp. 1 Centroid Attention Filters	22
Figure 1.8	Exp. 1 Centroid Efficiencies	24
Figure 1.9	Exp. 1 Centroid Selectivity Ratios	24
Figure 1.10	Exp. 2 Centroid Displays	29
Figure 1.11	Exp. 2 Visual Search Displays for Feature Targets	30
Figure 1.12	Exp. 2 Visual Search Displays for Conjunctive Targets	31
Figure 1.13	Exp. 2 Visual Search Reaction Times	34
Figure 1.14	Exp. 2 Centroid Attention Filters	37
Figure 1.15	Exp. 2 Centroid Efficiencies	38
Figure 1.16	Exp. 2 Centroid Selectivity Ratios	38
Figure 1.17	Replicated Exp. 1 Centroid Results	45
Figure 1.18	Replicated Exp. 1 Centroid Results (Aggregated)	46
Figure 1.19	Replicated Exp. 2 Centroid Results	47
Figure 1.20	Replicated Exp. 2 Centroid Results (Aggregated)	48
Figure 2.1	Example Exp.3 Centroid Trial	52
Figure 2.2	Exp. 3 Item Types	53
Figure 2.3	Exp. 3 Display Types	54
Figure 2.4	Exp.3 Relative Influence	60
Figure 2.5	Exp. 3 Efficiency	61
Figure 2.6	Exp. 3 Selectivity Ratio	62

Figure 3.1	Example Exp. 4 Trial for the 2-Item Task	76
Figure 3.2	Example Exp. 4 Trial for the 8-Item Task	77
Figure 3.3	Example Centroid Threshold Trial	79
Figure 3.4	Exp. 4 Luminance Levels	81
Figure 3.5	Exp. 4 Shape Levels	81
Figure 3.6	Exp. 4 Stimulus Displays for the 2-Item Task	83
Figure 3.7	Exp. 4 Stimulus Displays for the 8-Item Task	84
Figure 3.8	Comparison of Exp. 4 Conditions for Participant 1	86
Figure 3.9	Comparison of Exp. 4 Conditions for Participant 2	87
Figure 3.10	Comparison of Exp. 4 Conditions for Participant 3	88
Figure 3.11	Comparison of Exp. 4 Conditions for Participant 4	89
Figure 3.12	Exp. 4 Equisalience Plots	91

LIST OF TABLES

		Page
Table 1.1	Exp. 1 Paired-Samples T-tests of Search Slopes for Positive Trials	19
Table 1.2	Exp. 1 Paired-Samples T-tests of Search Slopes for Negative Trials	20
Table 1.3	Exp. 1 Paired-Samples T-tests of Centroid Efficiencies	25
Table 1.4	Exp. 1 Paired-Samples T-tests of Centroid Selectivity Ratios	26
Table 1.5	Exp. 2 Paired-Samples T-tests of Search Slopes for Positive Trials	35
Table 1.6	Exp. 2 Paired-Samples T-tests of Search Slopes for Negative Trials	36
Table 1.7	Exp. 2 Paired-Samples T-tests of Centroid Efficiencies	39
Table 1.8	Exp. 2 Paired-Samples T-tests of Centroid Selectivity Ratios	40
Table 2.1	Exp. 2 Display Type Efficiencies by Target Condition	64
Table 2.2	Exp. 2 Display Type Selectivity Ratios by Target Condition	65
Table 2.3	Exp. 2 Target Condition Selectivity Ratios by Display Type	66
Table 2.4	Exp. 2 Target Luminance and Shape Difference of Differences	67

ACKNOWLEDGMENTS

First and foremost, I would like to thank my advisers, Ted Wright and Charlie Chubb, for their encouragement, support, and mentoring. In particular, I would like to thank Charlie for taking a chance on me when he first welcomed me to the lab, and Ted for being my adviser long before he was officially recognized as such.

I would like to thank Don Hoffman for cultivating my interest in research as an undergraduate student, and for serving on my both my advancement and dissertation committees.

I would like to thank Dean Frances Leslie for advocating on behalf of graduate students in general and me in particular.

I would like to thank John Sommerhauser for always taking time out of his busy day to answer questions and provide invaluable instrumental support.

I would like to thank the School of Social Sciences and the Department of Cognitive Sciences, particularly Clara Schultheiss, Jessice Cañas-Castañeda, and Ruth Kim.

Finally, I would like to thank my friends, family, and colleagues for all their help along way. In particular, I would like to thank Chris Mitsch and Sierra Broussard, as I cannot imagine the last few years without your friendship.

CURRICULUM VITAE

A. Nicole Winter

EDUCATION

Ph.D. Psychology, Cognitive Neuroscience Concentration	**2017**
University of California, Irvine	*Irvine, California*
M.S. Cognitive Neuroscience	**2014**
University of California, Irvine	*Irvine, California*
B.A. Psychology and Social Behavior	**2012**
University of California, Irvine	*Irvine, California*

RESEARCH EXPERIENCE

Graduate Research Assistant	**2012-2017**
University of California, Irvine	*Irvine, California*

TEACHING EXPERIENCE

Teaching Assistant	**2012-2017**
University of California, Irvine	*Irvine, California*

ABSTRACT OF THE DISSERTATION

Visual Search Does Not Fully Characterize Feature-Based Selective Attention:
Evidence from the Centroid Paradigm

By

A. Nicole Winter

Doctor of Philosophy in Psychology

University of California, Irvine, 2017

Professor Charles Chubb, Chair

While the visual search task has been instrumental in the study of feature-based attention (i.e., how attentional mechanisms increase the salience of relevant features), no single methodology can tell the whole story. This dissertation investigates the contributions that a complementary methodology, the centroid task, can make to our understanding of feature-based attention. The first chapter directly compares the search and centroid tasks. It replicates the expected search results, finding that performance is always worse for conjunctive targets than it is for feature targets. However, it reports a different pattern of centroid results: conjunctive target centroid judgments can actually outperform constituent-feature centroid judgments. The second chapter examines the role of target-distractor similarity in centroid estimations. It finds that, given sufficiently salient feature contrasts, conjunctive target conditions are better than or equal to *both* constituent-feature conditions, suggesting that there is not necessarily any cost to conjunctive centroid judgments. The third chapter reviews an equisalience analysis of an 8-item centroid task

and an analogous 2-item task. The different equisalience functions for the two tasks suggest that they access information differently. Together, these chapters provide compelling evidence that the centroid paradigm allows us to study aspects of feature-based attention that visual search cannot capture.

INTRODUCTION

At any given moment, we are inundated with far more sensory input than we can possibly process—or, rather, more input than we can process fully. Attentional mechanisms, then, are necessary to determine which stimuli are selected for further processing and which are ignored. This selection matters because attention is a limited resource: the attentional selection of any one stimulus (or part of a stimulus) means there are fewer resources available to attend to other stimuli (or other parts of the same stimulus). Across all sensory modalities, the goal of attention is to prioritize the processing of relevant stimuli over that of irrelevant stimuli.

Visual attention is typically broken down into three main categories: spatial attention, object-based attention, and feature-based attention. Posner (1980) described spatial attention as a spotlight whose beam illuminates attended regions of a visual scene. Like a spotlight, the attentional beam moves through space and can be contracted to narrowly focus attention or expanded to more broadly disperse it. Shifts of spatial attention can be voluntary (endogenous) or reflexive (exogenous). Endogenous shifts take 300 msec or longer, while exogenous shifts typically occur in the range of 100-200 msec (e.g., Müller & Rabbitt, 1989). Object-based attention is the idea attention can select particular objects for further processing. In support of this idea, Kahneman, Treisman, and Gibbs (1992) report that observers are faster at making discriminations within a single, attended object than across multiple, even when controlling for distance. This dissertation, however, focuses on feature-based attention: how attention enhances the salience of particular features.

For an example of feature-based attention, consider the following scenario. You are attending a conference with a colleague who, on this particular day, is wearing a purple sweater. After attending different sessions in the morning, you decide to meet for lunch in hotel lobby. However, the hotel lobby is packed with other conference goers, and it is impossible for you to attend to all of them at once. Remembering the color of your colleague's sweater, you can efficiently search the crowded, complicated visual scene for just this color feature. This will enhance the salience of all purple items, including your colleague's sweater, allowing you to quickly locate her.

The way features guide visual attention has been the subject of much investigation, and Treisman's Feature Integration Theory (FIT) continues to be one of the most influential—if not *the* most influential—accounts of this process. Treisman and Gelade first proposed FIT in 1980, based primarily on a series of compelling visual search experiments. In the visual search task, a participant searches for a previously-defined target among an array of distractors. The target can be defined by a single feature (e.g., a red X among green Xs) or by a conjunction of feature (e.g., a red X among green Xs and red Os). The participant's task is to indicate as quickly and accurately as possible whether or not the target is present. Her reaction time is recorded and can be plotted as a function of display size—that is, the total number of items (targets and distractors) in the search array. Treisman and Gelade (1980) reported that, on positive (target present) trials, feature targets produce virtually no change in search time as the display size increases. However, they found that search time for conjunctive targets slows as more items are added to the display. According to FIT, these results reveal two stages of processing. The first (preattentive) stage involves parallel processing of features, creating individual feature maps. When the target is defined by a single feature, this first stage is sufficient for target detection: since all the items are processed in parallel, there is no performance cost to processing displays with more items, resulting in flat search slopes. The second (attentional) stage, however, is necessary for the detection of conjunctive targets. While the constituent features have already been processed in the first stage, attention is needed to bind those features together. So, for example, the first stage may have produced both a *red* map and an *X* map, but these maps—separately—provide no way of detecting an item that is both red and an X. It is then necessary have second stage of processing during which focused attention is deployed serially to each item location in order to integrate the relevant feature maps. Increasing display size, therefore, means that attention must be deployed to more locations, which in turn increases search time. (Treisman & Gelade, 1980; Treisman, 1985).

Treisman initially drew a sharp distinction between the processing demands of feature and conjunctive search, claiming that feature search could always be accomplished preattentively via parallel processing while conjunctive search always required serial allocation of attention to each item until the target is detected (e.g., Treisman & Gelade, 1980; Treisman, 1982). However, subsequent research would later reveal this to be too strong a position. For example, Houck and Hoffman (1986) found that color-form conjunctions could be registered preattentively and Nakayama and Silverman (1986) reported flat search slopes for stereo-motion and stereo-color conjunctions. Treisman (1988) responded by allowing for the possibility of early conjunctions—that is, for the possibility of preattentive registering of conjunctions in certain circumstances. She further added that feature inhibition was likely responsible for efficient conjunctive searches (Treisman & Sato, 1990). (For a comprehensive review of FIT, see Quinlan, 2003.)

Although it has undergone some revisions, FIT continues to be highly relevant to the study of feature-based attention. Similarly, the visual search task itself has made a lasting mark on the field—it is, in effect, the default methodology for investigating feature-based attention. The task has a relatively simple design and does not require any particularly fancy equipment. Search slopes are easily calculated (change in reaction time divided by change in display size) and easily interpreted (scanning rate per item). It is little wonder, then, why visual search reigns supreme.

However, all tasks have their own particular task demands; there is no escaping this fact, regardless of how useful or how popular a task may be. In this way, the prevalence of visual search may actually be a hindrance to the study of feature-based attention. It is possible that idiosyncrasies of visual search task are being confused with inherent properties of feature-based attention. So, without complementary methodologies, there is no way to know what is merely a task idiosyncrasy versus what is an actual characteristic of feature-based attention.

One promising complementary methodology is the centroid paradigm. As in the search task, targets can be defined by a single feature or by some combination of features. A participant in a centroid task briefly sees a stimulus cloud containing targets and distractors, and then estimates the centroid, or center of mass, of the targets. After 100 (or, in some cases, even fewer) centroid trials, we can obtain rich measures of performance as described by Sun, Chubb, Wright, and Sperling (2016). This dissertation presents evidence that some paradigmatic visual search results do not necessarily replicate in the centroid paradigm. This suggests that there is more to the story of feature-based attention than visual search can tell.

CHAPTER 1

CONJUNCTIVE TARGETS IN CENTROID AND SEARCH TASKS

1.1 INTRODUCTION

The feature-integration theory (FIT) of attention (Treisman & Gelade, 1980) continues to be hugely influential after more than three decades. This theory proposes that attention can select a particular feature, such as the color red, and heighten its salience in a visual scene. For instance, a teacher grading papers may need to locate her red pen in a cluttered office space. As she scans her desk, her attention focuses on each red item she comes across, while easily ignoring items of other colors. This heightened salience of all things red would be an example of what Treisman and Gelade called "feature-based attention" (FBA). The authors supported the concept of FBA using a series of visual search experiments.

In a typical visual search experiment, a participant is presented with an array of items and asked to indicate whether a previously-defined target item is present or absent. In some cases, the targets may be defined by a single feature, such as shape or color. For instance, the target could be an 'X' in a field of 'O' distractors (Figure 1.1a), or a green item in a field of red distractors (Figure 1.1b). When the target is defined in this way (by a single feature), it appears to "pop out"; no matter how many distractors are in the display, the participant can quickly spot a present target. In other cases, however, the target may be defined by a conjunction of features. For instance, the target could be a green 'X' in a field of red 'X' and green ‘O’ distractors (Figure 1.1c). These conjunctive targets no longer exhibit this “pop-out” effect and are much more difficult to spot.

The difficulty of finding a target can be quantified by measuring reaction time: how long it takes an observer to indicate whether the target was present or absent. Treisman and Gelade (1980) found that, when single-feature targets were present, reaction times remained virtually unchanged as the number of items in the display increased. When the target was defined by a conjunction of features, however, reaction time increased as a function of display size. These results, which have been replicated many times over (e.g. Bergen & Julesz, 1983; Egeth, Virzi, & Garbart, 1984; Nakayama & Silverman, 1986; Wolfe & Franzel, 1988), suggest a distinction between parallel and serial search. Participants can process all display items in parallel, Treisman and Gelade claim, for single-feature targets; however, for conjunctive targets, participants must scan each item individually. In this scanning process, attention is allocated to each item in order to bind the relevant features together (Treisman & Gelade, 1980).

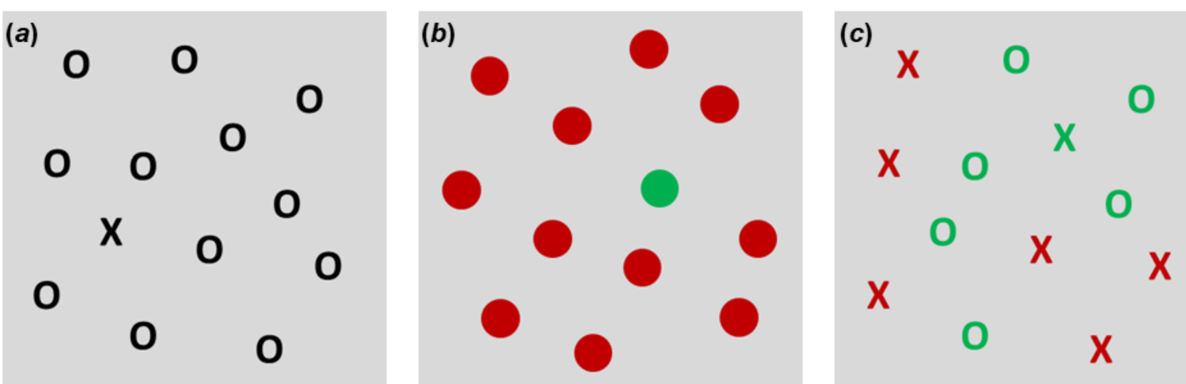


Figure 1.1: Feature vs. Conjunctive Targets in Visual Search Displays. *(a)* The target “X” is defined only by shape. *(b)* The target green circle is defined only by color. (*c*) The target green “X” is defined by both color and shape. When the target is defined by a single feature, as in (*a*) and *(b*), it appears to pop out from the distractors. However, when the target is defined by a conjunction of features, as in (*c*), it loses its pop-out effect.

Despite the many replications of this same basic pattern of results for single-feature and conjunctive targets, a few exceptions have been noted. For example, Nakayama and Silverman (1986) conducted a series of visual search experiments using the feature dimensions of color, motion, and stereoscopic disparity. As expected, the single-feature motion and color conditions produced flat reaction-time slopes, suggesting parallel search. One of the conjunction conditions, motion-color, produced the linearly increasing search slopes associated with serial search. However, the other two conjunction conditions, stereo-motion and stereo-color, produced flat reaction slopes indicative of parallel processing. In addition, Theeuwes and Kooi (1994) examined the feature dimensions of polarity, color, and shape in visual search. They also found flat (parallel) search slopes for all feature conditions. In the color-shape conjunction condition, they observed serial search slopes, but these slopes were relatively shallow. However, the slopes in the polarity-shape condition were flat, suggesting another case of parallel search for a conjunctive target. Even so, the visual search task, and its interpretation under FIT, has dominated much of the FBA literature.

However, the visual search task is not the only available methodology for studying FBA. The centroid paradigm (Sun, Chubb, Wright, & Sperling, 2016), for example, also requires the participant to use attention filters deployed over space. Consider the teacher from our previous example—the one searching for a red pen—and suppose that, during her search, she knocks over a container of blue thumbtacks, scattering them on the floor. Her task has now changed: instead of finding her red pen, she needs to track information about the spread of blue thumbtacks on the floor in order to avoid stepping on them. She could do this by deploying a blue-selective attention filter, binding the blue-feature of each

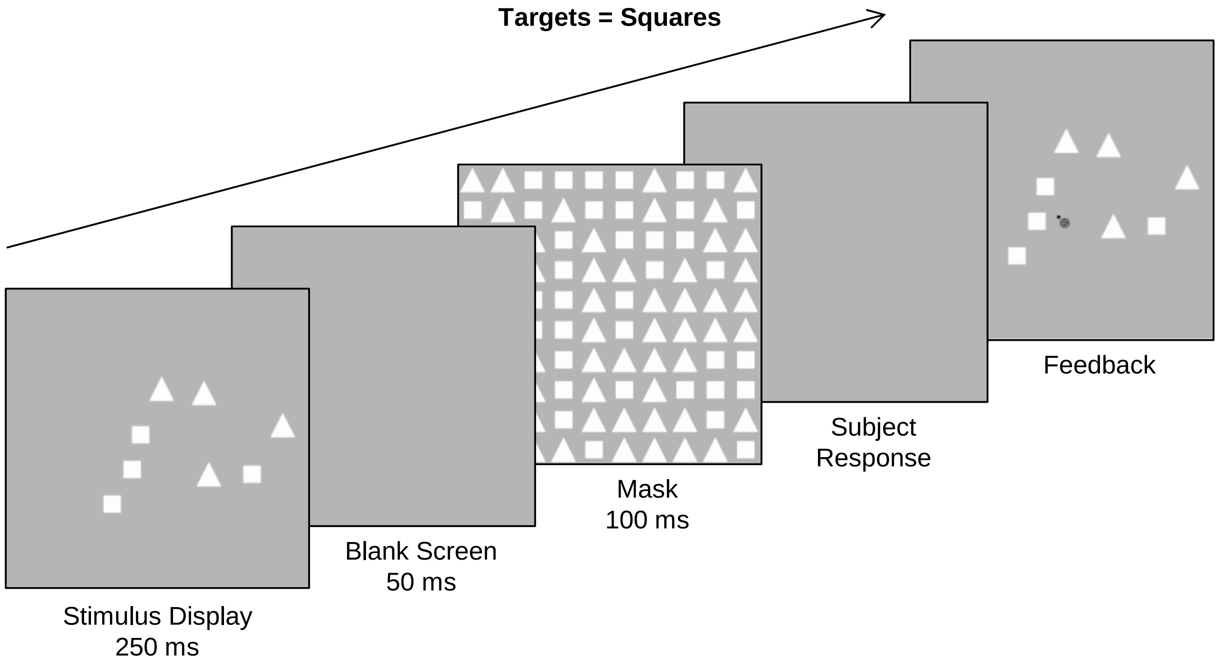


Figure 1.2: Example Centroid Trial. The participant briefly sees the stimulus display, followed by a blank screen and a mask. In this example, the squares are the target items, so her task is to estimation the centroid (or center of mass) of the squares. She responds via mouse click, then receives feedback showing her response (the small black square) and the actual centroid of the target items (the bullseye).

thumbtack with its location. She could also average the location information to estimate the epicenter of the area she wishes to avoid. The "epicenter" in this case would be the centroid, or center of mass of the thumbtacks. This centroid task can be used in controlled experiments to, in some sense, measure an observer's ability to construct and use such attention filters.

In a typical centroid task, a participant is briefly shown an array of items made up of targets and distractors. As in the visual search task, the targets can be defined by a single feature or a conjunction of features. After the stimulus display, a blank screen appears, followed by a visual mask. The purpose of the brief stimulus display time and the backwards mask is to prevent endogenous shifts of spatial attention. The participant then strives to mouse-click the centroid, or center of mass, of the targets. Figure 1.2 shows an example centroid trial. After 50-200 trials, we can accurately estimate (1) a lower bound on the number of items in the display she processes trial by trial to produce her responses, and (2) how much influence she gives each item type in her estimations. If performance on a particular centroid task is highly accurate, it suggests that there are populations of neurons that code for the feature or features that define the target.

Different tasks, of course, have different task demands. Given the prevalence of visual search in the FBA literature, we seek to examine the relationship between the centroid paradigm and visual search. In particular, do the same features that support "pop out" in visual search lead to improved performance on the centroid task? In this chapter, we will present preliminary evidence that suggests the answer is *no*, or at least *not always*. We replicate the standard visual search results: flat reaction times for single-feature targets and increasing reaction times for conjunctive targets as display size increases. However, in the centroid task, we find evidence for *improved* performance on conjunction conditions relative to single-feature conditions—that is, conditions in which target items all share the same single feature which, in the context of a search task, would suffice to produce pop-out.

1.2 EXPERIMENT 1: SIZE-COLOR CONJUNCTIONS

1.2.1 Methods

Experiment 1 consisted of both visual search and centroid tasks, completed in an ABBA order. For both tasks, the feature dimensions were size (big/small) and color (red/green), creating four distinct item types. All display items were squares subtending either 0.50 (big) or 0.29 (small) degrees of visual angle and presented against a gray background with Lxy triples (60.37, 0.32, 0.34). The exact color values for the red and green squares varied by participant (n=8) to achieve perceptual isoluminance; however, representative Lxy triples are (43.88, 0.38, 0.31) for a red item and (43.24, 0.26, 0.37) for a green item. Both tasks contained eight blocks, one for each of the eight target conditions (four single-feature, four conjunction). In the single-feature conditions, targets were defined either by size (big/small) or color (red/green) while, in the conjunction conditions, targets were defined by both size and color. We used a Latin square design for single-feature and conjunction conditions separately, then alternated whether the single-feature or conjunction conditions came first each session. The order of single-feature versus conjunction conditions was counterbalanced across participants.

The red and green used to create stimulus items were derived using a minimum motion procedure (Anstis & Cavanagh, 1983; Herrera, 2016). Psychometric data were collected to derive 20 lights of different hues, each (1) maximally saturated on the display device used in these experiments and (2) motion-equiluminant to the background gray used in the minimum motion stimulus displays. Each of these 20 lights was then projected to the corresponding point in the space spanned by the Stockman and Sharpe (1999) 2 deg. cone fundamentals, and the best fitting plane was taken as an estimate of the participant's equiluminant plane. The red and green used in the experiment for a given participant were the extreme points on the line in this plane with S-cone activation equal to S-cone activation produced by the background gray in the minimum motion procedure. Thus this red and green are drawn from opposite sides of the "constant-S" axis of DKL space

Figure 1.3: Exp. 1 Centroid Displays. The same centroid display could be used for any of the four single-feature targets (red, green, large, small) or any of the four conjunctive targets (red & large, red & small, green & large, green & small). All single-feature conditions had 8 items per display (4 targets, 4 distractors) and all conjunctive conditions had 16 items per display (4 targets, 12 distractors) in order to keep the number of targets constant.

(Derrington, Krauskopf, & Lennie, 1984). The background gray (60.37 cd/m^2) used in the visual search and centroid tasks was lighter than the equiluminant background gray (40.87 cd/m^2) used in the minimum motion procedure. Pilot data from our lab, however, suggest that the same pattern of results emerges whether or not the color of the stimulus items is isoluminant with the background. The example stimulus displays (Figures 1.3, 1.4, and 1.5) slightly exaggerate the item-background luminance difference for easier viewing. Participants with no prior centroid experience first completed 500 trials of target-only centroid training, in order to minimize (1) idiosyncratic centroid computations across participants and (2) random response noise for each participant individually (Sun et al., 2016). Centroid training consisted of four blocks (125 trials each) grouped by item type (black circles, black triangle, white circles, and white triangles). Stimulus cloud displays contained only one item type and 4, 8, 12, 16, or 20 items per display.

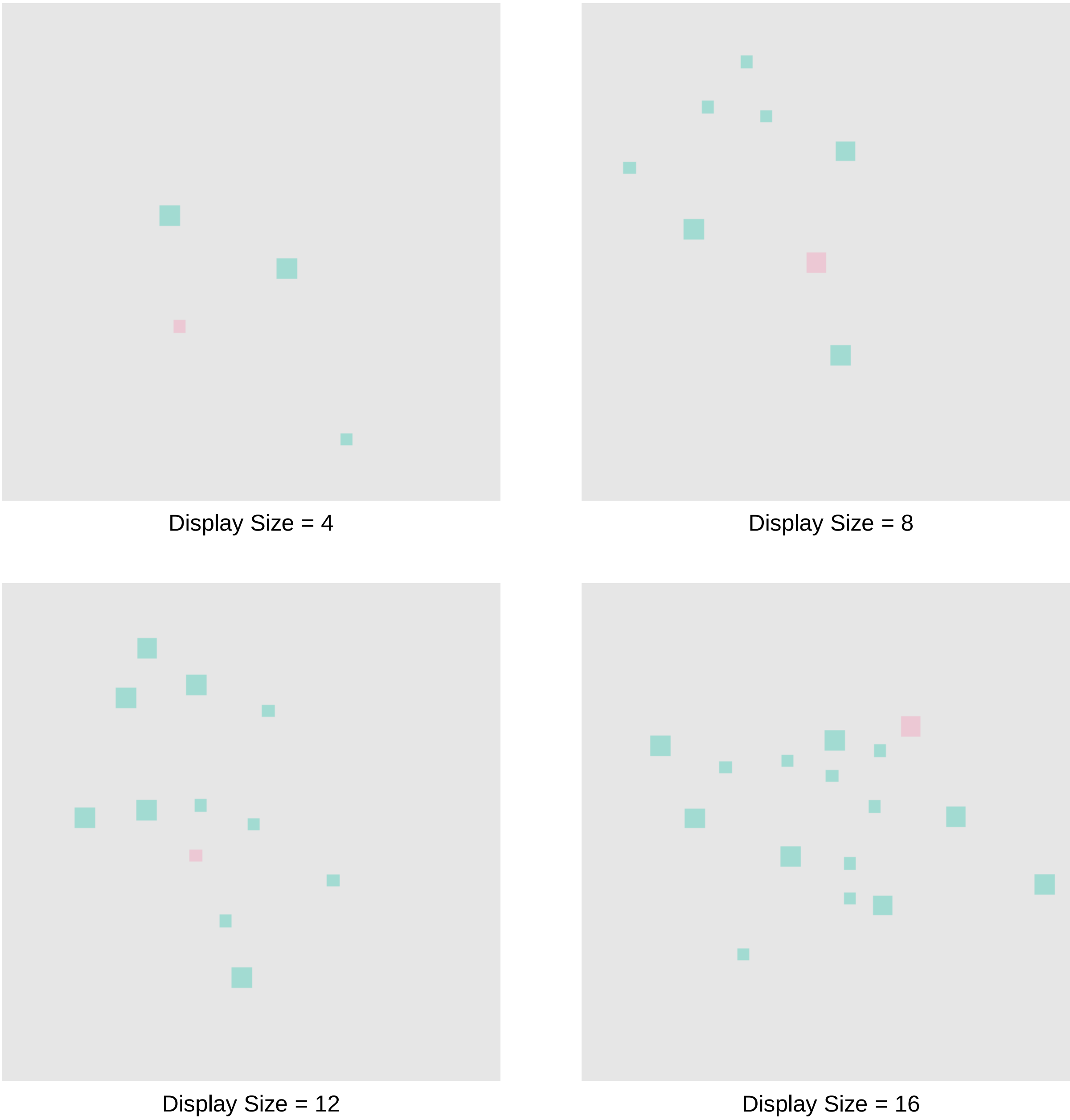


Figure 1.4: Exp. 1 Visual Search Displays for Feature Targets. The target in all four example displays is a red square, making color the relevant feature dimension and size the irrelevant feature dimension. The target item could take on either level of the irrelevant feature dimension; here, the red target could be either big or small.

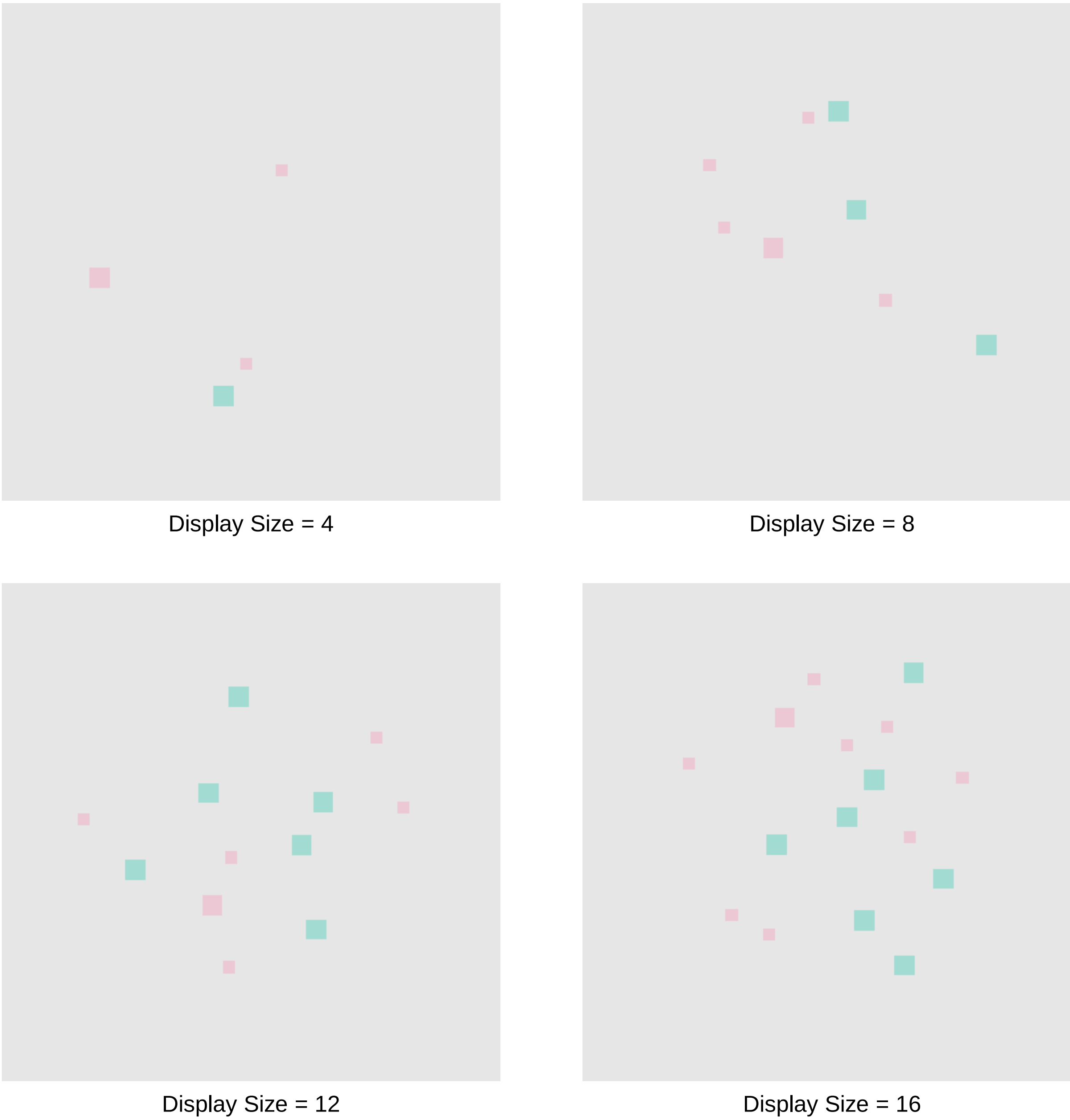


Figure 1.5: Exp. 1 Visual Search Displays for Conjunctive Targets. The target in all four example displays is a large, red square. Both color and size are relevant feature dimensions.

All participants began with a practice session identical to the experiment, except with randomized block orders and fewer trials. There were 25 practice trials per condition for the centroid task (200 practice trials total), and 40 practice trials per condition for the visual search task (320 practice trials total). Participants were undergraduates and graduate students at the University of California at Irvine with normal or corrected-to-normal vision.

In the centroid task, items were presented in a cloud stimulus display measuring 800x800 pixels, subtending approximately 13.75 degrees of visual angle. The dispersion of a stimulus cloud is given by

$$\text{Dispersion}(x, y) = \left[\frac{1}{2N_{cloud} - 1} \sum_{i=1}^{N_{cloud}} (x_i - \bar{X})^2 + (y_i - \bar{Y})^2\right]^{\frac{1}{2}}$$

where N_{cloud} is the total number of items in the cloud, $x = \left(x_1, x_2, \cdots, x_{N_{cloud}}\right)$ and $y = \left(y_1, y_2, \cdots, y_{N_{cloud}}\right)$ are the vectors of x- and y-coordinates of the items, and $\bar{X}$ ($\bar{Y}$) is the mean of vector x (y). Each cloud had a fixed dispersion of $133\frac{1}{3}$ pixels (2.30 degrees of visual angle), or one-sixth of the 800-pixel stimulus display.

Each cloud included an equal number of each item type: two of each on single-feature trials (for a total of 8 items), and four of each on conjunction trials (for a total of 16 items). The conjunction trials had twice as many items as the single-feature trials in order to keep the number of targets constant at four across all conditions. Figure 1.3 shows example stimulus clouds for the centroid task. The cloud was displayed for 250 ms, followed immediately by a blank screen for 50 ms, and then by a visual mask for 100 ms. The mask was a grid composed of the same item types as the cloud in order to maximize its effectiveness. Next, a cursor appeared at the center of the screen. The participant moved the cursor to enter her centroid estimation via mouse click. After every trial, the participant received feedback indicating both the location of her response (marked by a small, black dot) and the location of the target centroid (marked by a bulls-eye), overlaid on the original stimulus display. She advanced to the next trial by pressing the spacebar. There were 75 trials per block, eight blocks per session, and two experimental sessions for a total of 1,200 centroid trials per participant.

In the visual search task, we used exactly the same process to assign item locations that we used in the centroid task, except for varying the number of items in the array; display size was either 4, 8, 12, or 16 items. Half the trials were positive (target present) and half were negative (target absent). In single-feature conditions, both levels of the irrelevant dimension were present in equal numbers. For instance, in the red target block, half the display would be big, green squares and the other half would be small, green squares. On positive trials, one of the squares was selected at random to be the target and *only* its color changed, so the red target could be either big or small (see Figure 1.4).

In the conjunction conditions, both feature dimensions were relevant so, on positive trials, one of the distractors was selected at random to become the target. For example, in the red & big target block, there were big, green squares and small, red squares as distractors. If a big, green square was selected to be the target, we changed its color to red; if a small, red square was selected, we changed its size to be bigger (see Figure 1.5). All of these changes took place during the creation of the stimuli, so participants saw only the final product.

The task was to indicate, on each trial, whether or not the target was present. The participant pressed the *Z* key with her left hand to enter a "no" response or the *M* key with her right hand to enter a "yes" response. The search array remained on the screen until the participant entered her response, at which point she received visual correctness feedback. The feedback was displayed for 1000 ms, followed by a pause of either 250, 500, 750, or 1000 ms before the participant automatically advanced to the next trial. There were 120 trials per block, eight blocks per session, and two experimental sessions for a total of 1,920 visual search trials per participant.

1.2.2 Results

To analyze the visual search data, we first found each participant's median reaction time (RT) for each target condition, display size, and trial type (positive/negative), excluding incorrect trials. We then calculated the mean of the participants' median RTs and the slopes of the best-fitting line to the RT data for each condition and trial type (Figure 1.6). We conducted a series of paired-samples t-tests on the search slopes as described in Tables 1.1 and 1.2. The search slopes were flatter in the search-for-red and search-for-green conditions than in the search-for-large and search-for-small conditions. In addition, the single-feature target conditions had flatter slopes than the conjunctive target conditions. There was also a main effect of size, with search-for-large conditions producing flatter slopes than search-for-small conditions. Most of these differences, however, only reached statistical significance in the negative trial comparisons.

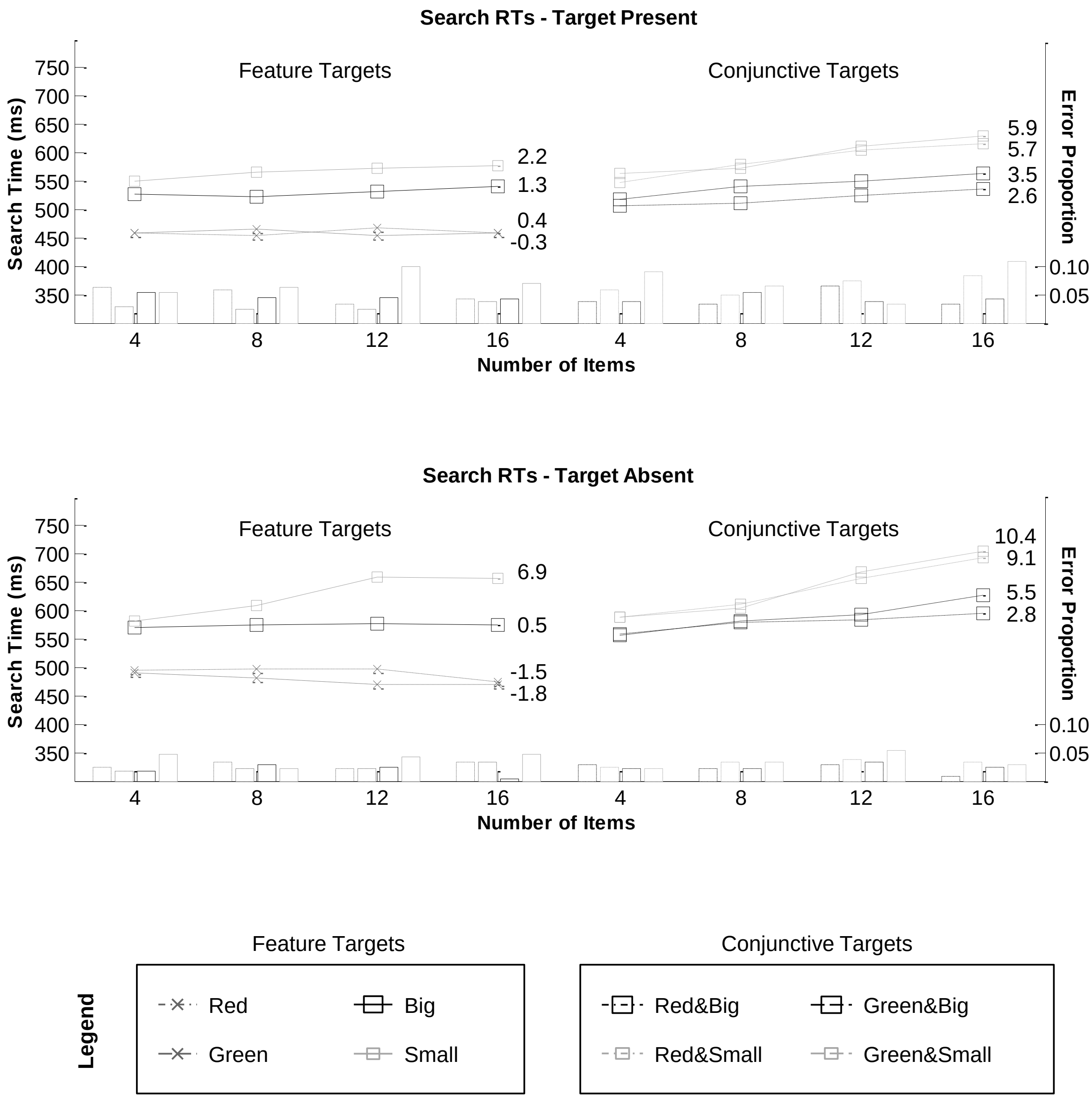


Figure 1.6: Exp. 1 Visual Search Reaction Times. The lines plot the mean of participants' median RTs for each target condition. The number the right of each line gives its slope. Overall, we found flatter RT slopes for single-feature conditions than for conjunction conditions. Notably, however, the RTs were longer and the slopes steeper for single-feature size conditions compared to the single-feature color conditions. The bars show the error proportion, which is the number of incorrect trials divided by the total number of trials in a condition. Incorrect trials were excluded from participants' median RTs.

(a) Comparisons between Feature Dimension Levels				
Contrast	$\mathbf{R-G}$	$\mathbf{B-S}$	$\frac{\mathbf{RB+RS}}{\mathbf{2}}-\frac{\mathbf{GB+GS}}{\mathbf{2}}$	$\frac{\mathbf{RB+GB}}{\mathbf{2}}-\frac{\mathbf{RS+GS}}{\mathbf{2}}$
Mean	0.710	-0.908	-0.564	-2.737
Standard Deviation	3.009	3.642	3.102	2.742
Upper Bound	3.225	2.137	2.030	-0.444
Lower Bound	-1.806	-3.952	-3.157	-5.029
T	0.667	-0.705	-0.514	-2.822
p	0.526	0.504	0.623	0.026
Bayes Factor	0.404	0.413	0.376	3.140

(b) Comparisons across Feature Dimensions			
Contrast	$\frac{\mathbf{R+G}}{\mathbf{2}}-\frac{\mathbf{B+S}}{\mathbf{2}}$	$\frac{\mathbf{R+G}}{\mathbf{2}}-\frac{\mathbf{RB+RS+GB+GS}}{\mathbf{4}}$	$\frac{\mathbf{B+S}}{\mathbf{2}}-\frac{\mathbf{RB+RS+GB+GS}}{\mathbf{4}}$
Mean	-1.725	-4.366	-2.642
Standard Deviation	2.383	1.924	3.174
Upper Bound	0.268	-2.757	0.011
Lower Bound	-3.717	-5.975	-5.295
T	-2.047	-6.417	-2.354
p	0.080	0.000	0.051
Bayes Factor	1.344	94.742	1.876

Table 1.1: Exp. 1 Paired-Samples T-tests of Search Slopes for Positive Trials. Letters refer to target conditions, with single letters indicating single-feature target conditions (e.g. 'R' = search-for-red) and two letters indicating conjunctive target conditions (e.g. 'RB' = search-for-red&big). Flatter (smaller) slopes indicate better performance while steeper (greater) slopes indicate worse performance. **(a)** The first and second columns compare the two color target conditions (R vs. G) and the two size target conditions (B vs. S), respectively. The third column compares the mean of the red conjunction conditions (RB and RS) with the mean of the green conjunction conditions (GB and GS) and, the fourth column compares the mean of big conjunction conditions (RB and GB) with the mean of small conjunction conditions (RS and GS). **(b)** The first column compares the mean of the color conditions with the mean of the size conditions, the second column compares the mean of the color conditions with the mean of the conjunction conditions, and the third column compares the mean of the size conditions with the mean of the conjunction conditions. These comparisons show that search slopes were flattest in the color conditions, next flattest in the size conditions, and steepest in the conjunction conditions, though not all these comparisons are statistically significant.

(a) Comparisons between Feature Dimension Levels				
Contrast	$\mathbf{R - G}$	$\mathbf{B - S}$	$\frac{\mathbf{RB + RS}}{\mathbf{2}} - \frac{\mathbf{GB + GS}}{\mathbf{2}}$	$\frac{\mathbf{RB + GB}}{\mathbf{2}} - \frac{\mathbf{RS + GS}}{\mathbf{2}}$
Mean	0.335	-6.455	-2.046	-5.554
Standard Deviation	2.683	2.960	2.176	5.651
Upper Bound	2.578	-3.981	-0.227	-0.830
Lower Bound	-1.908	-8.930	-3.866	-10.278
T	0.353	-6.168	-2.660	-2.780
p	0.734	0.000	0.032	0.027
Bayes Factor	0.354	77.807	2.628	2.999

(b) Comparisons across Feature Dimensions			
Contrast	$\frac{\mathbf{R + G}}{\mathbf{2}} - \frac{\mathbf{B + S}}{\mathbf{2}}$	$\frac{\mathbf{R + G}}{\mathbf{2}} - \frac{\mathbf{RB + RS + GB + GS}}{\mathbf{4}}$	$\frac{\mathbf{B + S}}{\mathbf{2}} - \frac{\mathbf{RB + RS + GB + GS}}{\mathbf{4}}$
Mean	-5.340	-8.580	-3.240
Standard Deviation	1.801	2.494	2.408
Upper Bound	-3.834	-6.495	-1.226
Lower Bound	-6.846	-10.664	-5.253
T	-8.385	-9.731	-3.805
p	0.000	0.000	0.007
Bayes Factor	380.279	854.462	8.985

Table 1.2: Exp. 1 Paired-Samples T-tests of Search Slopes for Negative Trials. Letters refer to target conditions, with single letters indicating single-feature target conditions (e.g. 'R' = search-for-red) and two letters indicating conjunctive target conditions (e.g. 'RB' = search-for-red&big). Flatter (smaller) slopes indicate better performance while steeper (greater) slopes indicate worse performance. **(a)** The first and second columns compare the two color target conditions (R vs. G) and the two size target conditions (B vs. S), respectively. The third column compares the mean of the red conjunction conditions (RB and RS) with the mean of the green conjunction conditions (GB and GS) and, the fourth column compares the mean of big conjunction conditions (RB and GB) with the mean of small conjunction conditions (RS and GS). **(b)** The first column compares the mean of the color conditions with the mean of the size conditions, the second column compares the mean of the color conditions with the mean of the conjunction conditions, and the third column compares the mean of the size conditions with the mean of the conjunction conditions. These comparisons show that search slopes were flattest in the color conditions, next flattest in the size conditions, and steepest in the conjunction conditions, and all these comparisons are statistically significant.

We used the methods described in Sun et al. (2016) to analyze the centroid task data. The model of Sun et al. (2016) assumes that the x- and y-coordinates of the participant's response on each trial are given by

$$R_x = \mu_x + Q_x \qquad \text{and} \qquad R_y = \mu_y + Q_y$$

for Q_x and Q_y independent, normally distributed random variables with mean 0 and some standard deviation σ and

$$\mu_x = \frac{\sum f_\varphi(\tau_i) x_i}{\sum f_\varphi(\tau_i)} \qquad \text{and} \qquad \mu_y = \frac{\sum f_\varphi(\tau_i) y_i}{\sum f_\varphi(\tau_i)}$$

where each sum is over all items i in cloud C, τ_i is the type of item i, and x_i and y_i are the *x*- and *y*- coordinates of item i, and f_φ is the attention filter achieved by the participant.

For a the attention condition with target filter φ, these methods enable us to estimate (1) the attention filter f_φ achieved by the participant in that condition, (2) the *Efficiency* with which the participant was able to deploy the filter f_φ, and (3) the *Data-drivenness* V of the participant's response-production process. The attention filter f_φ defines the relative influence exerted on the participant's responses by all four types of items occurring in the stimulus (large red, small red, large green and small green squares). The function f_φ is constrained to sum to 1; however, it is possible for f_φ to assign negative values to some item types. (Figure 1.7 plots the attention filter averaged across eight participants.) The participant's Efficiency in deploying f_φ is the minimum possible proportion of items that had to be included, on average, in the participant's centroid computation to achieve predicted responses of the accuracy observed. Efficiency is estimated by assuming that all

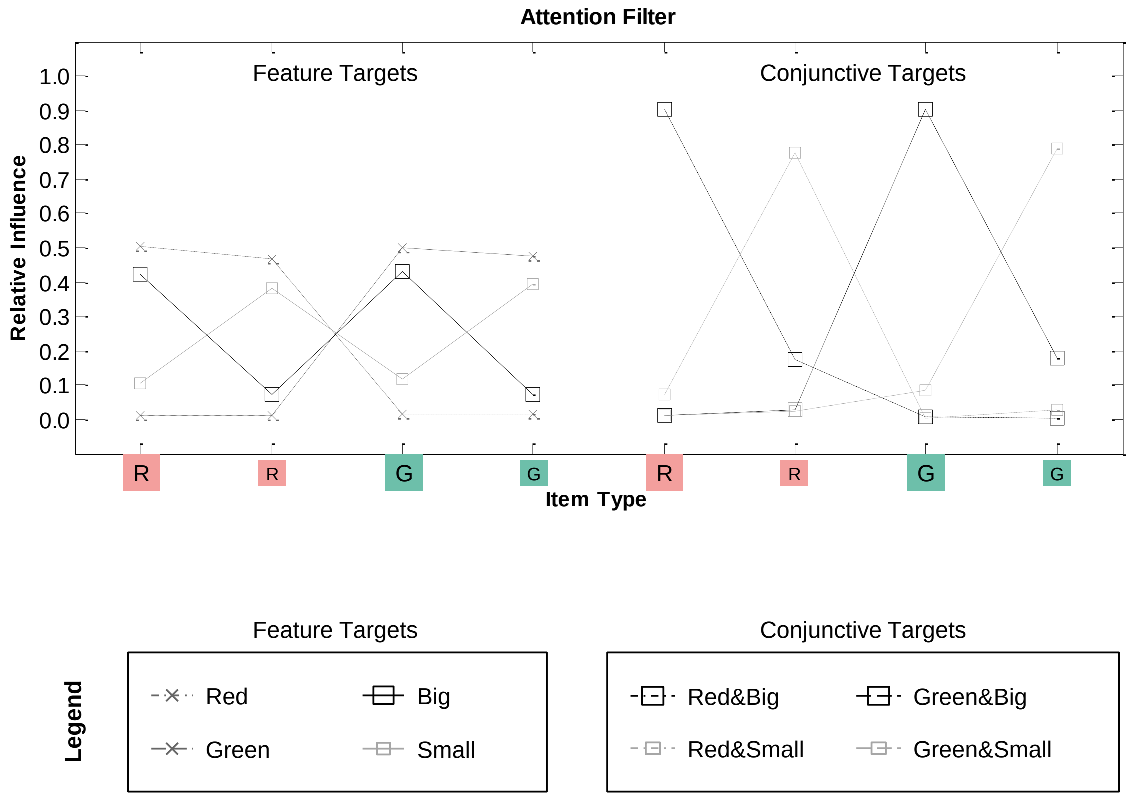


Figure 1.7: Exp. 1 Centroid Attention Filters. Each line shows the relative influence of each item type (averaged across eight participants) for a particular attention condition. The relative influence, or weight, of all the item types sums to 1. For all feature target conditions, the targets consist of two item types. The ideal attentional filter would be equally influenced by the two target item types (assigning them each a weight of 0.5) and not at all influenced by the two distractor item types (assigning them each a weigh of 0). Participants' actual performance follows these trends. For example, in the attend-to-red condition, the two red items (targets) both have weights of about 0.5 while the two green items (distractors) both have weights of about 0. For all conjunctive target conditions, the target consists of only one item type. The ideal attentional filter would be influenced only by the target item (assigning it a weight of 1), and not at all influenced by the three distractor item types (assigning them each a weigh of 0). Again, participants' actual performance is not far off. For example, in the attend-to-red&big target condition, the large red item type (target) has a weight of about 0.9, while the small red distractor has a weight of about 0.1 and the remaining green item types have a weight of about 0. In each conjunctive target condition, the distractor that shares the target's color exerts more influence on participants' centroid judgments compared to the other two distractor item types.

residual error (i.e., the deviations of responses predicted by the model from actual responses) is due to removing a fixed proportion Q of randomly chosen items from the display on each trial and applying the model to the decimated display without additional error; Efficiency is then taken to be equal to $1 - Q$. (Figure 1.8 plots Efficiency for each of target conditions averaged across eight participants.) Finally, Data-drivenness (V in Eq. 21 of Sun et al. 2016) reflects the degree to which the participant's response on each trial is determined by the stimulus presented on that trial as opposed to being drawn on each trial toward a fixed default location ($x_{default}, y_{default}$). In "binary" centroid tasks of the sort used in this paper in which the target filter assigns equal weight to a specific set of target item-types and weight 0 to the remaining distractor item-types, it is convenient to summarize the effectiveness of the attention filter f_φ achieved by the participant by the ratio of (numerator) the mean of $f_\varphi(t)$ taken across all target items t divided by (denominator) the mean of $\left|f_\varphi(d)\right|$ taken across all distractor items d. This statistic (selectivity ratio) provides a convenient index of the degree to which the attention filter achieved by the participant accentuates target items while filtering out distractor items. A selectivity ratio of ten or higher is considered excellent. We calculated the average efficiencies and selectivity ratios for participants across conditions (Figure 1.9).

We performed the same paired-samples t-tests on our centroid measures of efficiency (Table 1.3) and selectivity (Table 1.4) as we did on our search slope data. As in the visual search tasks, an asymmetry in centroid task performance was observed between the attend-to-color vs. the attend-to-size conditions. Just as reaction times were faster in search-for-red and search-for-green conditions than they were in the search-for-large and

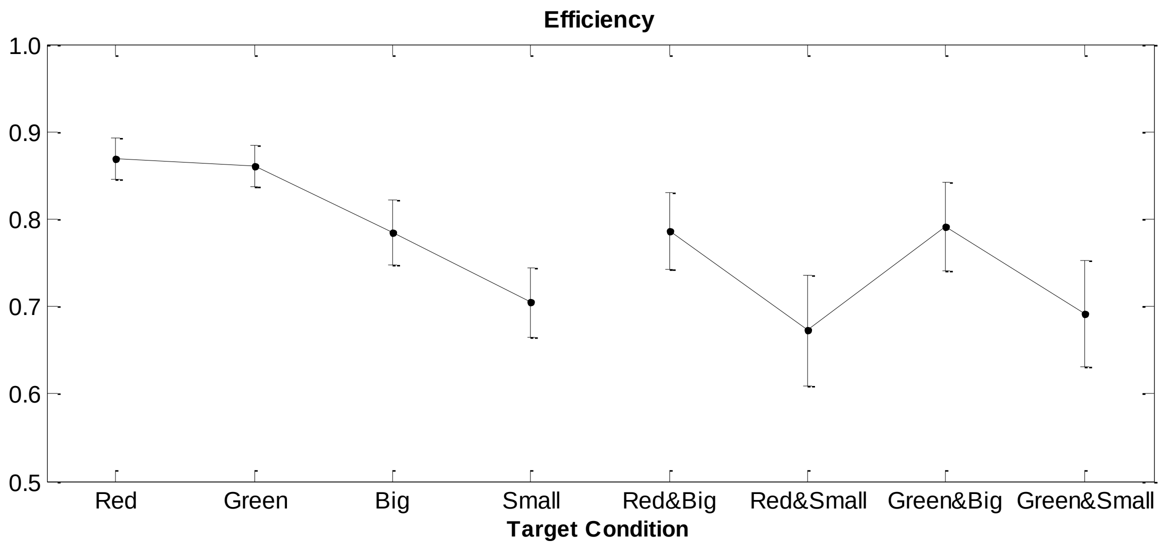


Figure 1.8: Exp. 1 Centroid Efficiencies. Efficiency estimates a lower bound on the proportion of items participant observed. The left side of the graph shows participants' average Efficiencies for feature target conditions, and the ride side shows participants' average Efficiencies for conjunctive target conditions. There were twice as many items in the conjunctive target conditions as in the feature target conditions, so lower Efficiency values for conjunctive targets no not necessarily indicate worse performance. Error bars reflect 95% confidence intervals.

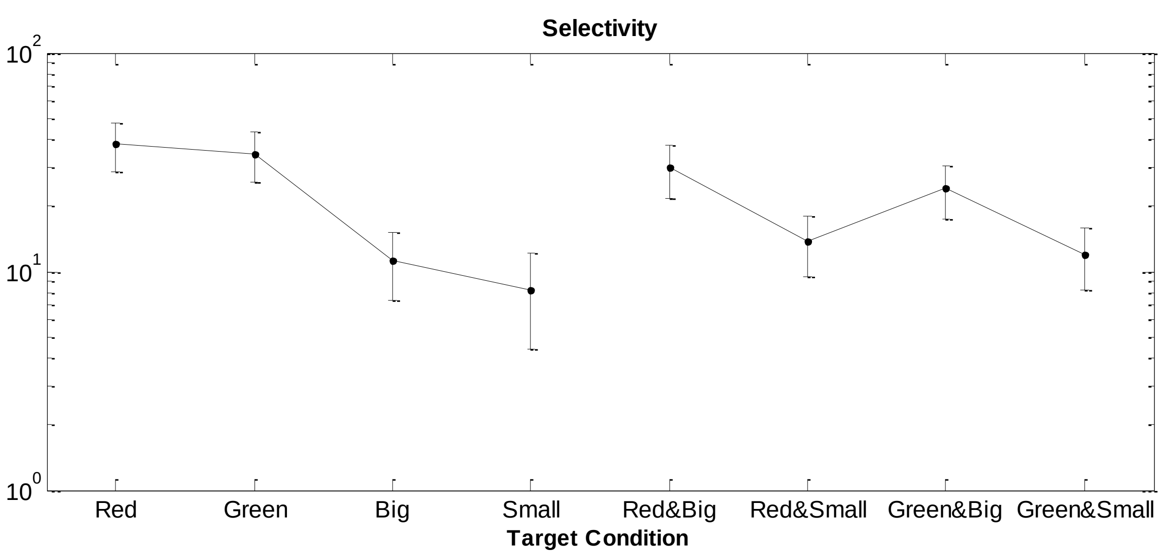


Figure 1.9: Exp. 1 Centroid Selectivity Ratios. The Selectivity Ratio is the mean target weight divided by the mean distractor weight, and offers a more direct way to compare performance on feature target conditions to performance on conjunctive target conditions. The left side of the graph shows participants' average Selectivity Ratios for feature target conditions, and the ride side shows participants' average Selectivity Ratios for conjunctive target conditions. Error bars reflect 95% confidence intervals.

(a) Comparisons between Feature Dimension Levels				
Contrast	$\mathbf{R - G}$	$\mathbf{B - S}$	$\frac{\mathbf{RB + RS}}{\mathbf{2}} - \frac{\mathbf{GB + GS}}{\mathbf{2}}$	$\frac{\mathbf{RB + GB}}{\mathbf{2}} - \frac{\mathbf{RS + GS}}{\mathbf{2}}$
Mean	0.008	0.080	-0.012	0.107
Standard Deviation	0.029	0.046	0.052	0.088
Upper Bound	0.032	0.118	0.031	0.181
Lower Bound	-0.016	0.042	-0.055	0.034
T	0.790	4.947	-0.657	3.444
p	0.455	0.002	0.532	0.011
Bayes Factor	0.434	27.312	0.402	6.158

(b) Comparisons across Feature Dimensions			
Contrast	$\frac{\mathbf{R + G}}{\mathbf{2}} - \frac{\mathbf{B + S}}{\mathbf{2}}$	$\frac{\mathbf{R + G}}{\mathbf{2}} - \frac{\mathbf{RB + RS + GB + GS}}{\mathbf{4}}$	$\frac{\mathbf{B + S}}{\mathbf{2}} - \frac{\mathbf{RB + RS + GB + GS}}{\mathbf{4}}$
Mean	0.120	0.129	0.009
Standard Deviation	0.050	0.090	0.048
Upper Bound	0.162	0.204	0.049
Lower Bound	0.078	0.054	-0.031
T	6.771	4.079	0.534
p	0.000	0.005	0.610
Bayes Factor	124.247	11.871	0.392

Table 1.3: Exp. 1 Paired-Samples T-tests of Centroid Efficiencies. Letters refer to target conditions, with single letters indicating single-feature target conditions (e.g. 'R' = attend-to-red) and two letters indicating conjunctive target conditions (e.g. 'RB' = attend-to-red&big). Larger efficiencies indicate better performance while smaller efficiencies indicate worse performance. **(a)** The first and second columns compare the two color target conditions (R vs. G) and the two size target conditions (B vs. S), respectively. The third column compares the mean of the red conjunction conditions (RB and RS) with the mean of the green conjunction conditions (GB and GS) and, the fourth column compares the mean of big conjunction conditions (RB and GB) with the mean of small conjunction conditions (RS and GS). **(b)** The first column compares the mean of the color conditions with the mean of the size conditions, the second column compares the mean of the color conditions with the mean of the conjunction conditions, and the third column compares the mean of the size conditions with the mean of the conjunction conditions. These comparisons show that efficiencies were highest in the color conditions, next highest in the size conditions, and lowest in the conjunction conditions, though not all these comparisons are statistically significant.

(a) Comparisons between Feature Dimension Levels				
Contrast	$\mathbf{R - G}$	$\mathbf{B - S}$	$\frac{\mathbf{RB + RS}}{\mathbf{2}} - \frac{\mathbf{GB + GS}}{\mathbf{2}}$	$\frac{\mathbf{RB + GB}}{\mathbf{2}} - \frac{\mathbf{RS + GS}}{\mathbf{2}}$
Mean	0.027	0.217	0.060	0.343
Standard Deviation	0.279	0.138	0.196	0.136
Upper Bound	0.260	0.332	0.224	0.456
Lower Bound	-0.206	0.102	-0.104	0.230
T	0.276	4.466	0.871	7.152
p	0.790	0.003	0.413	0.000
Bayes Factor	0.347	17.375	0.457	164.516

(b) Comparisons across Feature Dimensions			
Contrast	$\frac{\mathbf{R + G}}{\mathbf{2}} - \frac{\mathbf{B + S}}{\mathbf{2}}$	$\frac{\mathbf{R + G}}{\mathbf{2}} - \frac{\mathbf{RB + RS + GB + GS}}{\mathbf{4}}$	$\frac{\mathbf{B + S}}{\mathbf{2}} - \frac{\mathbf{RB + RS + GB + GS}}{\mathbf{4}}$
Log10 Mean	0.700	0.325	-0.375
Standard Deviation	0.317	0.179	0.151
Upper Bound	0.965	0.475	-0.248
Lower Bound	0.435	0.175	-0.502
T	6.242	5.122	-7.000
p	0.000	0.001	0.000
Bayes Factor	82.542	32.012	147.280

Table 1.4: Exp. 1 Paired-Samples T-tests of Centroid Selectivity Ratios. Letters refer to target conditions, with single letters indicating single-feature target conditions (e.g. 'R' = attend-to-red) and two letters indicating conjunctive target conditions (e.g. 'RB' = attend-to-red&big). Larger selectivity ratios indicate better performance while smaller selectivity ratios indicate worse performance. **(a)** The first and second columns compare the two color target conditions (R vs. G) and the two size target conditions (B vs. S), respectively. The third column compares the mean of the red conjunction conditions (RB and RS) with the mean of the green conjunction conditions (GB and GS) and, the fourth column compares the mean of big conjunction conditions (RB and GB) with the mean of small conjunction conditions (RS and GS). **(b)** The first column compares the mean of the color conditions with the mean of the size conditions, the second column compares the mean of the color conditions with the mean of the conjunction conditions, and the third column compares the mean of the size conditions with the mean of the conjunction conditions. These comparisons show that selectivity ratios were highest in the color conditions, next highest in the conjunction conditions, and lowest in the size conditions, and all these comparisons are statistically significant.

search-for-small conditions in the visual search tasks, both efficiency and selectivity were higher in the attend-to-red and attend-to-green centroid task conditions than they were in the attend-to-large and attend-to-small conditions, and there was a main effect of size with better performance on attend-to-large conditions than on attend-to-small conditions. In contrast to search results, however, centroid performance (as measured by both efficiency and selectivity) was better in conjunctive target conditions than in the constituent-feature target conditions of attend-to-large and attend-to-small. Although this difference was not significant as measured by efficiency, it should be noted again that efficiency refers to a proportion of the items, and the conjunctive target display had twice as many items as the single-feature target displays.

1.2.3 Conclusions

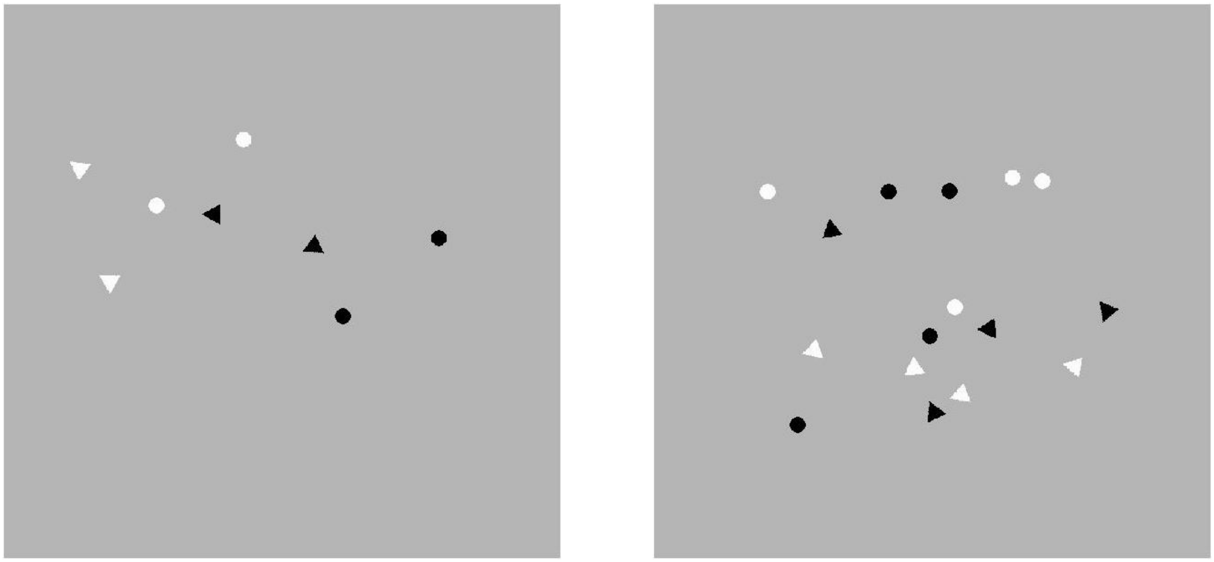
Our visual search results more or less follow the expected pattern given the extensive search literature. It seems targets defined by color are easier to find than targets defined by size, both of which are easier to find that targets defined by the conjunction of color and size. We can describe this pattern with the simple inequality: *Color* > *Size* > *Conjunction*.

Remarkably, the pattern of results is different in the centroid paradigm. Instead of all constituent-feature targets outperforming conjunctive targets, conjunctive targets are intermediate between the two constituent feature dimensions of the conjunction. Stated as an inequality, we find *Color* > *Conjunction* > *Size* in the centroid paradigm. Based on the FIT interpretation of visual search data, one might have predicted that a conjunctive centroid task would be an impossible one. However, it is not only possible, it is actually better than a constituent-feature centroid task, even when the single-feature task has half as many items per display. Given the surprising nature of these results, replication is crucial. We conducted Experiment 2 in order to provide converging evidence using the feature dimensions of luminance and shape.

1.3 EXPERIMENT 2: LUMINANCE-SHAPE CONJUNCTIONS

1.3.1 Methods

As in the first experiment, participants with no prior centroid experienced first completed 500 trails of centroid training. The design of Experiment 2 was essentially identical to that of Experiment 1: practice sessions followed by experiment sessions for all participants. The ABBA task order and Latin squares block order also remained the same. The main difference between the experiments was the feature dimensions: instead of size and color, they were luminance (black/white) and shape (circle/triangle). We chose circles and triangles to achieve a difference in shape comparably dramatic to the difference in luminance between black and white. The triangles were equilateral and appeared at random orientations. We matched the area of the shapes are closely as possible, so the area of both circles and triangles was approximately 400 pixels. The radius of the circle subtended approximately 0.21 degrees of visual angle, and the distance from the center of the triangle to a vertex subtended approximately 0.31 degrees of visual angle.



Feature Target Display Conjunctive Target Display

Figure 1.10: Exp. 2 Centroid Displays. The same centroid display could be used for any of the four single-feature targets (black, white, circle, triangle) or any of the four conjunctive targets (black circle, black triangle, white circle, white triangle). All single-feature conditions had 8 items per display (4 targets, 4 distractors) and all conjunctive conditions had 16 items per display (4 targets, 12 distractors) in order to keep the number of targets constant.

Again, in the centroid task, items were presented in a stimulus display cloud (800 x 800 pixels, 13.75 degrees of visual angle) with a fixed dispersion of one-sixth of the stimulus display size ($133\frac{1}{3}$ pixels, 2.30 degrees of visual angle). The background gray (47.63 cd/m^2) was darker than in Experiment 1 so that it its luminance would be almost exactly in between the white (107.6 cd/m^2) and black (0.28 cd/m^2) items. The single-feature conditions had eight items per display (two of each of the four item types) and the conjunction conditions had 16 (four of each of the four item types), so there were always four targets per display. Figure 1.10 shows examples of these centroid stimulus displays. The stimulus display appeared for 250 ms, the blank screen for 50 ms, and the visual mask for 100 ms, after which participants entered their responses and received feedback. There

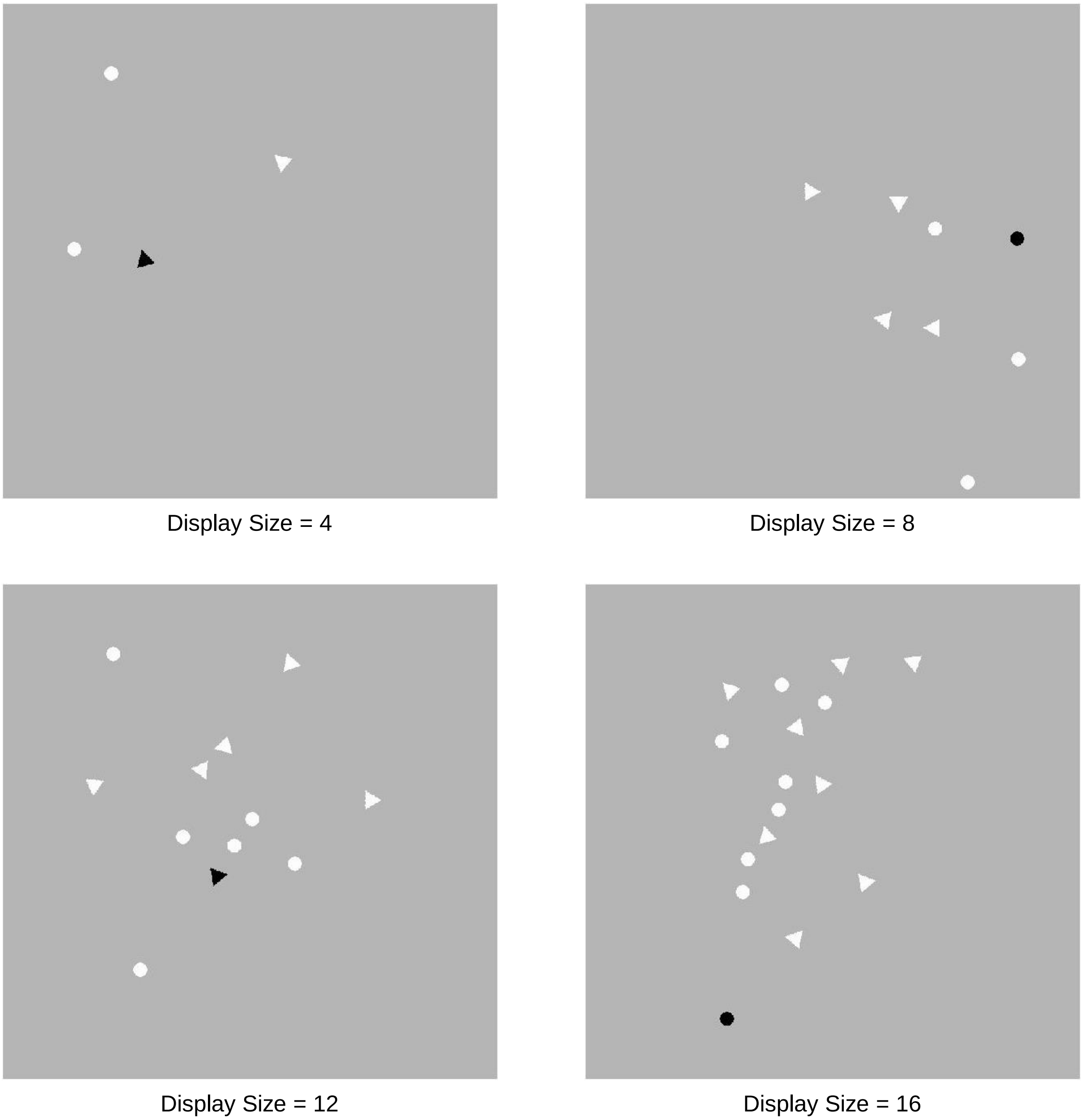


Figure 1.11: Exp. 2 Visual Search Displays for Feature Targets. The target in all four example displays is a black item, making luminance the relevant feature dimension and shape the irrelevant feature dimension. The target item could take on either level of the irrelevant feature dimension; here, the black target could be either a circle or a triangle.

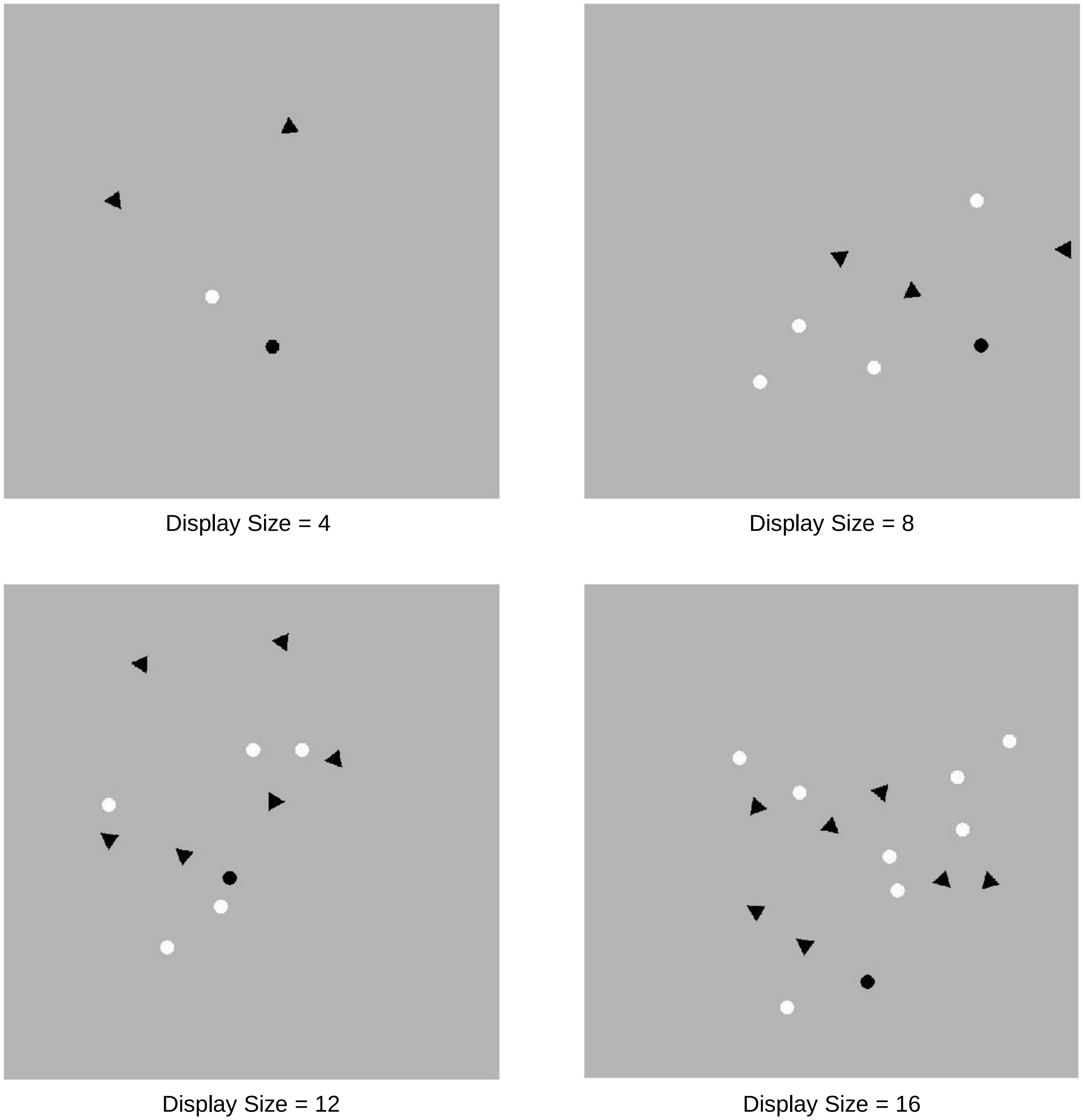


Figure 1.12: Exp. 2 Visual Search Displays for Conjunctive Targets. The target in all four example displays is a black circle. Both luminance and shape are relevant feature dimensions.

were 100 trials per block, eight blocks per session, and two sessions for a total of 1,600 centroid trials per participant (n=8).

The visual search arrays were also created the same way as in Experiment 1. We used the same process to assign item locations with the same fixed dispersion. Display size was 4, 8, 12, or 16 items; half the trials were positive and the other half were negative. When the target was defined by a single feature, the levels of the irrelevant feature dimensions appeared in equal numbers (see Figure 1.11). When the target was defined by a conjunction of features, both feature dimensions were relevant; so, on positive trials, one item was selected at random and modified to become the target (see Figure 1.12).

The task was to indicate, on each trial, whether or not the target was present. The participant pressed the *Z* key with her left hand to enter a “no” response or the *M* key with her right hand to enter a “yes” response. The search array remained on the screen until the participant entered her response, at which point she received visual correctness feedback. There were 160 trials per block, eight blocks per session, and two sessions for a total of 2,560 visual search trials per participant.

We originally collected data from nine participants, but one appeared not to understand the centroid task. This participant had a mean Efficiency of 0.3473 and Data-Drivenness of 0.4998, suggesting a strategy of clicking at random. We excluded this participant from future analyses and present data from the remaining eight participants in the following Results section.

1.3.2 Results

Again, we analyzed our visual search data by taking each participant's median RT on correct trials, then finding the mean for each target condition, display size, and trial type (positive/negative). We plot these means and the slopes of the best fitting lines for each condition and trial type (Figure 1.13) and performed a series of paired-samples t-tests on the search slopes as described in Tables 1.5 and 1.6. The search slopes were flatter in the search-for-black and search-for-white conditions than in the search-for-circles and search-for-triangles conditions. In addition, the single-feature target conditions had flatter slopes than the conjunctive target conditions.

For the centroid data, we calculated the attention filter for each item type by target condition using the Sun et al. (2016) methodology described earlier. The relative influence of each item type averaged across participants is shown in Figure 1.14. We also calculated the efficiencies (Figure 1.15) and selectivity ratios (Figure 1.16) for each target condition. Efficiency tracks error in that it provides an estimate of the number of items an ideal observer with the same attentional filter as participant would need to process in order to perform with the same error as that participant. In this way, it provides a lower bound on the number of items processed. The selectivity ratio is the average of the absolute values of the target item weights, divided by the average of the absolute values of the distractor item weights. This allows us to compare performance across feature target conditions (in which two item types are targets) to conjunction target conditions (in which only one item type is a target).

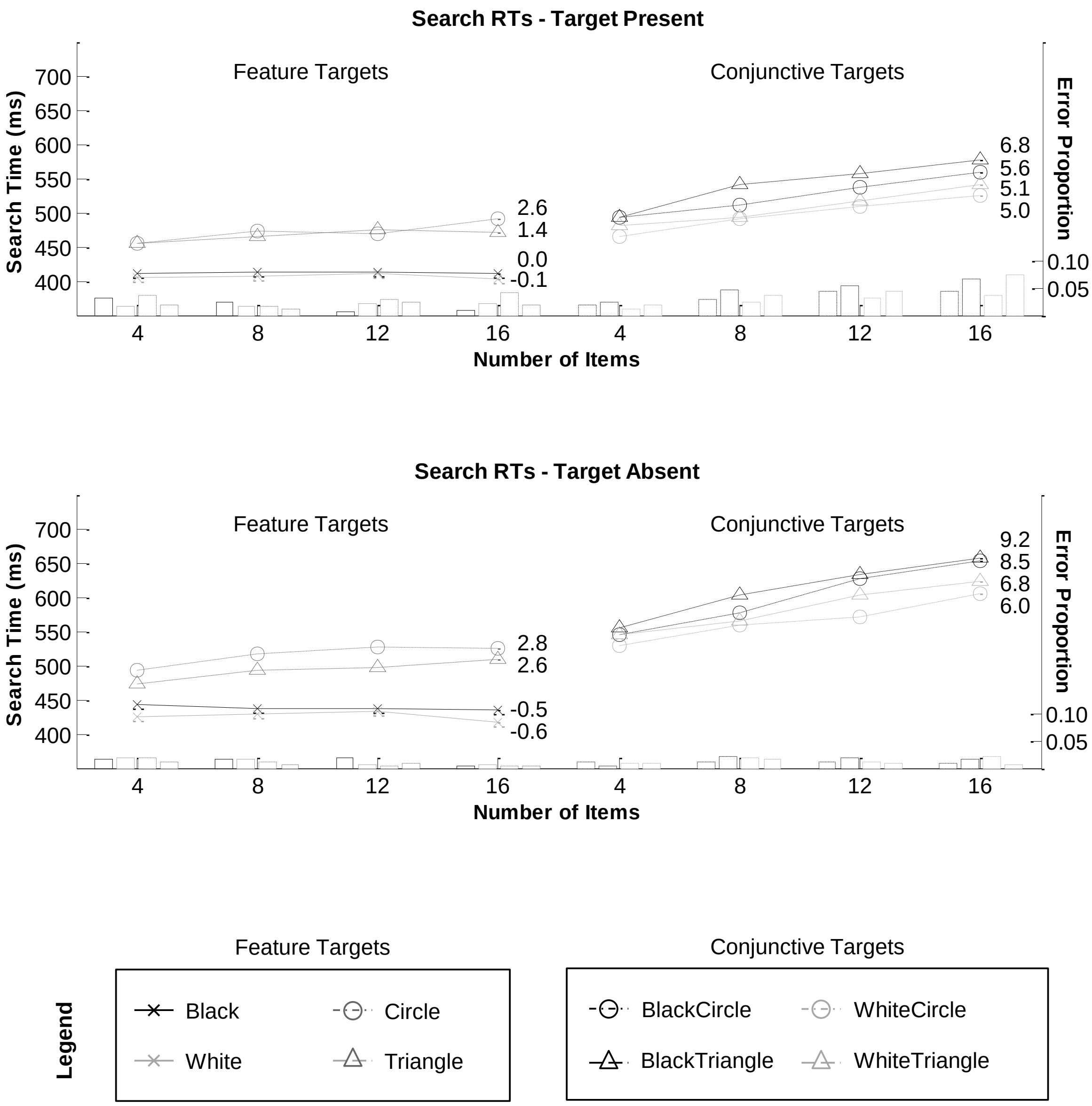


Figure 1.13: Exp. 2 Visual Search Reaction Times. The lines plot the mean of participants' median RTs for each target condition. . The number the right of each line gives its slope. Overall, we found flatter RT slopes for single-feature conditions than for conjunction conditions. Notably, however, the RTs were longer and the slopes steeper for single-feature shape conditions compared to the single-feature luminance conditions. The bars show the error proportion, which is the number of incorrect trials divided by the total number of trials in a condition. Incorrect trials were excluded from participants' median RTs.

(a) Comparisons between Feature Dimension Levels				
Contrast	$\mathbf{B - W}$	$\mathbf{C - T}$	$\mathbf{\frac{BC + BT}{2} - \frac{WC + WT}{2}}$	$\mathbf{\frac{BC + WC}{2} - \frac{BT + WT}{2}}$
Mean	0.107	1.226	1.133	-0.682
Standard Deviation	1.955	3.058	2.458	2.272
Upper Bound	1.742	3.783	3.188	1.217
Lower Bound	-1.527	-1.330	-0.923	-2.582
T	0.155	1.134	1.303	-0.850
p	0.881	0.294	0.234	0.424
Bayes Factor	0.340	0.555	0.641	0.451

(b) Comparisons across Feature Dimensions			
Contrast	$\mathbf{\frac{B + W}{2} - \frac{C + T}{2}}$	$\mathbf{\frac{B + W}{2} - \frac{BC + BT + WC + WT}{4}}$	$\mathbf{\frac{C + T}{2} - \frac{BC + BT + WC + WT}{4}}$
Mean	-2.087	-5.672	-3.584
Standard Deviation	1.730	2.067	1.812
Upper Bound	-0.641	-3.944	-2.069
Lower Bound	-3.534	-7.400	-5.099
T	-3.412	-7.761	-5.595
p	0.011	0.000	0.001
Bayes Factor	5.952	252.052	48.441

Table 1.5: Exp. 2 Paired-Samples T-tests of Search Slopes for Positive Trials. Letters refer to target conditions, with single letters indicating single-feature target conditions (e.g. 'B' = search-for-black) and two letters indicating conjunctive target conditions (e.g. 'BC' = search-for-black&circle). Flatter (smaller) slopes indicate better performance while steeper (greater) slopes indicate worse performance. **(a)** The first and second columns compare the two luminance target conditions (B vs. W) and the two size target conditions (C vs. T), respectively. The third column compares the mean of the black conjunction conditions (BC and BT) with the mean of the white conjunction conditions (WC and WT), and the fourth column compares the mean of circle conjunction conditions (BC and WC) with the mean of triangle conjunction conditions (BT and WT). **(b)** The first column compares the mean of the luminance conditions with the mean of the shape conditions, the second column compares the mean of the luminance conditions with the mean of the conjunction conditions, and the third column compares the mean of the shape conditions with the mean of the conjunction conditions. These comparisons show that search slopes were flattest in the luminance conditions, next flattest in the shape conditions, and steepest in the conjunction conditions, and all these comparisons are statistically significant.

(a) Comparisons between Feature Dimension Levels				
Contrast	$\mathbf{B-W}$	$\mathbf{C-T}$	$\frac{\mathbf{BC+BT}}{\mathbf{2}}-\frac{\mathbf{WC+WT}}{\mathbf{2}}$	$\frac{\mathbf{BC+WC}}{\mathbf{2}}-\frac{\mathbf{BT+WT}}{\mathbf{2}}$
Mean	-0.061	-0.207	2.454	-0.018
Standard Deviation	2.231	2.280	2.081	4.994
Upper Bound	1.804	1.699	4.194	4.157
Lower Bound	-1.926	-2.113	0.714	-4.194
T	-0.077	-0.257	3.335	-0.010
p	0.941	0.805	0.012	0.992
Bayes Factor	0.337	0.346	5.483	0.336

(b) Comparisons across Feature Dimensions			
Contrast	$\frac{\mathbf{B+W}}{\mathbf{2}}-\frac{\mathbf{C+T}}{\mathbf{2}}$	$\frac{\mathbf{B+W}}{\mathbf{2}}-\frac{\mathbf{BC+BT+WC+WT}}{\mathbf{4}}$	$\frac{\mathbf{C+T}}{\mathbf{2}}-\frac{\mathbf{BC+BT+WC+WT}}{\mathbf{4}}$
Mean	-3.298	-8.208	-4.909
Standard Deviation	2.501	2.650	3.132
Upper Bound	-1.207	-5.992	-2.291
Lower Bound	-5.389	-10.424	-7.528
T	-3.730	-8.759	-4.433
p	0.007	0.000	0.003
Bayes Factor	8.314	481.024	16.830

Table 1.6: Exp. 2 Paired-Samples T-tests of Search Slopes for Negative Trials. Letters refer to target conditions, with single letters indicating single-feature target conditions (e.g. 'B' = search-for-black) and two letters indicating conjunctive target conditions (e.g. 'BC' = search-for-black&circle). Flatter (smaller) slopes indicate better performance while steeper (greater) slopes indicate worse performance. **(a)** The first and second columns compare the two luminance target conditions (B vs. W) and the two size target conditions (C vs. T), respectively. The third column compares the mean of the black conjunction conditions (BC and BT) with the mean of the white conjunction conditions (WC and WT), and the fourth column compares the mean of circle conjunction conditions (BC and WC) with the mean of triangle conjunction conditions (BT and WT). **(b)** The first column compares the mean of the luminance conditions with the mean of the shape conditions, the second column compares the mean of the luminance conditions with the mean of the conjunction conditions, and the third column compares the mean of the shape conditions with the mean of the conjunction conditions. These comparisons show that search slopes were flattest in the luminance conditions, next flattest in the shape conditions, and steepest in the conjunction conditions, and all these comparisons are statistically significant.

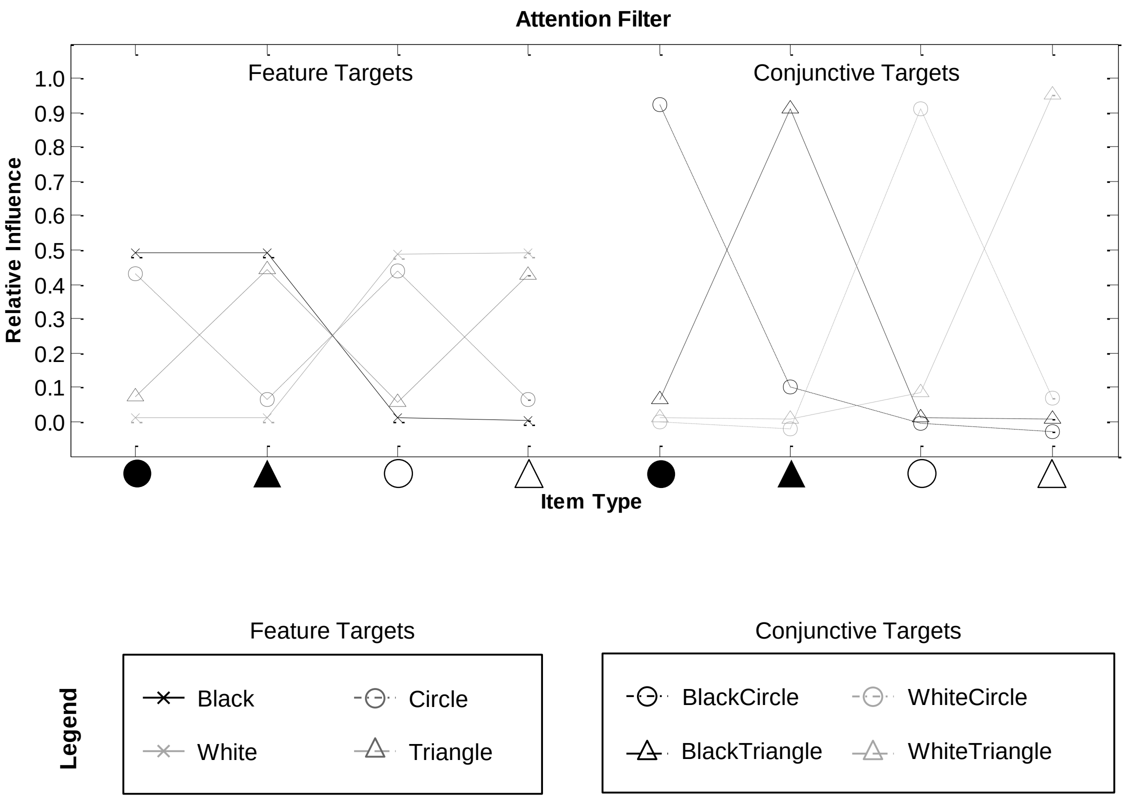


Figure 1.14: Exp. 2 Centroid Attention Filters. Each line shows the relative influence of each item type (averaged across eight participants) for a particular attention condition. The relative influence, or weight, of all the item types sums to 1. For all feature target conditions, the targets consist of two item types. The ideal attentional filter would be equally influenced by the two target item types (assigning them each a weight of 0.5) and not at all influenced by the two distractor item types (assigning them each a weigh of 0). Participants' actual performance follows these trends. For example, in the attend-to-black condition, the two black items (targets) both have weights of about 0.5 while the two white items (distractors) both have weights of about 0. For all conjunctive target conditions, the target consists of only one item type. The ideal attentional filter would be influenced only by the target item (assigning it a weight of 1), and not at all influenced by the three distractor item types (assigning them each a weigh of 0). Again, participants' actual performance is not far off. For example, in the attend-to-black&circle target condition, the black circle (target) has a weight of about 0.9, while the black triangle (distractor) has a weight of about 0.1 and the remaining white item types have a weight of about 0. In each conjunctive target condition, the distractor that shares the target's luminance exerts more influence on participants' centroid judgments compared to the other two distractor item types.

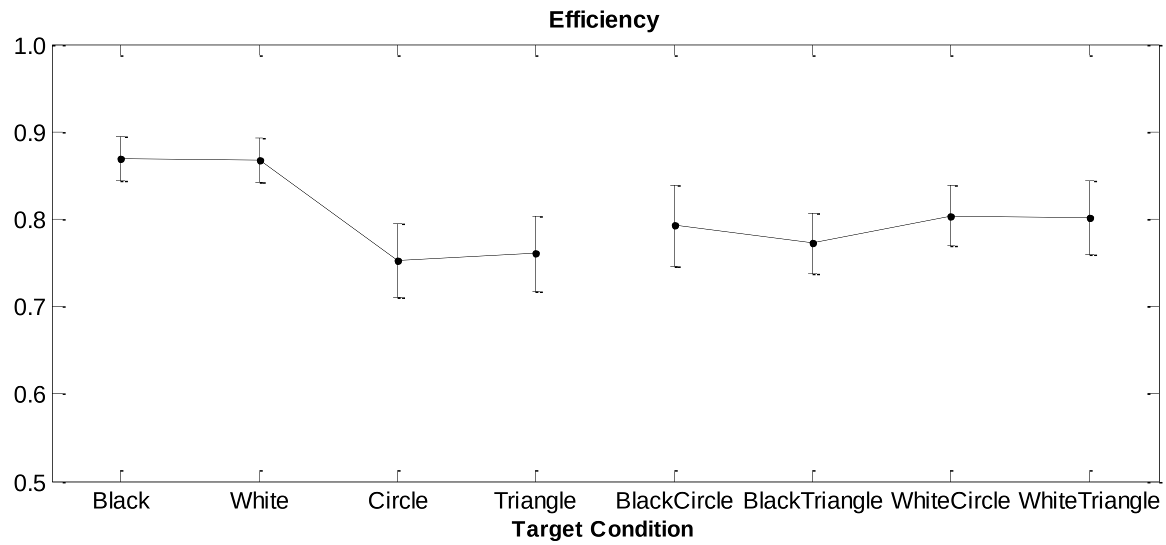


Figure 1.15: Exp. 2 Centroid Efficiencies. Efficiency estimates a lower bound on the proportion of items participant observed. The left side of the graph shows participants' average Efficiencies for feature target conditions, and the ride side shows participants' average Efficiencies for conjunctive target conditions. There were twice as many items in the conjunctive target conditions as in the feature target conditions, so lower Efficiency values for conjunctive targets no not necessarily indicate worse performance. Error bars reflect 95% confidence intervals.

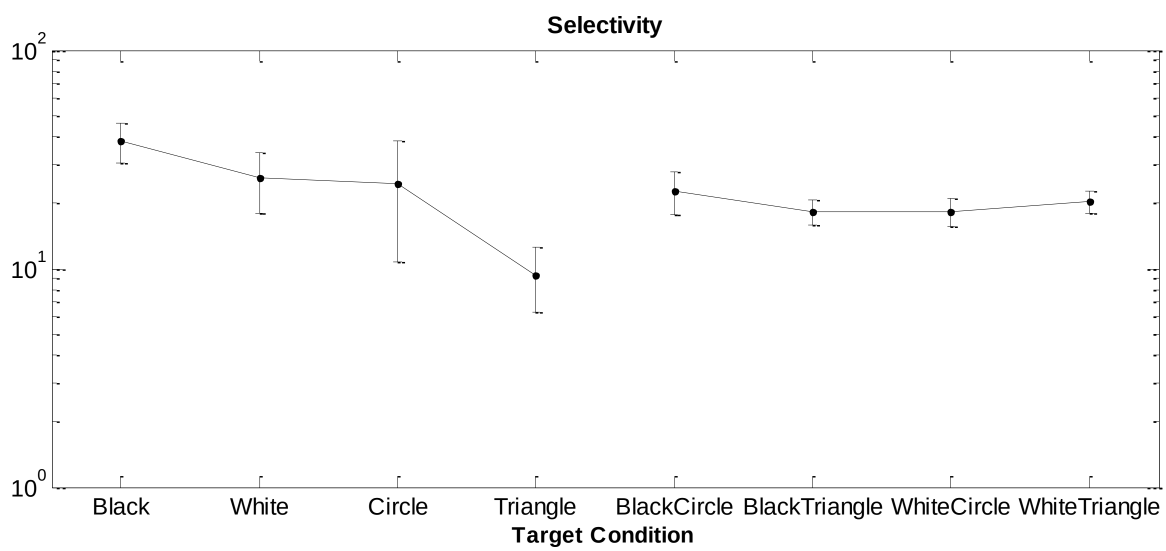


Figure 1.16: Exp. 2 Centroid Selectivity Ratios. The Selectivity Ratio is the mean target weight divided by the mean distractor weight, and offers a more direct way to compare performance on feature target conditions to performance on conjunctive target conditions. The left side of the graph shows participants' average Selectivity Ratios for feature target conditions, and the ride side shows participants' average Selectivity Ratios for conjunctive target conditions. Error bars reflect 95% confidence intervals.

(a) Comparisons between Feature Dimension Levels				
Contrast	$\mathbf{B - W}$	$\mathbf{C - T}$	$\frac{\mathbf{BC + BT}}{\mathbf{2}} - \frac{\mathbf{WC + WT}}{\mathbf{2}}$	$\frac{\mathbf{BC + WC}}{\mathbf{2}} - \frac{\mathbf{BT + WT}}{\mathbf{2}}$
Mean	0.001	-0.008	-0.020	0.011
Standard Deviation	0.015	0.036	0.018	0.018
Upper Bound	0.014	0.022	-0.005	0.027
Lower Bound	-0.011	-0.038	-0.035	-0.004
T	0.210	-0.659	-3.189	1.713
p	0.840	0.531	0.015	0.131
Bayes Factor	0.343	0.402	4.686	0.948

(b) Comparisons across Feature Dimensions			
Contrast	$\frac{\mathbf{B + W}}{\mathbf{2}} - \frac{\mathbf{C + T}}{\mathbf{2}}$	$\frac{\mathbf{B + W}}{\mathbf{2}} - \frac{\mathbf{BC + BT + WC + WT}}{\mathbf{4}}$	$\frac{\mathbf{C + T}}{\mathbf{2}} - \frac{\mathbf{BC + BT + WC + WT}}{\mathbf{4}}$
Mean	0.112	0.075	-0.036
Standard Deviation	0.056	0.050	0.025
Upper Bound	0.158	0.117	-0.015
Lower Bound	0.065	0.033	-0.057
T	5.656	4.250	-4.021
p	0.001	0.004	0.005
Bayes Factor	51.020	14.074	11.198

Table 1.7: Exp. 2 Paired-Samples T-tests of Centroid Efficiencies. Letters refer to target conditions, with single letters indicating single-feature target conditions (e.g. 'B' = search-for-black) and two letters indicating conjunctive target conditions (e.g. 'BC' = search-for-black&circle). Larger efficiency values indicate better performance while smaller efficiency values indicate worse performance. **(a)** The first and second columns compare the two luminance target conditions (B vs. W) and the two size target conditions (C vs. T), respectively. The third column compares the mean of the black conjunction conditions (BC and BT) with the mean of the white conjunction conditions (WC and WT), and the fourth column compares the mean of circle conjunction conditions (BC and WC) with the mean of triangle conjunction conditions (BT and WT). **(b)** The first column compares the mean of the luminance conditions with the mean of the shape conditions, the second column compares the mean of the luminance conditions with the mean of the conjunction conditions, and the third column compares the mean of the shape conditions with the mean of the conjunction conditions. These comparisons show that efficiencies were highest in the luminance conditions, next highest in the conjunction conditions, and lowest in the shape conditions, and all these comparisons are statistically significant.

(a) Comparisons between Feature Dimension Levels				
Contrast	$\mathbf{B - W}$	$\mathbf{C - T}$	$\frac{\mathbf{BC + BT}}{\mathbf{2}} - \frac{\mathbf{WC + WT}}{\mathbf{2}}$	$\frac{\mathbf{BC + WC}}{\mathbf{2}} - \frac{\mathbf{BT + WT}}{\mathbf{2}}$
Mean	0.207	0.145	0.009	-0.002
Standard Deviation	0.128	0.254	0.078	0.136
Upper Bound	0.314	0.358	0.074	0.112
Lower Bound	0.100	-0.067	-0.056	-0.116
T	4.586	1.619	0.332	-0.045
p	0.003	0.149	0.750	0.965
Bayes Factor	19.493	0.863	0.352	0.337

(b) Comparisons across Feature Dimensions			
Contrast	$\frac{\mathbf{B + W}}{\mathbf{2}} - \frac{\mathbf{C + T}}{\mathbf{2}}$	$\frac{\mathbf{B + W}}{\mathbf{2}} - \frac{\mathbf{BC + BT + WC + WT}}{\mathbf{4}}$	$\frac{\mathbf{C + T}}{\mathbf{2}} - \frac{\mathbf{BC + BT + WC + WT}}{\mathbf{4}}$
Log10 Mean	0.512	0.166	-0.346
Standard Deviation	0.494	0.269	0.336
Upper Bound	0.924	0.391	-0.065
Lower Bound	0.099	-0.059	-0.626
T	2.933	1.745	-2.914
p	0.022	0.124	0.023
Bayes Factor	3.547	0.979	3.474

Table 1.8: Exp. 2 Paired-Samples T-tests of Centroid Selectivity Ratios. Letters refer to target conditions, with single letters indicating single-feature target conditions (e.g. 'B' = search-for-black) and two letters indicating conjunctive target conditions (e.g. 'BC' = search-for-black&circle). Larger selectivity ratios indicate better performance while smaller selectivity ratios indicate worse performance. **(a)** The first and second columns compare the two luminance target conditions (B vs. W) and the two size target conditions (C vs. T), respectively. The third column compares the mean of the black conjunction conditions (BC and BT) with the mean of the white conjunction conditions (WC and WT), and the fourth column compares the mean of circle conjunction conditions (BC and WC) with the mean of triangle conjunction conditions (BT and WT). **(b)** The first column compares the mean of the luminance conditions with the mean of the shape conditions, the second column compares the mean of the luminance conditions with the mean of the conjunction conditions, and the third column compares the mean of the shape conditions with the mean of the conjunction conditions. These comparisons show that selectivity ratios were highest in the luminance conditions, next highest in the conjunction conditions, and lowest in the shape conditions, though not all these comparisons are statistically significant.

We conducted paired-samples t-tests (the same as the as we conducted on our search slope data) on efficiency (Table 1.7) and selectivity (Table 1.8). Considering both these measures, we find that centroid performance is better in attend-to-black and attend-to-white conditions than in attend-to-circles and attend-to-triangles conditions. Additionally, performance on the conjunctive target conditions is better than performance on the attend-to-circles and attend-to-triangles conditions. There were also several cases in which black targets outperformed white targets.

1.2.3 Conclusions

Again, our search results follow the expected pattern in which single-feature targets always outperform conjunctive targets. Expressed as an inequality, we find *Luminance* > *Size* > *Conjunction*.

Our centroid results, on the other hand, deviate from this expected pattern. Instead, we find *Luminance* > *Conjunction* > *Size*. Interestingly, this pattern is more compelling in the context of our efficiency data than it is in the context or out selectivity data. However, it is worth noting that all the selectivity ratios in the Experiment 2 were quite high, suggesting there may be a ceiling effect.

1.4 GENERAL DISCUSSION

We draw two main conclusions from our present findings.

(1) In visual search tasks, performance (as measured by reaction time slope) is better on all feature conditions compared to all conjunction conditions. In Experiment 1, reaction time slopes were flattest when participants searched for color targets. The next flattest slopes were observed for size targets, and the steepest slopes for color-size conjunctive targets. In Experiment 2, reaction time slopes were flattest when participants searched for luminance targets. The next flattest slopes were observed for shape targets, and the steepest slopes for luminance-shape conjunctive targets.
(2) In centroid tasks, performance (as measured by selectivity ratio) follows a different pattern: it is possible for performance on conjunction conditions to *exceed* performance on feature conditions. In Experiment 1, the attend-to-color conditions produced the greatest selectivity ratios, followed by the attend-to-conjunctions conditions, while the attend-to-size conditions produced the lowest. In Experiment 2, the attend-to-luminance conditions produced the greatest selectivity ratios, followed by the attend-to-conjunctions conditions, while the attend-to-shape conditions produced the lowest.

Our first conclusion, on its own, poses no challenge to FIT (or any current FBA theory, for that matter). Rather, these patterns of results precisely match the FIT's predictions about feature versus conjunctive targets. However, our second conclusion, in which performance on conjunction conditions exceeds performance on constituent-feature conditions in the centroid task, directly contradicts FIT's predictions. In previous visual search studies, there were a handful of cases in which performance on conjunction conditions was comparable to performance on single-feature conditions (e.g., Nakayama & Silverman, 1986; Theeuwes & Kooi, 1994). Here, however, we have evidence not just of comparable performance, but improved performance in conjunction conditions. These results motivate a reexamination of FIT and of the task demands of visual search. For instance, it is possible that particiapnts are relying on feature contrast rather than the features themselves. When there is only one target in a display and it is defined by a single feature, the contrast between the target and distractors is often highly salient. However, the conjunctive targets remove the contrast clue, which could explain poorer performance for conjunction conditions in visual search. By using multiple targets in the centroid paradigm, we can rule out the possibility that participants are merely relying on feature contrast, rather than the features themselves, to perform the task. (For more on how the centroid paradigm can eliminate feature contrast clues, see Inverso, Sun, Chubb, Wright, & Sperling, 2016.)

Our centroid conjunction results also call into question the role of attention in binding features together. Given the relatively shallow search slopes we found even for conjunctive targets, participants should have had enough time to scan every item in centroid conjunction displays and bind together the relevant features, according to FIT. However, the shallow slopes would also suggest that they would have more than enough to scan every individual item in the centroid feature displays, even though they should not need to in order to perform the attend-to-size and attend-to-shape feature tasks. What, then, explains the improved performance on conjunction conditions in the centroid task? One possibility is that there are obligatory filters for certain features in certain contexts. Color, for example, could be one such feature. The display could be segmented into red and green items whether the viewer wants it segmented this way or not. When the task requires her to filter based on both color and shape, half the work is already done for her, thanks to the obligatory filter. When she must filter by color alone, her performance is superb. But when the task requires her to filter only by shape, the (task-irrelevant) obligatory color filter gets in the way, impairing performance. This explanation, however, cannot yet explain our visual search results—in which there appears to be no such interference from task-irrelevant obligatory filters—or what is relevantly different about the "context" of search versus centroid tasks.

Another possible explanation of our results is the Guided Search model (described in its most current form in Wolfe, 2007). The Guided Search (GS) model proposes two stages of attention in which output from the first stage can guide attention in the second stage, after passing through a bottleneck of processing. The first stage can process basic features of the visual scene in parallel. Through feature channels tuned to particular properties (such as the color red), it can locate particular regions to be processed in more detail in the second stage. These regions are represented in the form of "peaks" in activation maps. The second stage employs an asynchronous diffusion model to process up to a few items at a time, making a decision about whether each item is a target or a distractor. One possibility is that, in our centroid tasks, the stimulus clouds were displayed long enough to allow for a guided search in the centroid conditions. However, a guided search in which the first stage identifies the locations of all red items in a red-big conjunction task (for instance), cannot fully explain improved performance on conjunction conditions relative to the size feature conditions. At best, the guided search would limit the second stage to processing only eight items—the same number of items in the feature centroid displays. Still, we addressed this concern by repeating the centroid tasks of Experiments 1 and 2 using a short display time and more items per cloud for one participant. The stimulus cloud was now displayed for only 180 ms, followed by a blank for 10 ms screen and a visual mask for 20 ms. The number

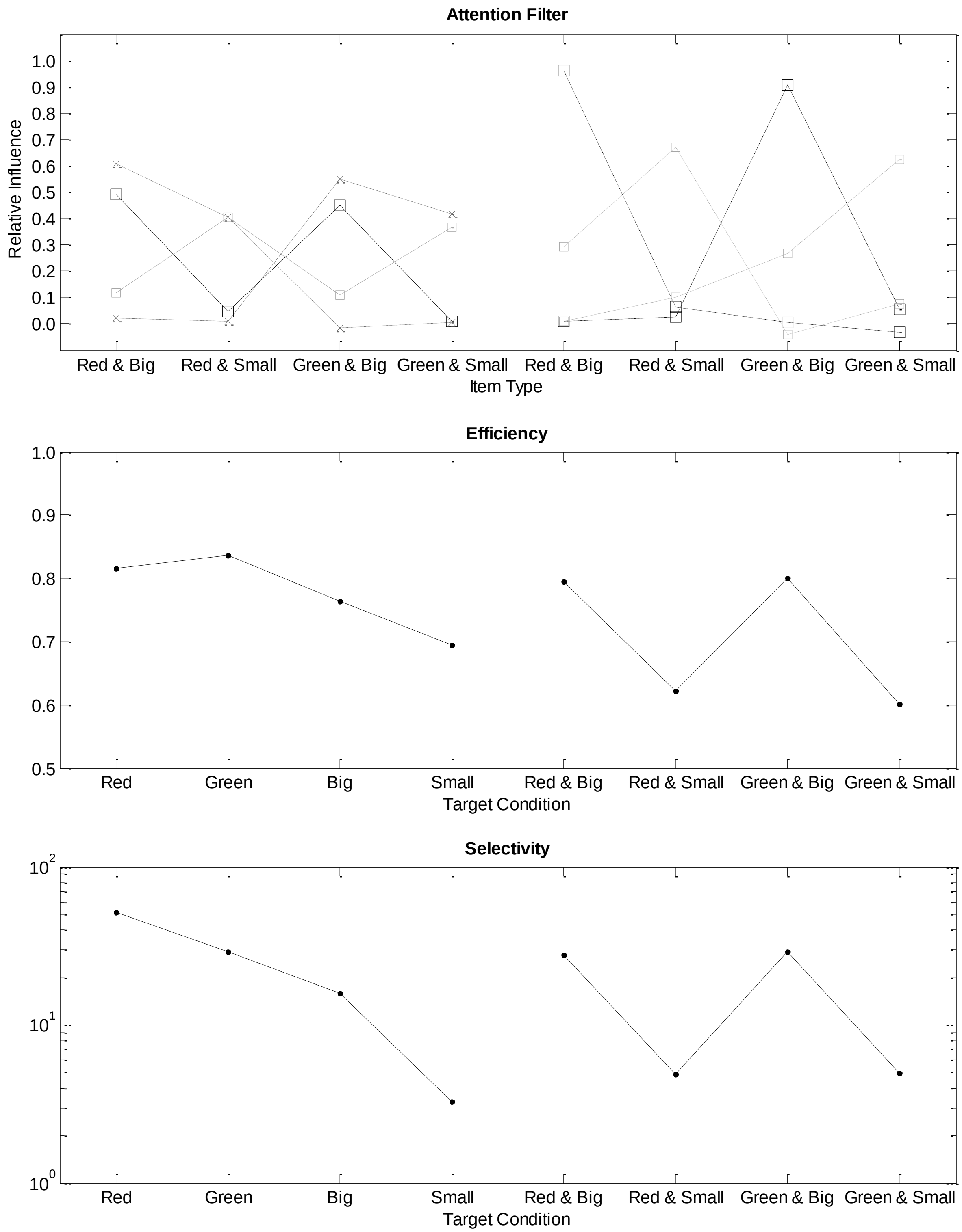


Figure 1.17: Replicated Exp. 1 Centroid Results. One participant repeated the centroid task with a shorter display time and more items per display. The basic pattern of results from Exp. 1 remains more or less intact, although the one participant's idiosyncrasies are more noticeable.

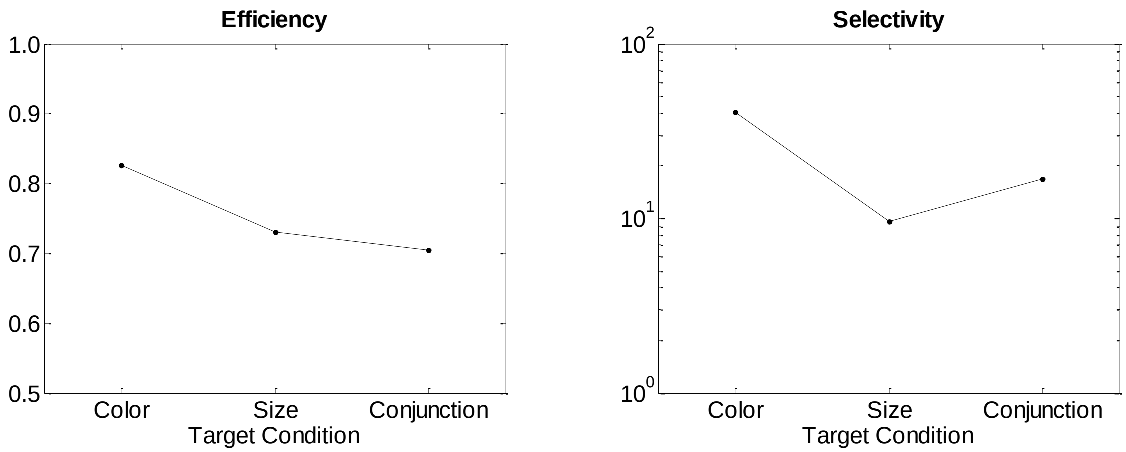


Figure 1.18: Replicated Exp. 1 Centroid Results (Aggregated). These are the same Efficiencies and Selectivity Ratios shown in Figure 1.17, but aggregated by relevant feature dimension(s). 'Color' is the average of the red and green target conditions, 'Size' is the average of the big and small target conditions, and 'Conjunction' is the average of all four conjunction conditions. While Efficiency is slightly lower in the Conjunction conditions than in the Size conditions, there were twice as many items in the former. Selectivity in the Conjunction condition is worse than in the Color condition, but better than in the Size condition.

of items was increased to 12 (three times each of the four item types) items in feature conditions, and to 24 (six times the number of item types) items in conjunction conditions, so each condition had 6 targets. We also reduced the size of each item in order to fit the additional items in the cloud. Figure 1.17 shows the participant's complete results from the Experiment 1 replication, while Figure 1.18 shows efficiency and selectivity aggregated over the attend-to-color, attend-to-size, and attend-to-conjunctions conditions. Similarly, Figure 1.19 shows the participant's complete results from the Experiment 2 replication, and Figure 1.20 shows efficiency and selectivity aggregated over the attend-to-luminance, attend-to-shape, and attend-to-conjunctions conditions. Compared to the results of Experiments 1 and 2, these results are certainly noisier, but the basic pattern is more or less the same. That is, we still find evidence of improved performance on conjunctive target conditions relative to some constituent-feature target conditions.

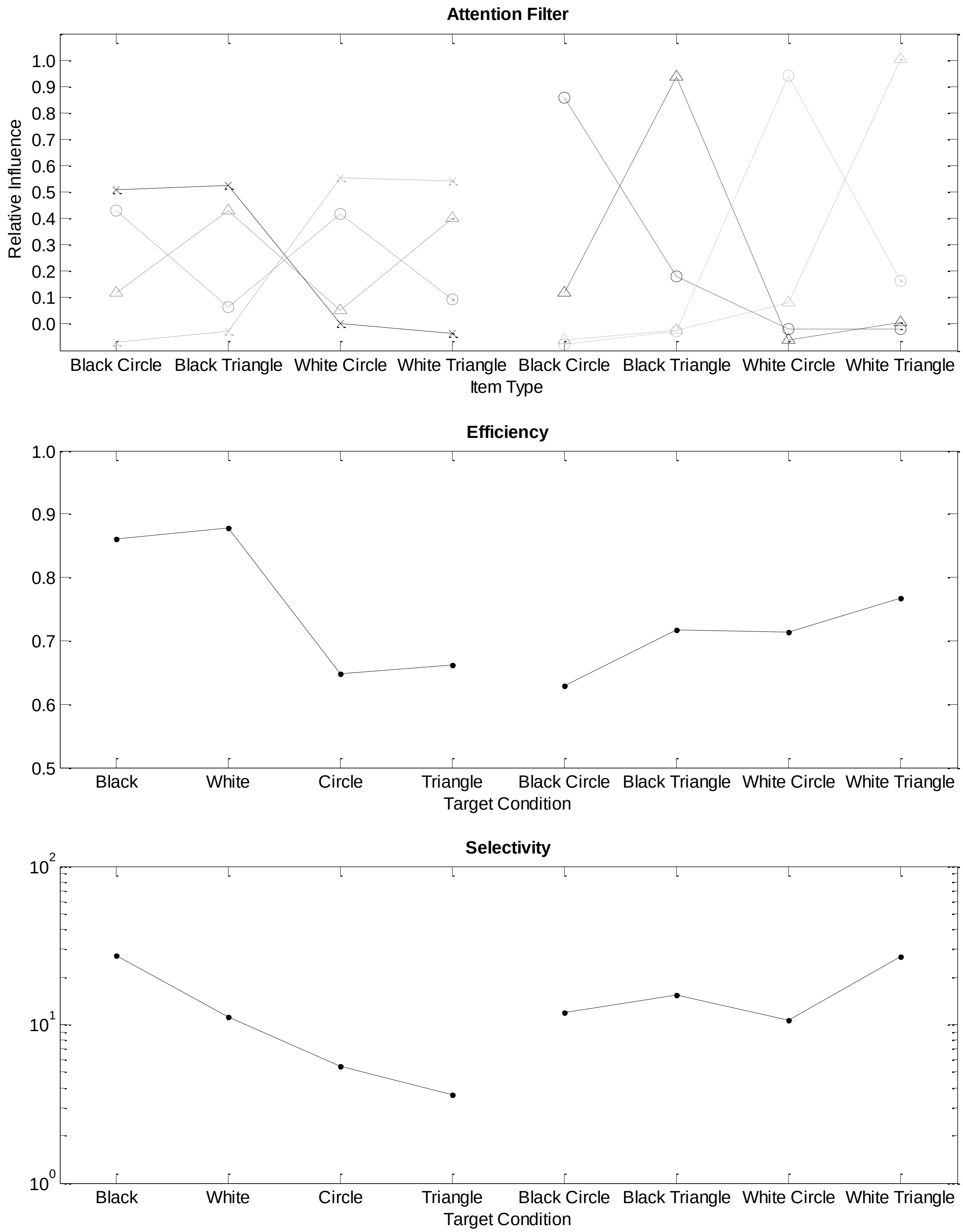


Figure 1.19: Replicated Exp. 2 Centroid Results. One participant repeated the centroid task with a shorter display time and more items per display. The basic pattern of results from Exp. 2 remains more or less intact, although the one participant's idiosyncrasies are more noticeable.

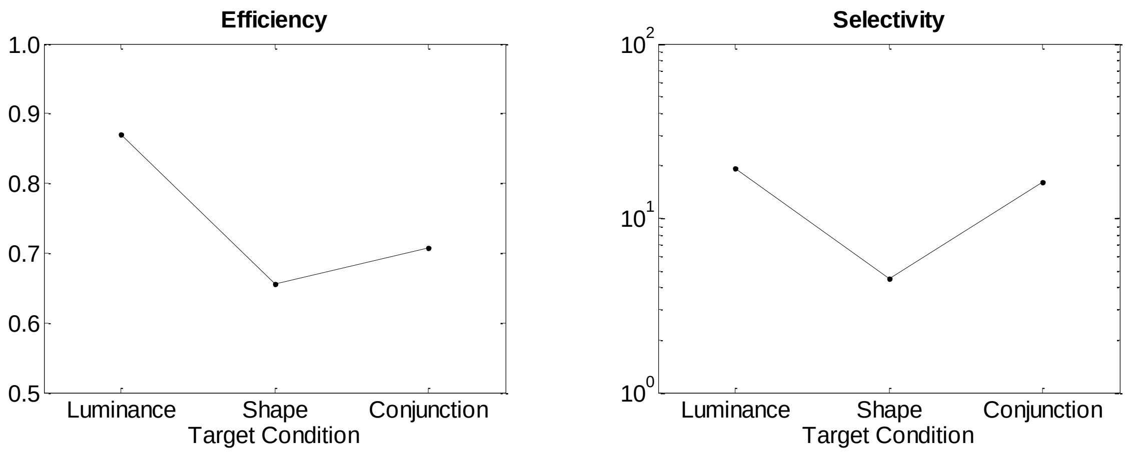


Figure 1.20: Replicated Exp. 2 Centroid Results (Aggregated). These are the same Efficiencies and Selectivity Ratios shown in Figure 1.19, but aggregated by relevant feature dimension(s). 'Luminance' is the average of the black and white target conditions, 'Shape' is the average of the circle and triangle target conditions, and 'Conjunction' is the average of all four conjunction conditions. Both Efficiency and Selectivity in the Conjunction condition are worse than in Luminance condition but better than in Shape condition.

Another difficulty for the GS model is how to explain the differences in performance on the visual search and centroid tasks. If guided search can improve performance on conjunction conditions relative to some feature conditions in the centroid task, why are those same conjunction conditions not easier in the visual search task? One possibility is that the second stage in GS accumulates information at different rates for the different tasks. However, this naturally leads to the question of *why* these rates should differ when the targets are defined in exactly the same way (so the same top-down mechanisms could be use) and the displays were are similar as the tasks the would allow.

Finally, it is worth noting that the centroid results are not merely explained by the difficulty of disjunctive targets, since *all* the feature targets used in both Experiment 1 and Experiment 2 were disjuncts. If disjunctive targets were that much more difficult, we would expect conjunctive targets to consistently out-perform disjunctive ones. Instead, we found that all feature (disjunctive) targets were easier to find than all conjunctive targets in visual search, and that conjunction target conditions were easier than some feature (disjunctive) target conditions but harder than other feature (disjunctive) target conditions in the centroid task.

This reordering of performance by task is truly the most interesting—and difficult to explain—finding we present. Existing accounts of feature-based attention can explain either the search results or the centroid results, but none, to our knowledge, can explain both. One reason for this difficulty is the ubiquitous use of visual search in the FBA literature. With one task dominating the current research, it is unclear how to separate out the peculiarities of visual search task demands from inalienable properties of feature-based attention. We hope to motivate a reexamination of existing FBA accounts using converging evidence from multiple tasks, and we propose the centroid paradigm as one viable supplement to visual search.

CHAPTER 2

EXPERIMENT 3: TARGET-DISTRACTOR SIMILARITY

2.1 INTRODUCTION

In visual search, “feature targets” (e.g., a red square among green squares) exhibit a “pop-out” effect, making them easy to detect irrespective of the number of distractors in the display. By contrast, conjunctive targets (e.g., a red square among green squares and red disks) lose the “pop-out” effect and become increasingly difficult to detect as display size increases (Treisman & Gelade, 1980). This improved performance for feature search compared to conjunctive search has been replicated in a variety of contexts (e.g. Bergen & Julesz, 1983; Egeth, Virzi, & Garbart, 1984; Nakayama & Silverman, 1986; Wolfe & Franzel, 1988) and has informed much of the feature-based attention literature.

However, our previous work offers preliminary evidence that a different pattern of results emerges in the centroid paradigm (Winter, Wright, Chubb, & Sperling, 2016). We ran two experiments, each with visual search and centroid tasks. In the first experiment, we presented red and green, large and small items on a gray background. Targets could be defined by either color (red or green), size (large or small), or by a conjunction of color and size (red & large, red & small, green & large, or green & small). In the search task, performance was best in the search-for-color conditions, followed by the search-for-size conditions, and worst in the search-for-conjunction conditions. However, in the centroid task, performance was best in the attend-to-color conditions, followed by the attend-to-conjunction conditions, and worst in the attend-to-size conditions. In the second experiment, we presented black and white circles and triangles items on a gray background. Targets could be defined by either luminance (black or white), shape (circle or triangle), or by a conjunction of color and size (black circle, black triangle, white circle, or white triangles). Again, we found the expected results for the search task (performance was best in the search-for-luminance conditions, next best in the search-for-shape conditions, and worst in the search-for-conjunction conditions) and some unexpected results in the centroid task (performance was best in the attend-to-luminance conditions, next best in the attend-to-conjunction conditions, and worst in the attend-to-shape conditions). In short, we found evidence from both experiments for improved performance on conjunctive centroid judgments compared to constituent-feature centroid judgments.

It is important to keep in mind, though, that our previous work offers only preliminary evidence. Given the surprising nature of our conjunctive-target findings, replication is particularly important. So, in the present study, we seek to replicate and extend our previous centroid results, while also addressing some limitations of the previous experiments. The major limitation was the use of the gray background. In each experiment, the gray background was intermediate between the two levels of one feature dimension; that is, the gray background fell in between the red and green color levels in the first experiment, and between the black and white luminance levels in the second. It is possible, then, that the gray background allowed for an automatic segmentation of redder/greener (darker/lighter) items in the centroid task, thus facilitating performance on color-size (luminance-shape) conjunction conditions relative to size (shape) feature conditions. To remove this potential confound in the current study, we presented various gray items on a white background. In addition, we varied the similarity of the levels of each feature dimension. We found that not only are conjunctive targets better than one constituent-feature target, but there are also cases in which conjunctive targets are better than or equal to both constituent-feature targets.

2.2 METHODS

In the centroid paradigm, participants briefly view a stimulus cloud then strive to mouse-click the centroid, or center or mass, of the target items while ignoring any distractors. The

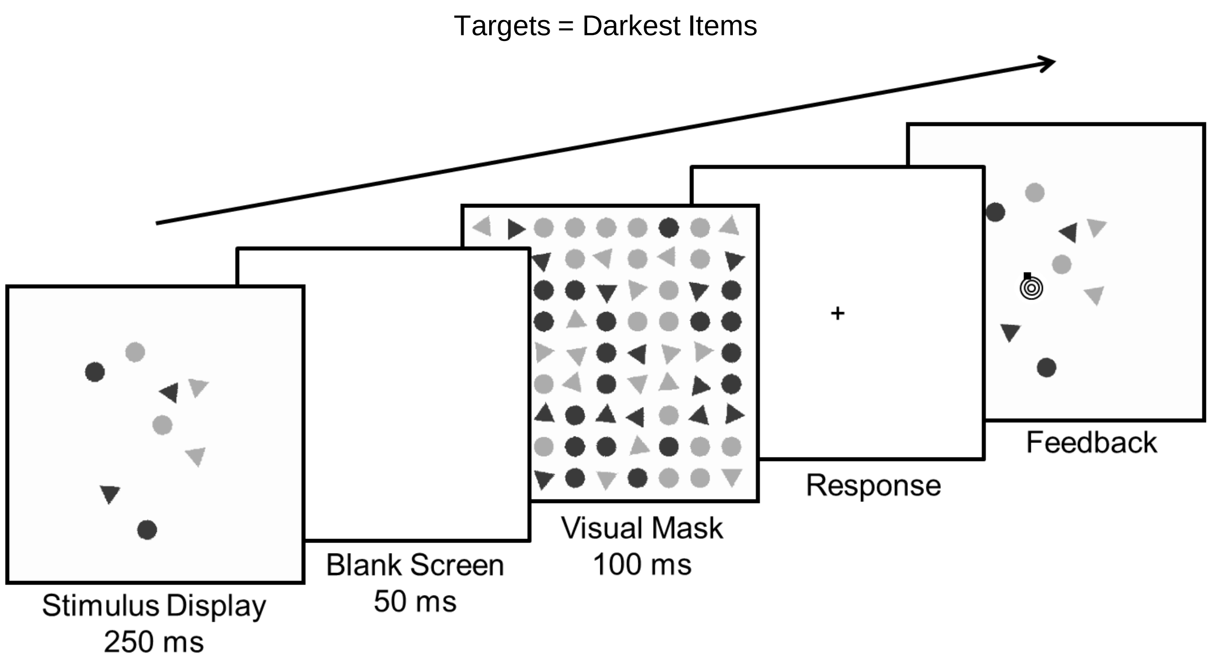


Figure 2.1: Example Exp. 3 Centroid Trial. The participant's task is to estimate the centroid, or center of mass, of the targets (for example, the darkest items) while ignoring the distractors. She briefly views the stimulus display containing the items, which is followed by a blank screen and a visual mask. She then enters her response via mouse-click and receives feedback before advancing to the next trial. The feedback screen shows her response as a small, black square and the actual centroid as a bullseye.

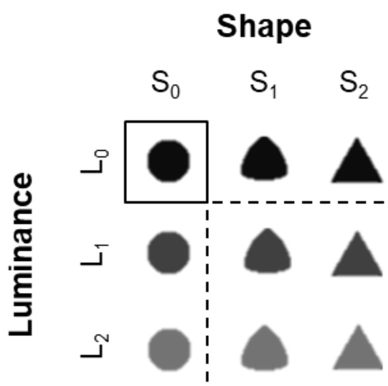


Figure 2.2: Exp. 3 Item Types. Each display contained four different items types. The darkest luminance level (L_0, top row) was always present, as was the most circular shape level (S_0, leftmost column). As a result, the boxed item was present in every display type. The fixed L_0 luminance level could be paired with either L_1 items (as in $\scriptstyle LS$ and $\scriptstyle L\displaystyle S$ displays) or with L_2 items (as in $L\scriptstyle S$ and LS displays). Likewise, the fixed S_0 shape level could be paired with either S_1 items (as in $\scriptstyle LS$ and $L\scriptstyle S$ displays) or with S_2 items (as in $\scriptstyle L\displaystyle S$ and LS displays). In the Targets=Darkest conditions, targets always consisted of items from the top row. Targets consisted of items in the top row in the Targets=Darkest condition, and targets consisted of items in the leftmost column in the Targets=Circles condition. In the Targests=DarkestCircles condition, the boxed item in the top left corner was the only target item type.

target items can be defined by a single feature (such as luminance) or a combination of features (such as luminance and shape). Figure 2.1 shows a sample centroid trial for this experiment. A more detailed description of centroid methodology, as well as other applications, can be found in Sun et al, 2015.

2.2.1 Display Types

In the current study, there were eight different display types. Every item used in these displays had one of three luminances and one of three shapes. The possible luminances were L_0 (3.846 cd/m^2), L_1 (22.16 cd/m^2), and L_2 (43.37 cd/m^2), all of which were presented on a white background (107.6 cd/m^2). The possible shapes were S_0 (a circle whose radius subtended approximately 0.2417 degrees of visual angle), S_1 (a triangular shape with rounded edges that subtended approximately 0.2935 degrees of visual angle from its center to each vertex), and S_2 (a triangle that subtended approximately 0.3281 degrees of visual angle from its center to each vertex), all of which were presented within a display region that subtended approximately 13.7467 degrees of visual angle. Figure 2.2 shows all possible item types.

Let $t(j,k)$ be the item-type with luminance L_j, j = 0,1,2 and shape S_k, k = 0,1,2. Then displays of type

> *LS*8 (*LS*16) contained 2 (4) each of items of type $t(0,0)$, $t(0,1)$, $t(1,0)$ and $t(1,1)$.
>
> *LS*8 (*LS*16) contained 2 (4) each of items of type $t(0,0)$, $t(0,2)$, $t(1,0)$ and $t(1,2)$.

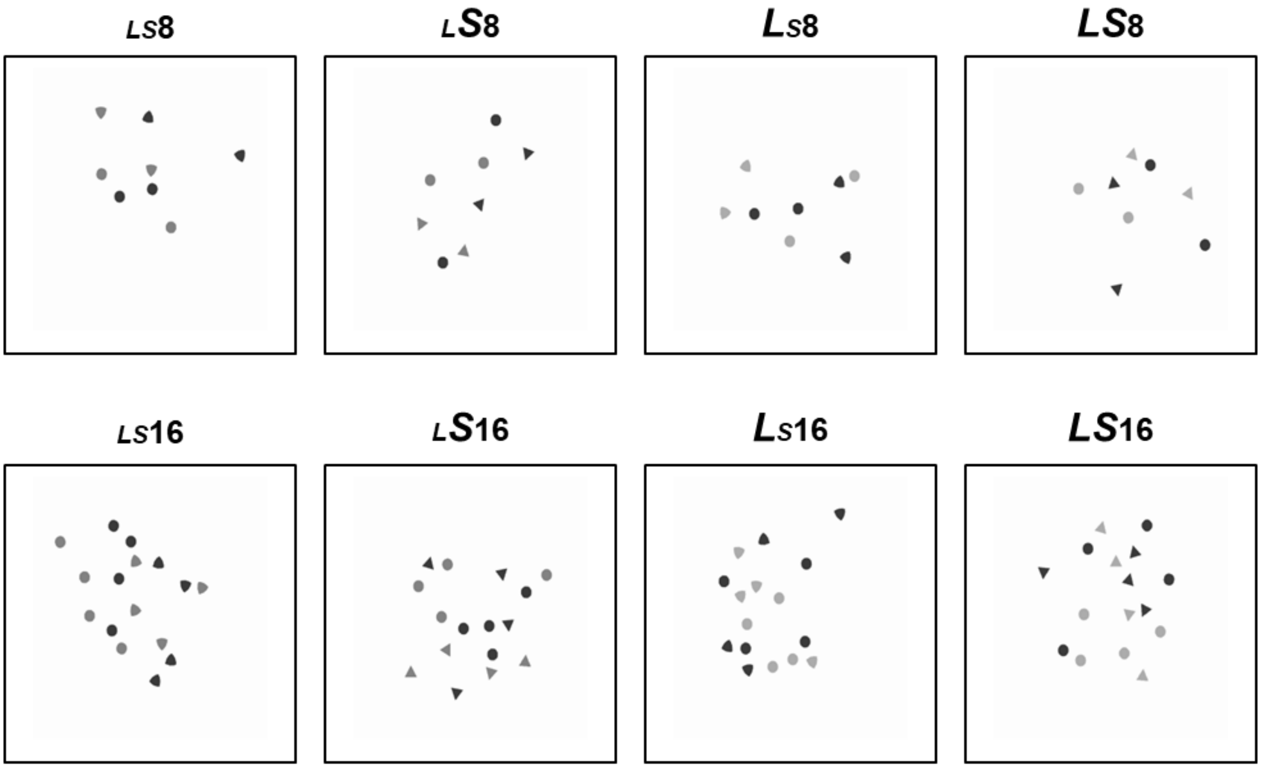


Figure 2.3: Exp. 3 Display Types. Each stimulus display contained four items types, consisting of two luminance levels and two shape levels. The L_0 luminance level was always present and could be paired either the L_1 level (as in *LS* and *LS* displays) or the L_2 level (as in *Ls* and *LS* displays); likewise the S_0 shape level was always present and could be paired either the S_1 level (as in *LS* and *Ls* displays) or the S_2 level (as in *LS* and *LS* displays). In the feature target conditions, in which targets were defined by either luminance or shape, the displays always contained eight items—four targets and four distractors. In order to keep the number of targets constant, we doubled the number of items in the display types for the conjunction condition, in which targets were defined by both luminance and shape. The resulting conjunction condition displays had 16 items—four targets and 12 distractors. Otherwise, the 8-item and 16-item displays were identical.

$L{\scriptstyle S}8$ ($L{\scriptstyle S}16$) contained 2 (4) each of items of type $t(0,0)$, $t(0,1)$, $t(2,0)$ and $t(2,1)$.

$LS8$ ($LS16$) contained 2 (4) each of items of type $t(0,0)$, $t(0,2)$, $t(2,0)$ and $t(2,2)$.

An example of each of these eight display types is shown in Figure2. 3.

2.2.2 Experimental Conditions

A given condition was defined by (1) the type of display used in that condition, and (2) the attention instruction for that condition. All attention instruction conditions were tested in separate blocks, and the different display type conditions were intermixed throughout those blocks. In each condition, the attention construction designated some item-types (in the display type used in that condition) as targets and the others as distractors. The task was to mouse-click the centroid of the target-item locations ignoring distractors.

Attention instructions:

1. Targets=Darkest: target items were the two types with luminance L_0.
2. Targets=Circles: target items were the two types with shape S_0.
3. Targets=DarkestCircles: target items were of type $t(0,0)$.

We tested the following conditions:

1. With each of attention instructions Targets=Darkest and Targets=Circles, each participant was tested in separately blocked conditions with display types
    a. ${\scriptstyle LS}8$
    b. ${\scriptstyle L}S8$

c. $L{\scriptstyle S}8$

d. $LS8$

2. With attention instruction Targets=DarkestCircles, each participant was tested in separately blocked conditions with display types
   a. ${\scriptstyle LS}16$
   b. ${\scriptstyle L}S16$
   c. $L{\scriptstyle S}16$
   d. $LS16$

In the four conditions using the Targets=Darkest attention instruction, all targets will have luminance L_0, and all distractors will have a different fixed luminance (the distractor luminance will be L_1 (L_2) in conditions using displays ${\scriptstyle LS}8$ or ${\scriptstyle L}S8$ ($L{\scriptstyle S}8$ or $LS8$)). Similarly, in the four conditions using the Targets=Circles attention instruction, target items will all be circles, and all distractors will have a fixed non-circular shape. Thus in all of the Targets=Darkest and Targets=Circles conditions, targets differ from distractors along a single feature-dimension (either in luminance or in shape). For this reason, we call these conditions "feature" conditions.

By contrast, in each of the four conditions using the Targets=DarkestCircles attention instruction, the targets are identical to some of the distractors in luminance and they are identical to other of the distractors in shape. Thus, in order to correctly identify an item as a target item, one must confirm both that it has shape S_0 and also luminance L_0. For this reason, we call these conditions, "conjunction" conditions.

The reader will note that all of the conjunction conditions use displays with twice as many items as the displays used in the feature conditions. Less obvious is the fact that the items in the 16-item displays are more densely packed in space than are the items in the 8-item displays. The point of these manipulations is to keep the number and spatial distribution of target items identical in the feature and conjunction conditions.

The items were presented in a cloud stimulus display measuring 800x800 pixels. The dispersion on the stimulus cloud is given by:

$$\text{Dispersion}(x, y) = \left[\frac{1}{2N_{cloud} - 1} \sum_{i=1}^{N_{cloud}} (x_i - \bar{X})^2 + (y_i - \bar{Y})^2\right]^{\frac{1}{2}}$$

where N_{cloud} is the total number of items in the cloud, $x = \left(x_1, x_2, \cdots, x_{N_{cloud}}\right)$ and $y = \left(y_1, y_2, \cdots, y_{N_{cloud}}\right)$ are the vectors of *x*- and *y*-coordinates of the items, and $\bar{X}$ $(\bar{Y})$ is the mean of vector x (y). Each cloud had a fixed dispersion of $116\frac{2}{3}$ (2.3019 degrees of visual angle), or one-sixth of the 800-pixel (13.7467 degrees of visual angle) stimulus display.

Participants were eight undergraduate and graduate students at the University of California, Irvine. Those with no prior centroid experience completed 500 trials of target-only centroid training prior to the experiment. All participants began the experiment with 120 practice trials identical to the experimental trials (40 trials per target condition). After the practice trials, there were two experimental sessions with 600 trials each, for a total of 1,200 trials. There were 400 trials per target condition and 100 trials per target condition and display type. The two feature conditions (Targets=Darkest and Targets=Circles) were grouped, so the conjunction condition (Targets=DarkestCircles) appeared either at the very beginning or the very end of the session. The order of the feature conditions were counterbalanced across the two sessions for each participant, as was the order placement of the conjunction condition at either the beginning or the end of the session. So, a participant who completed the attention conditions in a Targets=Darkest, Targets=Circles, Targets=DarkestCircles order during the first session would complete the attend conditions in a Targets=DarkestCircles, Targets=Circles, Targets=Darkest order during the second session.

2.3 RESULTS

We used the methods described in Sun et al. (2016) to analyze our centroid data. The model of Sun et al. (2016) assumes that the x- and y-coordinates of the participant's response on each trial in a given condition C are given by

$$R_x = \mu_x + Q_x \qquad \text{and} \qquad R_y = \mu_y + Q_y$$

for Q_x and Q_y independent, normally distributed random variables with mean 0 and some standard deviation σ and

$$\mu_x = \frac{\sum f_C(\tau_i) x_i}{\sum f_C(\tau_i)} \qquad \text{and} \qquad \mu_y = \frac{\sum f_C(\tau_i) y_i}{\sum f_C(\tau_i)}$$

where each sum is over all items i in the stimulus cloud, τ_i is the type of item i, and x_i and y_i are the x- and y- coordinates of item i, and f_C is the attention filter achieved by the participant.

These methods enable us to estimate (1) the attention filter f_C achieved by the participant in condition C, and (2) the *Efficiency* with which the participant was able to deploy the filter f_C. The attention filter f_C defines the relative influence exerted on the participant's responses by all four types of items occurring in the stimulus. Figure 2.4 shows the attention filters achieved in all 12 conditions averaged across eight participants. The function f_C is constrained to sum to 1; however, it is possible for f_C to assign negative values to some item types. The participant's Efficiency in deploying f_C is the minimum possible proportion of items that had to be included, on average, in the participant's centroid computation to achieve predicted responses of the accuracy observed. Efficiency is estimated by assuming that all residual error (i.e., the deviations of responses predicted by the model from actual responses) is due to removing a fixed proportion Q of randomly chosen items from the display on each trial and applying the model to the decimated display without additional error; Efficiency is then taken to be equal to $1 - Q$. See Figure 2.5 for the average Efficiency for all 12 conditions.

In "binary" centroid tasks of the sort used in this paper in which the target filter assigns equal weight to a specific set of target item-types and weight 0 to the remaining distractor item-types, it is convenient to summarize the effectiveness of the attention filter f_C achieved by the participant by the ratio of (numerator) the mean of $f_C(t)$ taken across all target items t divided by (denominator) the mean of $|f_C(d)|$ taken across all distractor items$d$. This statistic (Selectivity Ratio) provides a convenient index of the degree to which the attention filter achieved by the participant accentuates target items while filtering out distractor items. A Selectivity Ratio of ten or higher is considered excellent. We calculated the average Selectivities for participants across the 12 conditions (shown in Figure 2.6).

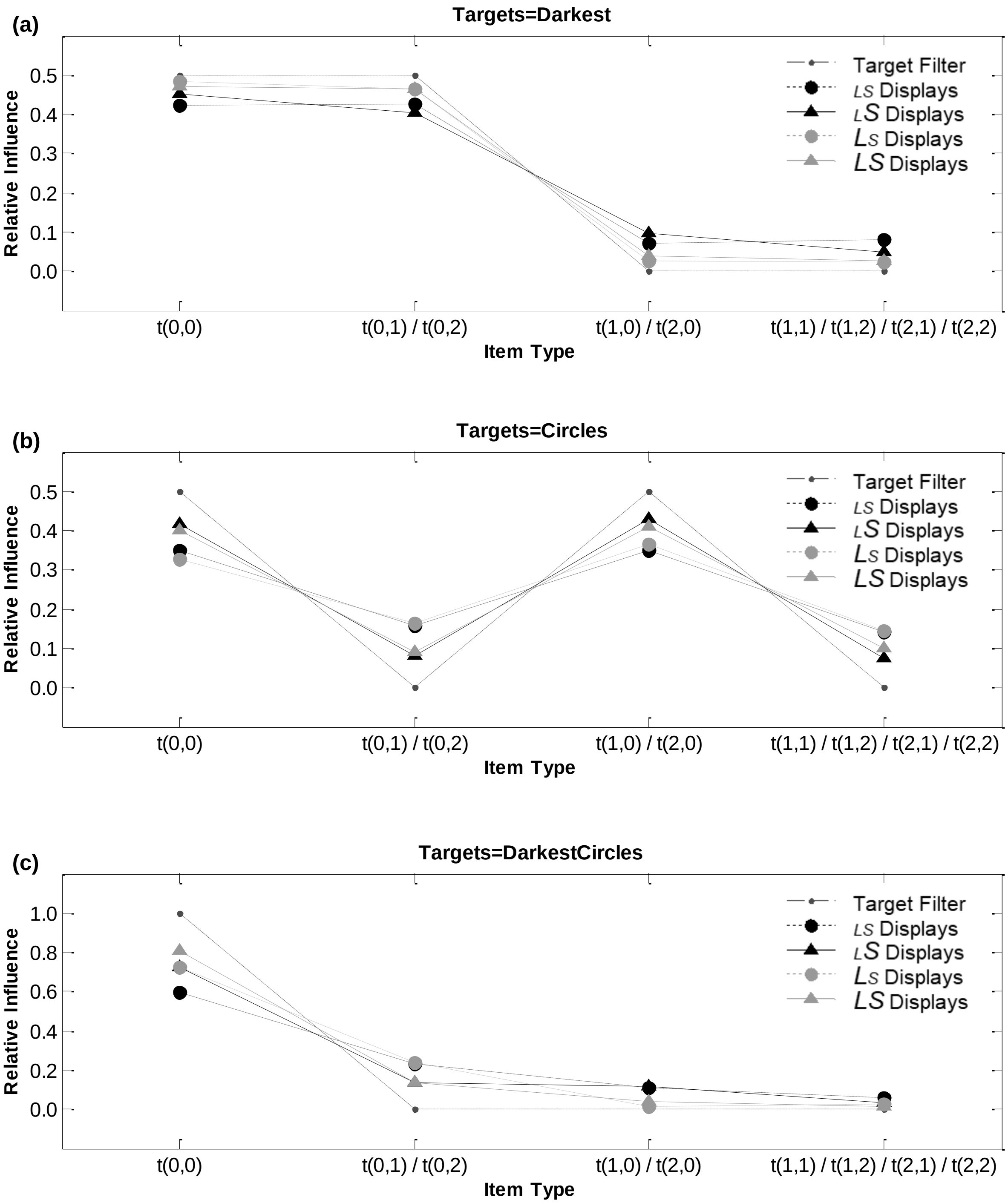


Figure 2.4: Exp. 3 Relative Influence. The plots above show the relative influence of each item type on participants' centroid estimations, averaged across eight participants. The dashed line in each graph represents the target filter. **(a)** In the Targets=Darkest condition, the targets are items $t(0,0)$ and $t(0,1)/t(0,2)$. **(b)** In the Targets=Circles condition, the targets are items $t(0,0)$ and $t(1,0)/t(2,0)$. **(c)** In the Targets=DarkestCircles condition, the targets are item $t(0,0)$.

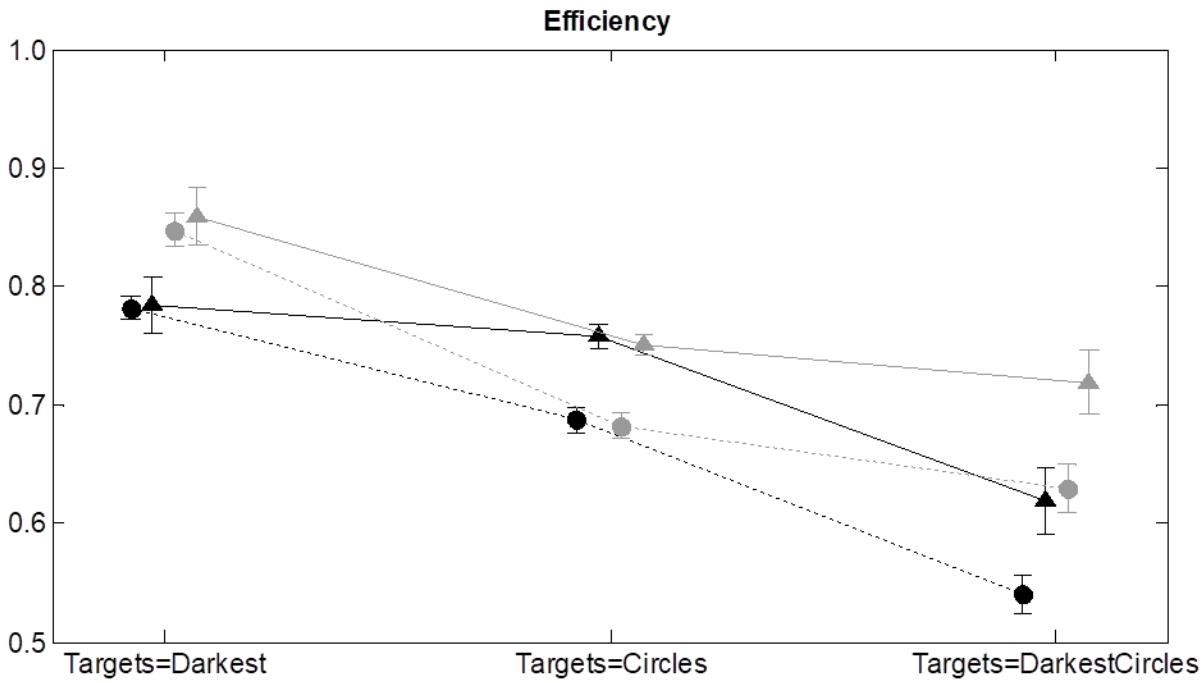


Figure 2.5: Exp. 3 Efficiency. The plot above shows the Efficiency, or the lower bound on the proportion of items observed, for each display type and target condition, averaged across eight participants. In the Targets=Darkest condition, Efficiency is higher when the dark targets are among lighter distractors (*Ls* and *LS*). In the Targets=Circles condition, Efficiency is higher when the circular targets are among more triangular distractors (*LS* and *LS*). In the Targets=DarkestCircles condition, Efficiency is highest when the darker, circular targets are among lighter, more triangular distractors (*LS*) and lowest when among less light, less triangular distractors (*LS*). For each display type, Efficiency appears to be lowest in the Conjunction condition. However, there were 16 items in that condition, compared to only eight in the Luminance and Shape conditions. Since Efficiency estimates a proportion of items observed, the lower Conjunction Efficiencies don't necessarily indicate poorer performance.

For our statistical analyses, we began with paired-samples *t*-tests of participants' Efficiencies across display types within each target condition. For each target condition, we grouped the display types according to the levels of the relevant feature dimension(s). So, in the Target=Darkest conditions, we compared the two display types with L_1 distractors (*LS*8and *LS*8) to the two display types L_2 distractors (*Ls*8 and *LS*8) and found that mean Efficiency was significantly higher for the L_2-distractor displays than for the L_1-distractor displays. In the Targets=Circles condition, we compared the two display types with S_1 distractors (*LS*8and *Ls*8) to the two display types with S_2 distractors (*LS*8 and *LS*8) and

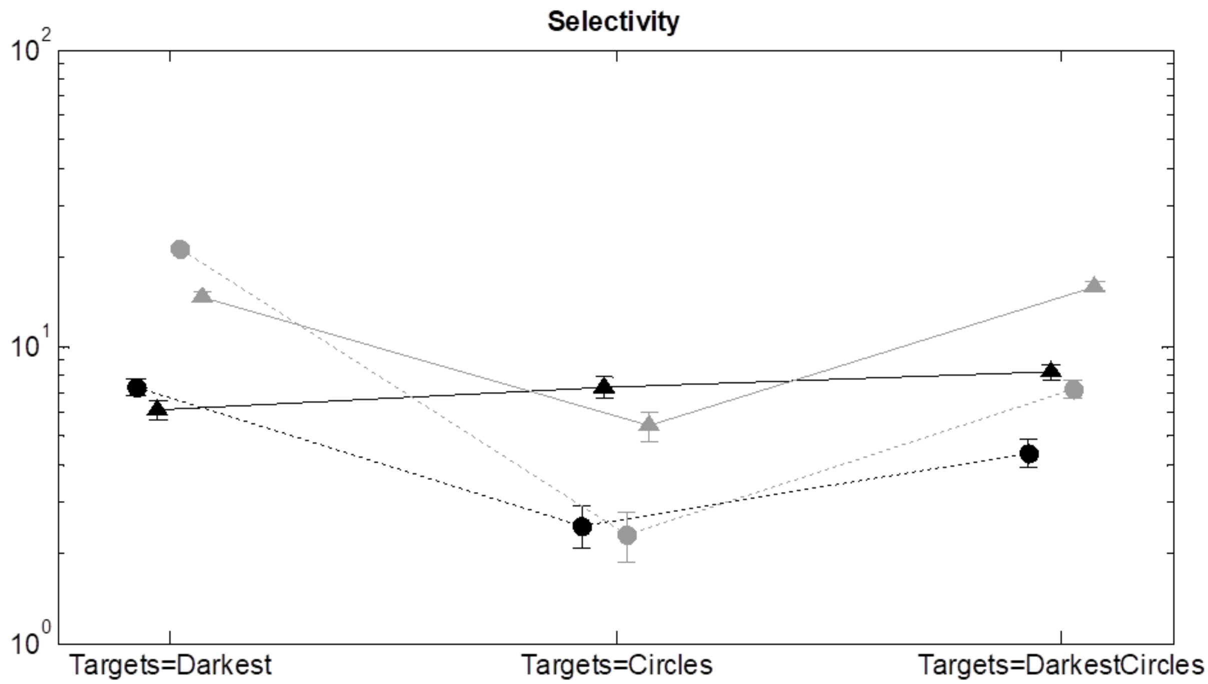


Figure 2.6: Exp. 3 Selectivity Ratio. The Selectivity Ratio is calculated by dividing the mean target weight by the mean distractor weight. The plot above shows the Selectivity for each display type and target condition, averaged across eight participants. In the Targets=Darkest condition, Selectivity is higher when the dark targets are among lighter distractors ($L{\scriptstyle S}$ and LS). In the Targets=Circles condition, Selectivity is higher when the circular targets are among more triangular distractors (${\scriptstyle L}S$ and LS). In the Targets=DarkestCircles condition, Selectivity is highest when the darker, circular targets are among lighter, more triangular distractors (LS) and lowest when among less light, less triangular distractors (${\scriptstyle LS}$). For each display type, the Selectivity in the conjunction condition is greater than or equal to the Selectivity in at least one constituent-feature condition.

found that the mean Efficiency was significantly higher for S_2-distractor displays than for the S_1-distractor displays. In the Targets=DarkestCircles condition, both feature dimensions were relevant, so we compared the display type with L_1 and S_1 distractors (${\scriptstyle LS}16$) to the display with L_2 and S_2 distractors ($LS16$) and found that the mean Efficiency for the $LS16$ displays was significantly higher than for the ${\scriptstyle LS}16$ displays. We also compared the display type with L_1 and S_1 distractors (${\scriptstyle LS}16$) to the display types with L_1 and S_2 distractors (${\scriptstyle L}S16$) and L_2 and S_1 distractors ($L{\scriptstyle S}16$) and found the that Efficiency was significantly higher for the ${\scriptstyle L}S16$ and $L{\scriptstyle S}16$displays. Additionally, a comparison between the display type with L_2

and S_2 distractors (LS16) and the display types with L_1 and S_2 distractors (${\scriptstyle L}S$16) and L_2 and S_1 distractors ($L{\scriptstyle S}$16) revealed that Efficiency was significantly higher in the former. Table 2.1 summarizes all our Efficiency comparisons.

We performed the same paired-samples t-tests on participants' Selectivity Ratios and found similar trends. In the Targets=Darkest condition, the mean Selectivity was lower for the L_1-distractor displays (${\scriptstyle LS}$8and ${\scriptstyle L}S$8) than for the L_2-distractor displays ($L{\scriptstyle S}$8 and LS8). In the Targets=Circles condition, the mean Selectivity was lower for the S_1-distractor displays (${\scriptstyle LS}$8and $L{\scriptstyle S}$8) than for the S_2-distractor displays (${\scriptstyle L}S$8 and LS8). In the Targets=DarkestCircles condition, the mean Selectivity for the ${\scriptstyle LS}$16 display condition was significantly lower than both the LS16 display condition and the ${\scriptstyle L}S$16 and $L{\scriptstyle S}$16 display conditions, and the mean Selectivity was higher for the LS16 display condition than the ${\scriptstyle L}S$16 and $L{\scriptstyle S}$16 display conditions. All our Selectivity comparisons are summarized in Table 2.2.

We performed an additional set of paired-samples t-tests for participants' Selectivity Ratios across target conditions for each display type. Table 2.3 shows a complete list of these comparisons, but perhaps the most interesting result is that, for each display type, Selectivity was significantly higher in the Targets=DarkestCircles condition than in one of the constituent feature target conditions. That is, for the ${\scriptstyle LS}$, $L{\scriptstyle S}$, and LS display types, Selectivity was higher in the Targets=DarkestCircles condition than in the Targets=Circles condition; for the ${\scriptstyle L}S$ display type, Selectivity was higher in the Targets=DarkestCircles

Targets=Darkest Efficiency			
Contrast	$\frac{L{\scriptstyle S} + LS}{2} - \frac{{\scriptstyle LS} + {\scriptstyle L}S}{2}$	${\scriptstyle L}S - {\scriptstyle LS}$	$LS - L{\scriptstyle S}$
Mean	0.070	0.002	0.011
Std. Dev.	0.037	0.056	0.047
Upper Bound	0.102	0.049	0.050
Lower Bound	0.039	-0.044	-0.027
T	5.302	0.117	0.698
p	0.001	0.910	0.508
Bayes Factor	37.573	0.338	0.411

Targets=Circles Efficiency			
Contrast	$\frac{{\scriptstyle L}S + LS}{2} - \frac{{\scriptstyle LS} + L{\scriptstyle S}}{2}$	$L{\scriptstyle S} - {\scriptstyle LS}$	$LS - {\scriptstyle L}S$
Mean	0.069	0.066	0.075
Std. Dev.	0.043	0.037	0.040
Upper Bound	0.105	0.097	0.108
Lower Bound	0.033	0.035	0.042
T	4.548	5.012	5.353
p	0.003	0.002	0.001
Bayes Factor	18.799	28.981	39.295

Targets=DarkestCircles Efficiency				
Contrast	$LS - {\scriptstyle LS}$	$L{\scriptstyle S} - {\scriptstyle L}S$	$\frac{{\scriptstyle L}S + L{\scriptstyle S}}{2} - {\scriptstyle LS}$	$LS - \frac{{\scriptstyle L}S + L{\scriptstyle S}}{2}$
Mean	0.179	0.010	0.084	0.095
Std. Dev.	0.112	0.094	0.077	0.055
Upper Bound	0.272	0.089	0.148	0.142
Lower Bound	0.085	-0.068	0.019	0.049
T	4.521	0.310	3.050	4.869
p	0.003	0.766	0.019	0.002
Bayes Factor	18.319	0.350	4.030	25.421

Table 2.1: Exp. 2 Display Type Efficiencies by Target Condition. Within each target condition, display types were grouped based on the similarity/difference of the relevant feature dimension(s). Efficiency is higher when the two levels of the relevant feature dimension(s) are different and lower when they are similar. The upper and lower bounds of the 95% confidence interval are provided.

Targets=Darkest Selectivity Ratio			
Contrast	$\frac{L{\scriptstyle S} + LS}{2} - \frac{{\scriptstyle LS} + {\scriptstyle L}S}{2}$	${\scriptstyle L}S - {\scriptstyle LS}$	$LS - L{\scriptstyle S}$
Log10 Mean	0.980	-0.180	-0.381
Std. Dev.	0.888	0.579	0.340
Upper Bound	1.722	0.304	-0.098
Lower Bound	0.238	-0.664	-0.665
T	3.122	-0.879	-3.177
p	0.017	0.408	0.016
Bayes Factor	4.358	0.459	4.626

Targets=Circles Selectivity Ratio			
Contrast	$\frac{{\scriptstyle L}S + LS}{2} - \frac{{\scriptstyle LS} + L{\scriptstyle S}}{2}$	$L{\scriptstyle S} - {\scriptstyle LS}$	$LS - {\scriptstyle L}S$
Log10 Mean	0.962	1.080	0.879
Std. Dev.	0.633	0.880	0.948
Upper Bound	1.491	1.816	1.672
Lower Bound	0.433	0.345	0.086
T	4.299	3.472	2.622
p	0.004	0.010	0.034
Bayes Factor	14.769	6.344	2.520

Targets=DarkestCircles Selectivity Ratio				
Contrast	$LS - {\scriptstyle LS}$	$L{\scriptstyle S} - {\scriptstyle L}S$	$\frac{{\scriptstyle L}S + L{\scriptstyle S}}{2} - {\scriptstyle LS}$	$LS - \frac{{\scriptstyle L}S + L{\scriptstyle S}}{2}$
Log10 Mean	1.293	-0.132	0.559	0.735
Std. Dev.	0.681	0.290	0.245	0.637
Upper Bound	1.863	0.111	0.764	1.267
Lower Bound	0.724	-0.374	0.353	0.202
T	5.370	-1.282	6.438	3.262
p	0.001	0.241	0.000	0.014
Bayes Factor	39.884	0.629	96.306	5.070

Table 2.2: Exp. 2 Display Type Selectivity Ratios by Target Condition. Within each target condition, display types were grouped based on the similarity/difference of the relevant feature dimension(s). Selectivity is higher when the two levels of the relevant feature dimension(s) are different and lower when they are similar, following the same pattern as the Efficiency data in Table 1. The upper and lower bounds of the 95% confidence interval are provided.

LS Selectivity Ratio			
Contrast	**D − C**	**D − DC**	**C − DC**
Log10 Mean	1.074	0.510	-0.564
Std. Dev.	0.490	0.619	0.243
Upper Bound	1.484	1.028	-0.361
Lower Bound	0.664	-0.007	-0.768
T	6.198	2.331	-6.554
p	0.000	0.053	0.000
Bayes Factor	79.698	1.830	105.350

LS Selectivity Ratio			
Contrast	**D − C**	**D − DC**	**C − DC**
Log10 Mean	-0.178	-0.294	-0.116
Std. Dev.	0.757	0.333	0.770
Upper Bound	0.455	-0.016	0.528
Lower Bound	-0.811	-0.573	-0.760
T	-0.665	-2.496	-0.427
p	0.527	0.041	0.682
Bayes Factor	0.404	2.193	0.363

LS Selectivity Ratio			
Contrast	**D − C**	**D −D C**	**C − DC**
Log10 Mean	2.229	1.098	-1.131
Std. Dev.	0.604	0.714	0.330
Upper Bound	2.734	1.695	-0.855
Lower Bound	1.724	0.501	-1.407
T	10.438	4.348	-9.688
p	0.000	0.003	0.000
Bayes Factor	1260.015	15.495	833.877

LS Selectivity Ratio			
Contrast	**D − S**	**D − DC**	**S − DC**
Log10 Mean	0.997	-0.084	-1.081
Std. Dev.	1.076	0.927	0.703
Upper Bound	1.897	0.691	-0.493
Lower Bound	0.097	-0.859	-1.669
T	2.620	-0.257	-4.347
p	0.034	0.805	0.003
Bayes Factor	2.514	0.438	15.480

Table 2.3: Exp. 2 Target Condition Selectivity Ratios by Display Type. For each display type, we compared the selectivity ratios of the target conditions. The Targets=Darkest condition is abbreviated D, the Targets=Circles condition C, and the Targets=DarkestCircles condition DC. In each display type, the Selectivity of the conjunction condition is significantly greater than that of at least one constituent feature condition. In the *LS* and *LS* display types, the Selectivity of the conjunction condition is greater than or equal to that of both constituent feature conditions. The upper and lower bounds of the 95% confidence interval are provided.

condition than in the Targets=Darkest condition. Furthermore, in the *LS* and *LS* displays, Selectivity was at least as high as in the conjunction target condition as it was in *both* constituent feature target conditions. That is, the Targets=DarkestCircles Selectivity was at least as high as the Targets=Circles Selectivity in addition to being significantly higher than the Targets=Darkest Selectivity in the *LS* displays, and it was at least as high as the Targets=Darkest Selectivity as well as being significantly higher than the Targets=Circles Selectivity in *LS* displays. It is worth noting that we performed these analyses on our

Efficiency	
Contrast	$\left(\left[\frac{\mathrm{D}(\boldsymbol{L}{\scriptstyle S}) + \mathrm{D}(\boldsymbol{LS})}{2}\right] - \left[\frac{\mathrm{D}({\scriptstyle LS}) + \mathrm{D}({\scriptstyle L}\boldsymbol{S})}{2}\right]\right) - \left(\left[\frac{\mathrm{C}({\scriptstyle L}\boldsymbol{S}) + \mathrm{C}(\boldsymbol{LS})}{2}\right] - \left[\frac{\mathrm{C}({\scriptstyle LS}) + \mathrm{C}(\boldsymbol{L}{\scriptstyle S})}{2}\right]\right)$
Mean	0.001
Std. Dev.	0.033
Upper Bound	0.028
Lower Bound	-0.027
T	0.046
p	0.965
Bayes Factor	0.337

Selectivity Ratio	
Contrast	$\left(\left[\frac{\mathrm{D}(\boldsymbol{L}{\scriptstyle S}) + \mathrm{D}(\boldsymbol{LS})}{2}\right] - \left[\frac{\mathrm{D}({\scriptstyle LS}) + \mathrm{D}({\scriptstyle L}\boldsymbol{S})}{2}\right]\right) - \left(\left[\frac{\mathrm{C}({\scriptstyle L}\boldsymbol{S}) + \mathrm{C}(\boldsymbol{LS})}{2}\right] - \left[\frac{\mathrm{C}({\scriptstyle LS}) + \mathrm{C}(\boldsymbol{L}{\scriptstyle S})}{2}\right]\right)$
Log10 Mean	0.009
Std. Dev.	0.635
Upper Bound	0.540
Lower Bound	-0.522
T	0.040
p	0.969
Bayes Factor	0.337

Table 2.4: Exp. 2 Luminance and Shape Difference of Differences. We compared the difference of differences for the Targets=Darkest and Targets=Circles conditions to investigate whether the performance cost was the same when the levels of the relevant feature dimension were similar compared to when they were different. Displays types (*LS*, *LS*, *LS*, and *LS*) are written as functions of the target conditions (D and C). The upper and lower bounds of the 95% confidence interval are provided. These comparisons suggest that this performance cost of target-distractor similarity on the relevant feature dimension is likely the same for both these feature target conditions.

Selectivity Ratio data but not on our Efficiency data because Efficiency gives a lower bound on the *proportion* of items observed. Since the conjunction target condition displays contained twice as many items, the comparison between conjunction and feature target conditions would not be a meaningful one.

Finally, we compared the difference of differences in performance (as measured by both Efficiency and Selectivity Ratio) for the Luminance and Shape conditions. For this analysis, we grouped the display types in which the levels of the relevant feature dimensions were similar (L_1-distractor displays in the Targets=Darkest condition, S_1-distractor displays in the Targets=Circles condition) and the levels in they were different (L_2-distractor displays in the Targets=Darkest condition, S_2-distractor displays in the Targets=Circles condition). As shown in Table 2.4, it appears that the performance cost of similarity on the relevant dimension was the same for both feature target conditions.

2.4 DISCUSSION AND CONCLUSIONS

First and foremost, our present study replicates our previous finding that performance on the centroid task, as measured by Selectivity Ratio, is better when targets are defined by a conjunction of features than by just one of those constituent features. This finding alone deviates sharply from the visual search literature, which consistently finds that performance is considerably worse for conjunctive targets than feature targets. There are some exceptions in the search literature, however, in which conjunctive targets are sometimes found to be no worse than feature targets (Theeuwes & Kooi, 1994; Nakayama & Silverman, 1986). Evidence of *improved* performance for conjunctive targets in the centroid task, then, is a strikingly new result. Our replication provides reassurance that this surprising result was not merely an accident of chance.

Second, by varying the similarity along both feature dimensions, we were able to find a display type ($_LS$) in which Selectivity on both the Target=Circle and Target=DarkestCircle conditions is higher than Selectivity on the Target=Darkest condition. In our previous study, it was always the case that Luminance always outperformed Conjunction, which outperformed Shape. The new pattern, in which Shape and Conjunction outperform Luminance in ${}_{L}S$ displays, suggests that there is not necessarily anything special about the feature dimensions of luminance that aids performance in conjunction conditions. Rather, it seems that performance on conjunction conditions depends more on appropriately large differences between levels of either feature dimension.

Finally, our results suggest there is not necessarily any performance cost in the Conjunction condition relative to *either* of the constituent-feature target conditions. Specifically, we see this in the ${}_{L}S$ and LS display types. In the ${}_{L}S$ display type, Conjunction was at least as good as Shape, and both were better than Luminance. In the LS display type, Conjunction was at least as good as Luminance, and both were better than Shape. This is all the more striking when we consider that the displays themselves were arguably more complicated in the Conjunction condition (16-item displays, compared to eight-item displays in the feature target conditions). If, for instance, we had run the Shape condition using both the eight-item and 16-item displays, we could expect performance on the 16-item displays to be worse than on the eight-item displays. That is, the 16-item displays themselves were likely more difficult, completely independent of the target condition. And yet, the Conjunction conditions, even though they used more challenging displays, showed no performance cost compared to both constituent-feature target condition in the ${}_{L}S$ and LS display types. Indeed, it even showed a performance *benefit* compared to one constituent-feature target condition.

Based solely on the visual search literature, one might predict that conjunctive-target centroid tasks would be outright impossible to do. That participants could perform even reasonably well on a conjunctive centroid task would be an interesting discovery in its own right. However, our finding that conjunctive targets are better than or equal to *both* constituent-feature targets is unprecedented in light of the visual search literature. This provides compelling evidence that visual search cannot tell the whole story of feature-based attention. Every task comes with its own task demands, and neither search nor centroids are exceptions. Overreliance on a single methodology runs the risk of conflating particular task demands with inherent properties of feature-based attention—such as the supposed primacy of feature targets over conjunctive targets. The centroid paradigm is an alternative way to study feature-based attention, which can help us better understand what patterns of results are merely caused by task demands versus what is true of feature-based attention more generally. Of course, the use of even more methodologies would be better still.

CHAPTER 3

EXPERIMENT 4: EQUISALIENCE ANALYSIS

3.1 INTRODUCTION

Feature-based attention research has for decades relied heavily on the visual search task, in which observers identify a target in a field of distractors. Treisman & Gelade's (1980) Feature Integration Theory was an early and hugely influential account of the differences between feature search (in which a target is defined by a single feature, e.g. color) and conjunction search (in which a target is defined by a conjunction of features, e.g. color and shape). In short, Feature Integration Theory suggests that feature search can be accomplished by parallel processing of all items in the display, while conjunction search demands serial processing of the items, so that attentional processes can bind the features of each item together. From Treisman & Gelade's Feature Integration Theory, to Wolfe's ever-evolving Guided Search model (Wolfe, Cave, & Franzel, 1989; Wolfe, 1994; Wolfe & Gancarz, 1997; and Wolfe & Gray, 2007), to completely new ways of thinking about feature-based attention (e.g. Becker, Folk, & Remington, 2010; Buetti, Cronin, Madison, Wang, & Lleras, 2016), the visual search task remains the primary means of investigation. Indeed, many psychology conferences have sessions—often multiple—dedicated entirely to visual search. Clearly, the field is deeply indebted to this particular methodology.

However, by relying so heavily on a single task, it becomes increasingly difficult to separate out idiosyncratic task demands of visual search from inherent characteristics of feature-based attention. For a more complete picture, alternative methodologies need to be investigated. One such alternative is the centroid paradigm as developed by Sun, Chubb, Wright, and Sperling (2016). In the centroid paradigm, observers briefly view a stimulus cloud. The cloud can have target and distractor items, with the targets defined by one or more features. The stimulus cloud is then masked to prevent shifts in spatial attention, and the observer estimates the centroid—or center of mass—of the target items. Performance on the task indicates how well observers are able to attend to the targets while ignoring distractors.

Already, there are compelling reasons to believe that the centroid task is importantly different from visual search. While previous studies (e.g. Foster, & Ward, 1991) indicated that targets defined by line orientation produced "pop out" in visual search, Inverso, Sun, Chubb, Wright, and Sperling (2016) found that performance on an orientation centroid task declined rapidly as the number of items in the display increased. Previously, we directly compared performance on visual search and centroid tasks, in which targets were defined by either a single feature or by a conjunction of features (Winter, Wright, Chubb, & Sperling, 2016). We found the expected pattern of results for visual search: performance was better for all feature target conditions compared to all conjunctive target conditions. However, our centroid results revealed a different pattern of performance. Participants performed better when estimating the centroids of conjunctive targets (e.g. black circles) than when estimating the centroids of constituent-feature targets (e.g. circles). These centroid results pose a direct challenge to Feature Integration Theory, which has been supported almost entirely by visual search studies but makes claims about feature-based attention generally. That is, Feature Integration Theory cannot explain the improved performance for conjunctive targets relative to feature targets in the centroid task. Indeed, the typical visual search results of flat slopes for feature targets and steeper slopes for conjunctive targets have been replicated so many times that even alternative theories of feature-based attention typically offer some sort of explanation for this phenomenon. Importantly, however, the scope of these explanations ought to be limited to feature-based attention *in the context of visual search* when they rely solely on evidence from the search task.

It is possible that centroids are processed differently than are other types of visual stimuli, which may contribute to the divergent pattern of results described above. For instance, Zhou, Chu, Li, and Zhan (2006) found that centroids automatically capture attention and are not subject to inhibition of return. Furthermore, it remains an open question how exactly observers estimate centroids in the first place. One seemingly plausible explanation is that they count the target items then average those items' locations. If this were indeed the case, we would expect observers to perform better when judging numerosity (which requires only counting) than when judging centroids (which, under this explanation, requires both counting and averaging). However, Inverso, Chubb, Wright, Shiffrin, and Sperling (2016) report higher efficiencies for centroid estimations than for numerosity judgments. If there is, in fact, something special about the way centroids are processed, the natural next question is *what*. One possibility is that the centroid paradigm provides access to different sorts of information than do other tasks, or at least that it uses the same information differently. Here, we test this possibility using the equisalience analysis outlined by Wright, Chubb, Winkler, and Stern (2013).

To perform Wright et al.'s (2013) equisalience analysis, a researcher measures a participant's performance in a task for a range of levels of feature dimension X and a range a levels of feature dimension Y. She can then fit a Weibull function to the participants data for each feature dimension, and use an inverse Weibull function to estimate the physical levels of each feature dimension at which the participant performs with various levels of accuracy (for instance, 10% accuracy, 20% accuracy, 30% accuracy, and so on). Then, she can create a graph with the physical levels of feature dimension X on one axis and the physical levels of feature dimension Y on the other and plot the point that corresponds to the level at which the participant performed with, for instance, 10% accuracy for each dimension. By doing this for a range of performance levels, she finds the X-to-Y equisalience function for the task. Repeating this process for another task allows the researcher to compare the X-to-Y equisalience functions for the two tasks. Different X-to-Y equisalience functions for different tasks suggest that those tasks has access to different sorts of information.

In the current study, we ask whether the luminance-to-shape equisalience function for a centroid task using 4 target items and 4 distractor items is the same as the luminance-to-shape equisalience function for an analogous task using just a single target item and a single distractor item. Given that the tasks differ only in the number of items per display (two versus eight), one might reasonably expect them to use the same kind of information in the same ways. In fact, this is exactly what we had expected: we initially developed the 2-

item task with the intention of using it to create centroid stimulus displays in which luminance and shape discriminability were equally matched for each participant. However, our pilot data made it immediately clear that the matched luminance and shape levels in the 2-item task did not produced matched levels in the centroid task. For a range of distractor luminances that were easily discriminable from the target luminance in the centroid task, the "matched" distractor shapes were virtually impossible to distinguish from the target shape. This unexpected discovery motivated the present study, in which we compare the equisalience functions of the 2-item and 8-item tasks. We find that the equisalience functions are indeed different, suggesting that the two tasks differ in their access to information carried by item luminance versus item shape.

3.2 METHODS

3.2.1 Participants

The participants were four UC Irvine graduate students, all of whom had extensive centroid training prior to our experiment. The methods were approved by the UC Irvine IRB, and all participants provided signed consent.

3.2.2 Procedure

The experiment consisted of two staircased tasks: a 2-item (two-alternative forced-choice) task and an 8-item (centroid) task. For both tasks, we ran three staircases (1up,2down; 1up,3down; 1up,4down) with intermixed trials in each block in order to accurately estimate a range of points along the psychometric function.

In the 2-item task, one target item and one distractor item appeared for 180 msec, followed by a blank screen for 20 msec and a mask for 100 msec. The target was defined by either

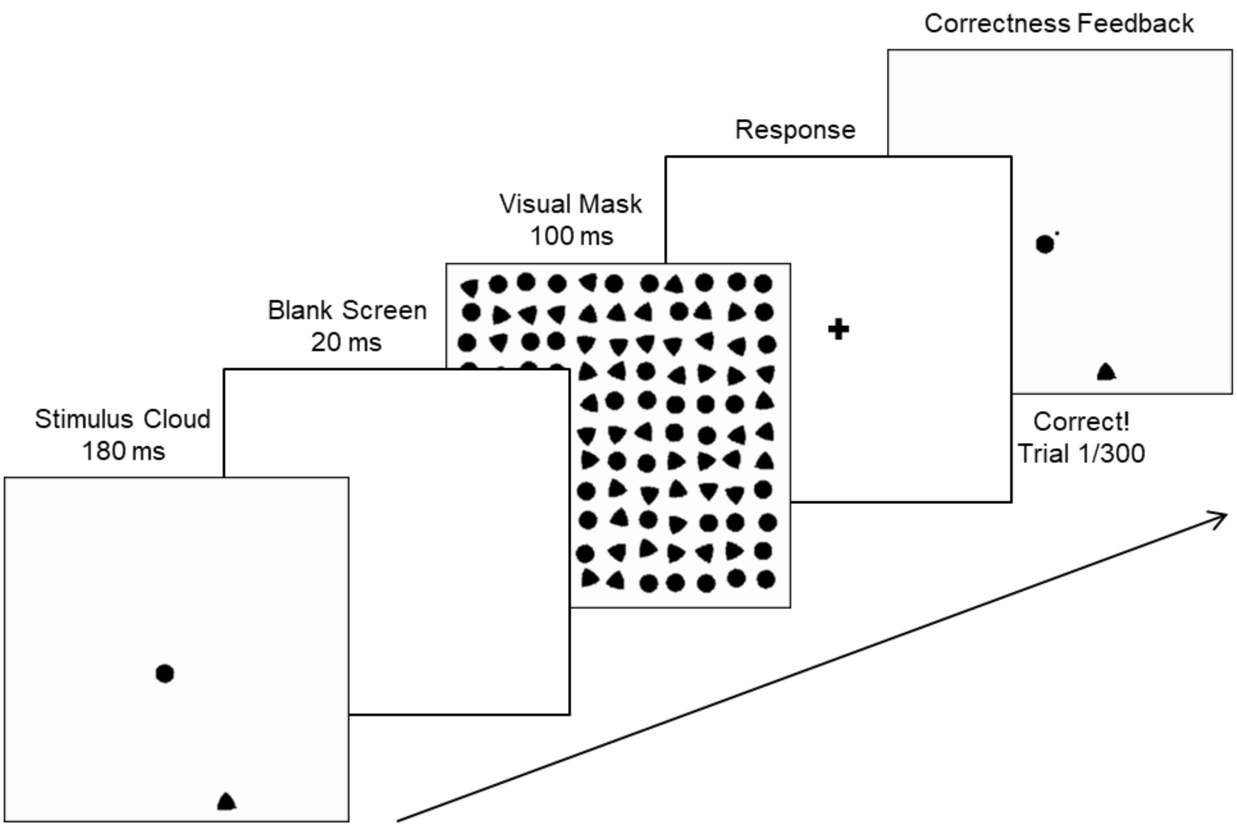


Figure 3.1: Example Exp. 4 Trial for the 2-Item Task. A stimulus cloud containing two items, one target (the more circular item) and one distractor (the more triangular item), is briefly presented and masked. The participant enters her response by moving the cursor (+) and clicking on the location where the target appeared. She then receives visual correctness feedback. The feedback screen shows the original stimulus cloud with a small black dot where she clicked. Beneath the display, written feedback of either "Correct!" or "Incorrect" appears, along with trial progress information. Her response is considered correct if it is closer to the target than to distractor.

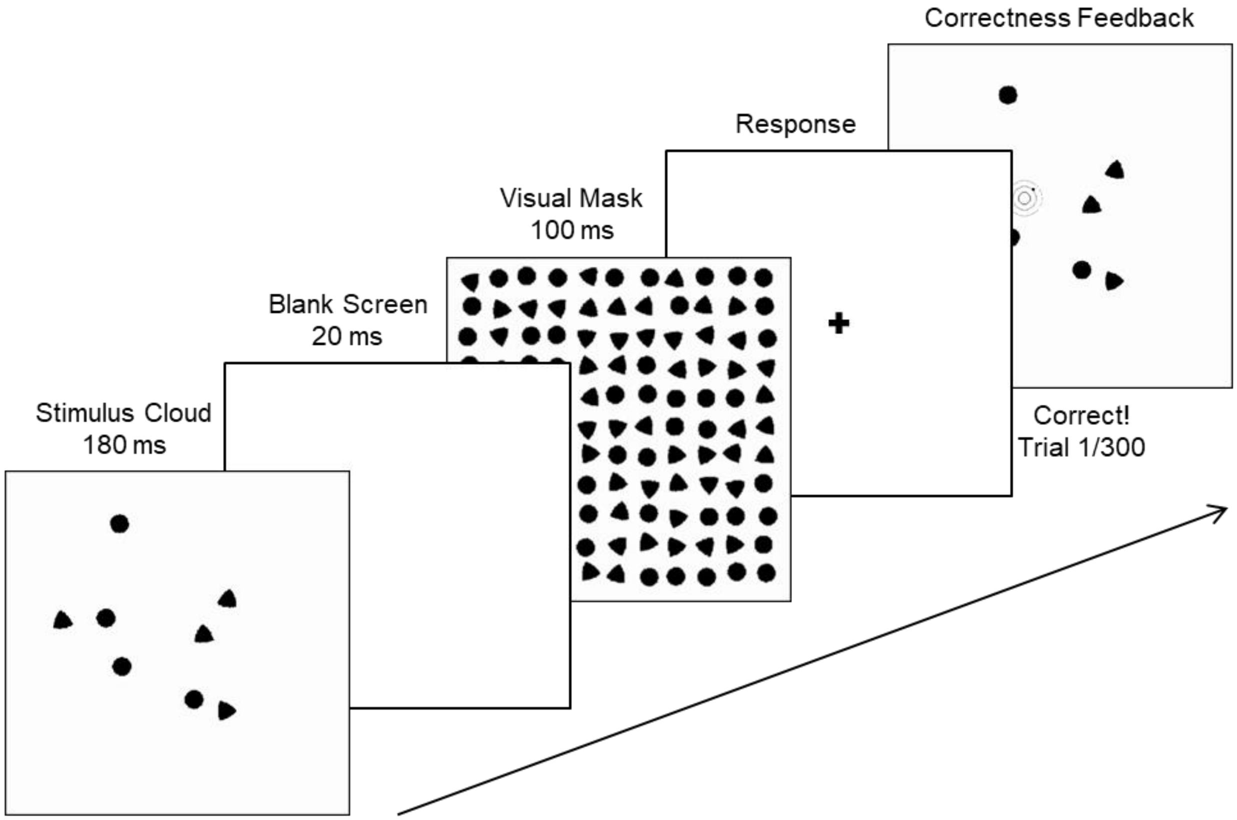


Figure 3.2: Example Exp. 4 Trial for the 8-Item Task. A stimulus cloud containing eight items, four targets (the more circular items) and four distractors (the more triangular items), is briefly presented and masked. The participant enters her response by moving the cursor (+) and clicking on the centroid, or center of mass, of the target items. She then receives visual correctness feedback. The feedback screen shows the original stimulus cloud with a bullseye over the actual centroid a small black dot indicating her centroid estimation. Beneath the display, written feedback of either "Correct!" or "Incorrect" appears, along with trial progress information. Her response is considered correct if it is within a threshold distance from the actual centroid and incorrect otherwise.

luminance (in which case the target was always the darker of the two items) or shape (in which case the target was always the more circular of the two items). The mask was then replaced by a response screen with the cursor positioned in the center of display. The participant was instructed to move the cursor and click on the location where the target item had appeared. She then received visual correctness feedback before advancing to the next trial by pressing the space bar. Her response was considered correct if it was closer to the target than to the distractor and otherwise incorrect. Figure 3.1 shows an example of a 2-item task trial in which the target was defined by shape.

In the 8-item task, the stimulus cloud consisted of eight items: four targets and four distractors. As in the 2-item task, the stimulus cloud appeared for 180 msec, followed by a blank screen for 20 msec and a backward mask for 100 msec. Again, targets were defined by either luminance (i.e., the four darker items) or shape (i.e., the four circular items). The participant was then asked to estimate the centroid, or center of mass, of the four target items while ignoring the distractors. The response screen was the same as in the 2-item task: blank except for a cursor in the center. The participant moved the cursor and clicked to indicate her response, then received visual correctness feedback. Figure 3.2 shows an example of an 8-item task trial in which the target was defined by shape.

In order to run the 8-item task as staircased procedure, we needed a way to distinguish between correct responses and incorrect ones. However, this posed a challenge since typical analysis of centroid data involves continuous measures of error and influence. (See Sun et al, 20l6, for a more detailed description of centroid methodology and analysis.) Our solution was to determine a threshold distance from the true centroid of the targets: if the participant's response was within this threshold distance from the actual centroid, the trial was considered a success. Otherwise, the trial was a failure. We estimated a threshold distance for each participant in a centroid threshold task prior to the experiment. In the centroid threshold task (Figure 3.3), the stimulus cloud contained four targets and no distractors. Participants estimated the centroid of these target-only displays and received feedback after every trial. There were two sessions, each with two 60-trial blocks, for a

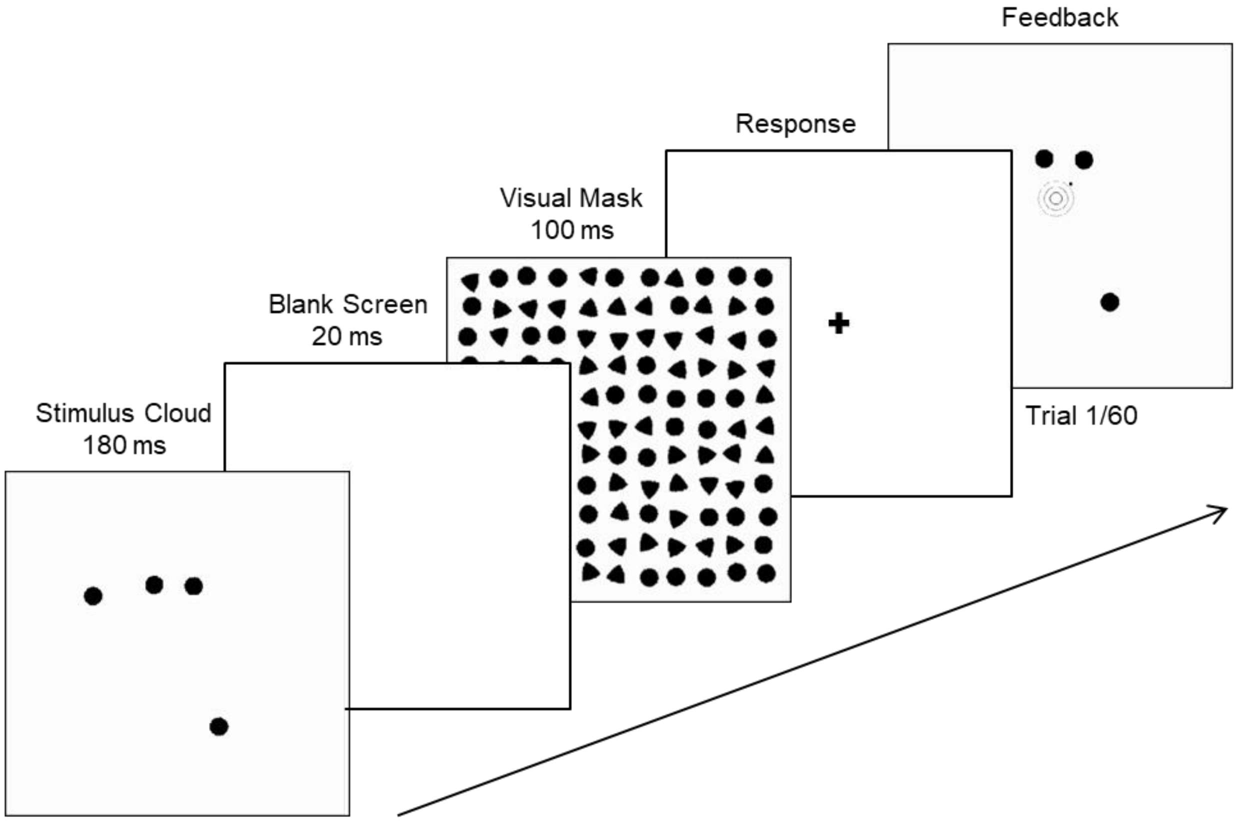


Figure 3.3: Example Centroid Threshold Trial. A stimulus cloud containing four items, all targets, is briefly presented and masked. The participant enters her response by moving the cursor (+) and clicking on the centroid, or center of mass, of all four items. She then receives centroid feedback showing the actual centroid as a bullseye and her centroid estimation as a small black dot. Beneath the feedback display, her trial progress information appears.

total of 240 trials. For each participant, we calculated the distance between her response and the actual centroid on each trial, giving us a distribution of error. We then defined the participant's threshold distance as the value that was larger than 95% of the values in the error distribution. We later used this participant-specific threshold to determine whether a response was correct or incorrect in the staircased 8-item centroid task.

We introduced one further modification to the 8-item task in order to make it more effective as an adaptive staircase procedure. In a typical centroid experiment, there are no restrictions on the distance between the target centroid and the distractor centroid. If the responses of the participant are influenced by the distractors as well as the target items, then this random variation in the location of the distractor centroid will induce corresponding, random, trial-by-trial variation in the distance of the participant's response from the target centroid. The effect of such stimulus-driven response noise would be to artificially flatten the psychometric function we seek to measure. To ensure a more robust estimate of the psychometric function, the distance between the target centroid and the distractor centroid was always at least 1.5 times the participant's threshold distance.

3.2.3 Stimuli

In both the 2-item and 8-item tasks, items were presented in a stimulus region measuring 600x600 pixels circumscribed by a thin, square black frame. The dispersion of a stimulus cloud is given by

$$\text{Dispersion}(x, y) = \left[\frac{1}{2N_{cloud} - 1} \sum_{i=1}^{N_{cloud}} (x_i - \bar{X})^2 + (y_i - \bar{Y})^2\right]^{\frac{1}{2}}$$

where N_{cloud} is the total number of items in the cloud, $x = \left(x_1, x_2, \cdots, x_{N_{cloud}}\right)$ and $y = \left(y_1, y_2, \cdots, y_{N_{cloud}}\right)$ are the vectors of x- and y-coordinates of the items, and $\bar{X}$ ($\bar{Y}$) is the mean of vector x (y). Each cloud had a fixed dispersion of 100 pixels (approximately 1.73 degrees of visual angle), or one-sixth of the 600-pixel (approximately 10.33 degrees of visual angle) stimulus display. By fixing the dispersion across the two tasks, we ensured that each item (regardless of task) had an equal chance of falling in the center or the periphery of the participant's visual field.

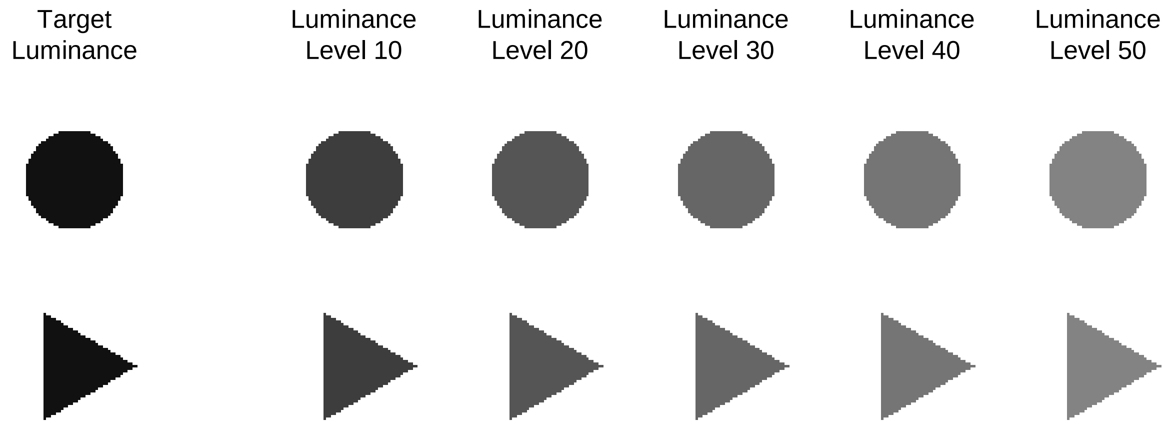


Figure 3.4: Exp. 4 Luminance Levels. Luminance discriminations were made in the context of either circular items (top row) or triangular items (bottom row). The target luminance (leftmost column) was fixed and always darker than the distractor luminance (all other columns). Lower distractor luminance levels (e.g. Level 10) indicate greater target-distractor similarity while higher distractor luminance levels (e.g. Level 50) indicate less target-distractor similarity.

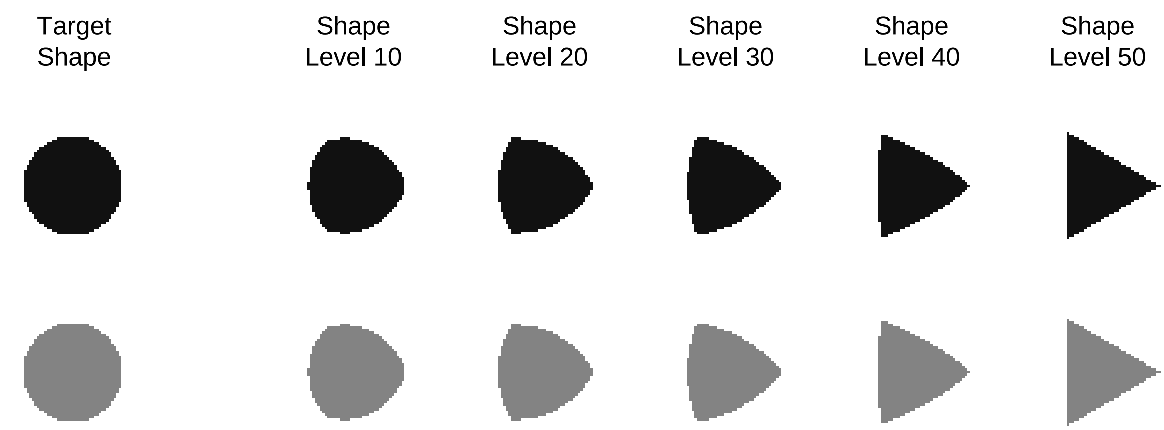


Figure 3.5: Exp. 4 Shape Levels. Shape discriminations were made in the context of either darker items (top row) or lighter items (bottom row). The target shape (leftmost column) was fixed and always more circular than the distractor shape (all other columns). Lower distractor shape levels (e.g. Level 10) indicate greater target-distractor similarity while higher distractor shape levels (e.g. Level 50) indicate less target-distractor similarity.

Also in both tasks, the target level was fixed and the distractor level changed to be more or less similar to the target. This meant that, in luminance discriminations, target items always had the same nearly-black luminance (0.41 cd/m^2) from trial to trial while distractor items increased or decreased in luminance in increments of approximately 0.43 cd/m^2 and presented against a white background (107.6 cd/m^2); similarly, in shape discriminations, the target item always had the same nearly-circular shape while the distractor items appeared more or less triangular. The shape of the items was formed by the intersection of three circles. For the most circular items, the three intersecting circles used to form the shape had almost complete overlap. For more triangular items, the radii of the circles were longer and the centers were spaced further apart. The distance from the item's center to a vertex subtended between 0.24 (for the most circular items) and 0.33 (for the most triangular items) degrees of visual angle. Figure 3.4 shows examples of various luminance levels, and Figure 3.5 examples of various shape levels.

3.2.4 Design

The relevant feature dimension was tested at two levels of the irrelevant feature dimension. So, participants made luminance discriminations between identically-circular items in one condition, and identically-triangular items in the other condition. Likewise, participants made shape discriminations between identically-dark items in one condition and identically-light items in the other condition. Figure 3.6 provides examples of such stimuli for each feature dimension and condition for the 2-item task, and Figure 3.7 provides equivalent examples for the 8-item task. For both tasks, the two conditions of

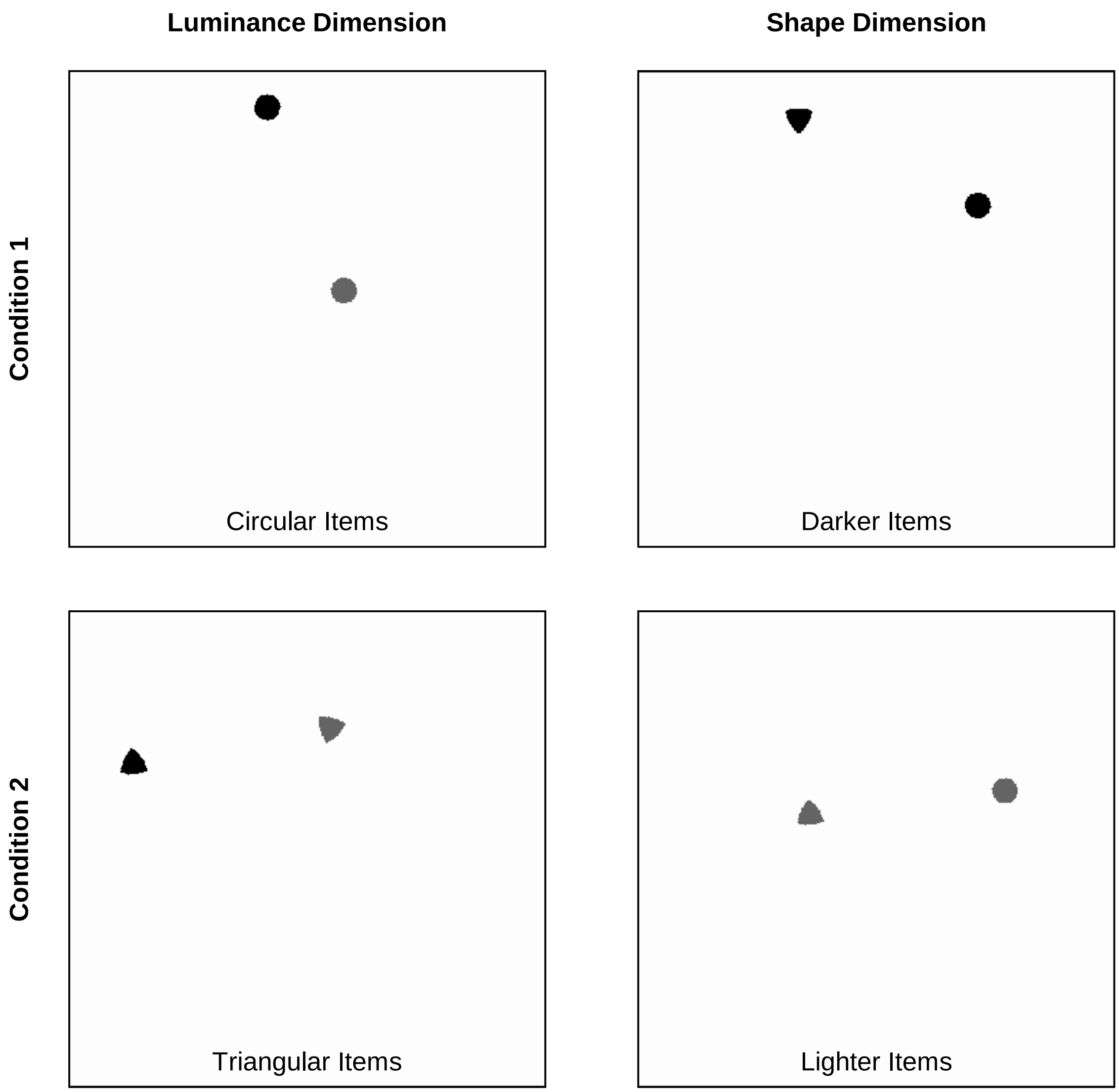


Figure 3.6: Exp. 4 Stimulus Displays for the 2-Item Task. In the 2-item task, only two items were displayed—one target and one distractor. In the luminance feature dimension (left column), the target was always darker than the distractor. Luminance discrimination was tested at two levels of the irrelevant feature dimension: shape. In the first luminance condition, both the target and distractor were nearly-perfect circles; in the second luminance condition, both the target and distractors were more triangular. In the shape feature dimension (right column), the target was always more circular than the distractor. Shape discrimination was also tested at two levels of the irrelevant feature dimension: luminance. In the first shape condition, both the target and distractor were a darker, nearly black gray; in the second shape condition, both the target and distractors were a lighter gray.

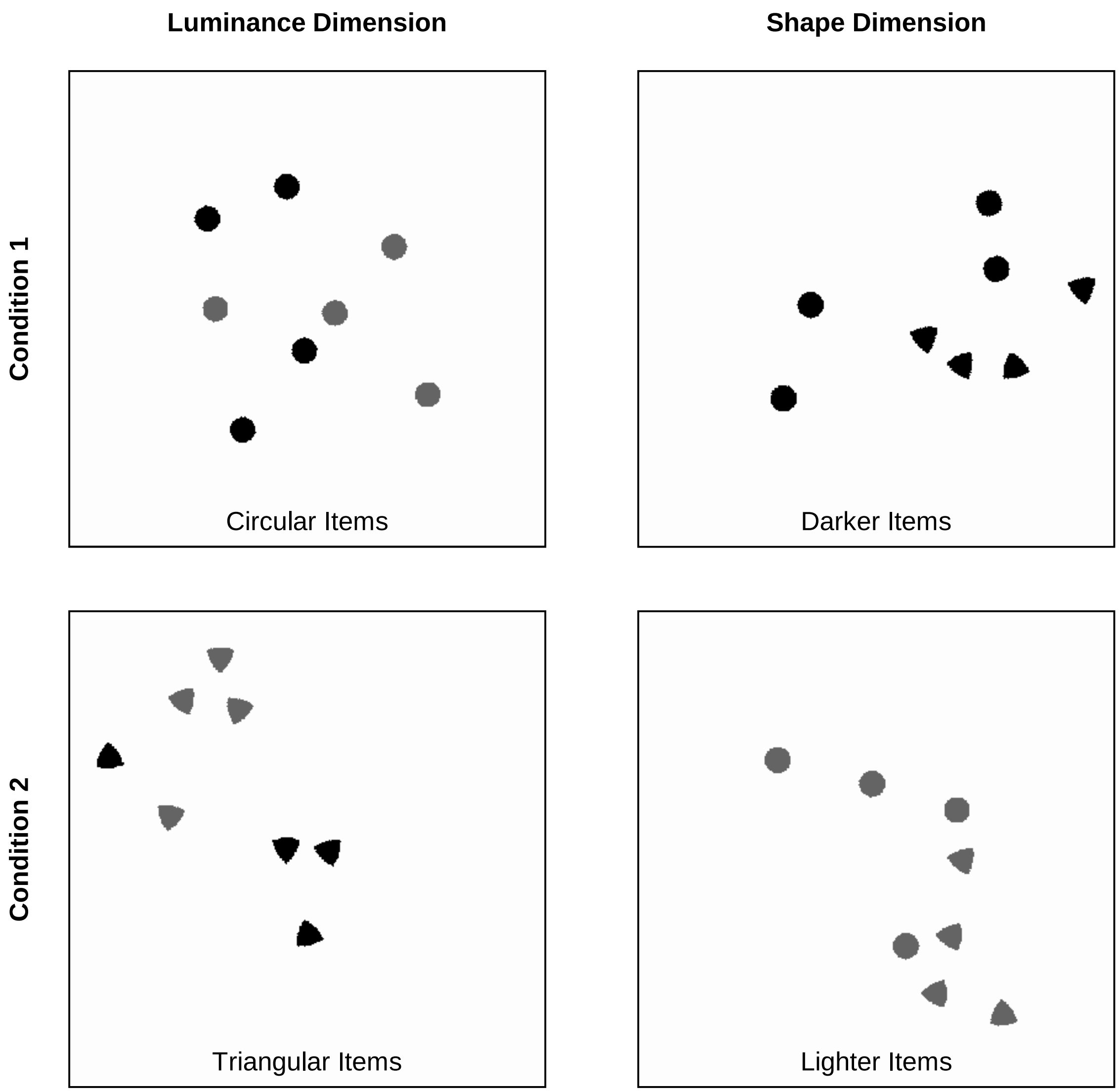


Figure 3.7: Exp. 4 Stimulus Displays for the 8-Item Task. In the 8-item task, eight items were displayed—four targets and four distractors. In the luminance feature dimension (left column), the targets were always darker than the distractors. Luminance discrimination was tested at two levels of the irrelevant feature dimension: shape. In the first luminance condition, both targets and distractors were nearly-perfect circles; in the second luminance condition, both targets and distractors were more triangular. In the shape feature dimension (right column), the targets were always more circular than the distractors. Shape discrimination was also tested at two levels of the irrelevant feature dimension: luminance. In the first shape condition, both targets and distractors were a darker, nearly black gray; in the second shape condition, both targets and distractors were a lighter gray.

each feature dimension were blocked. There were two blocks per session and four sessions in the experiment. There were 2,400 total trials, with 1,200 trials per task, 600 trials per task and feature dimension, and 300 trials per task, feature dimension, and condition block. The tasks followed an AABB order paired with feature dimensions in a CDDC order to make up the four sessions. The condition order was randomized. The task order and feature dimension order were counterbalanced across the four participants.

3.3 RESULTS

The first analysis compared the two conditions within a feature dimension and task. We sought to determine whether luminance discriminations in the context of circular items produced a different psychometric function than did luminance discriminations in the context of triangular items, and whether shape discriminations in the context of darker items produced a different psychometric function than did shape discriminations in the context of lighter items. For example, to test whether the psychometric functions for luminance derived using circular items versus using triangular items were different, we used a likelihood ratio test to compare the fit provided by a nested model in which the two Weibull functions share the same threshold parameter α and steepness parameter β versus a fuller model in which the two Weibull functions can have different threshold parameters α_{circ} and α_{tri} and steepness parameters β_{circ} and β_{tri}. Figures 3.8 through 3.11 show each participant's psychometric functions and the results of the likelihood ratio tests for all four pairs of task and feature dimension. Interestingly, these likelihood ratio tests did indeed decisively reject the null hypothesis that Weibull functions being compared shared the

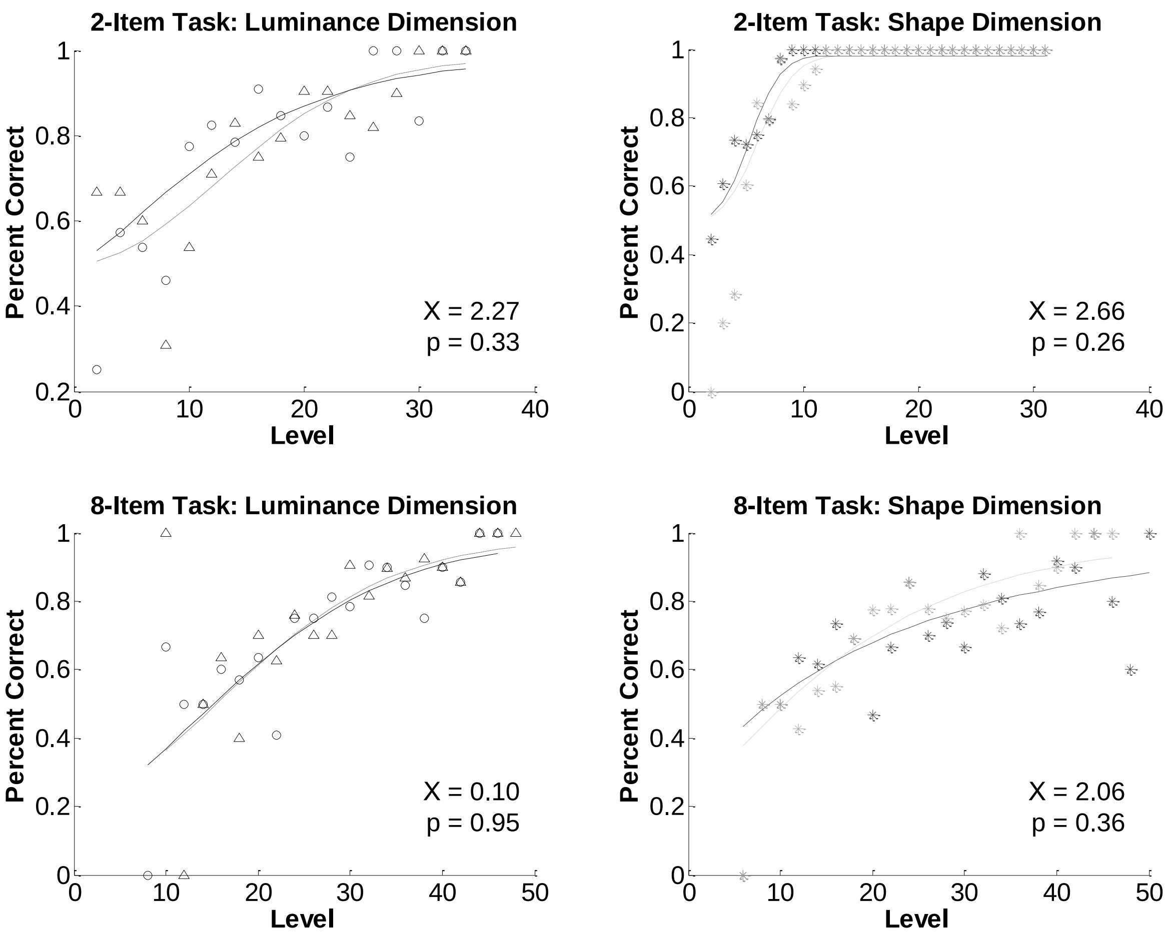


Figure 3.8: Comparison of Exp. 4 Conditions for Participant 1. The plots shows Participant 1's psychometric functions for each condition. In the Luminance dimension, circle markers show the participant's raw data for Condition 1 (circular items) and the solid line the line of best fit; triangle markers show the participant's raw data for Condition 2 (triangular items) and the dashed line the line of best fit. In the Shape Dimension, black markers show the participant's raw data for Condition 1 (darker items) and the solid black line the line of best fit; gray markers show the participant's raw data for Condition 2 (lighter items) and the dashed gray the line of best fit. We performed likelihood ratio tests for all conditions and give the chi-square score (X) for each. For this participant, none of the differences were significant.

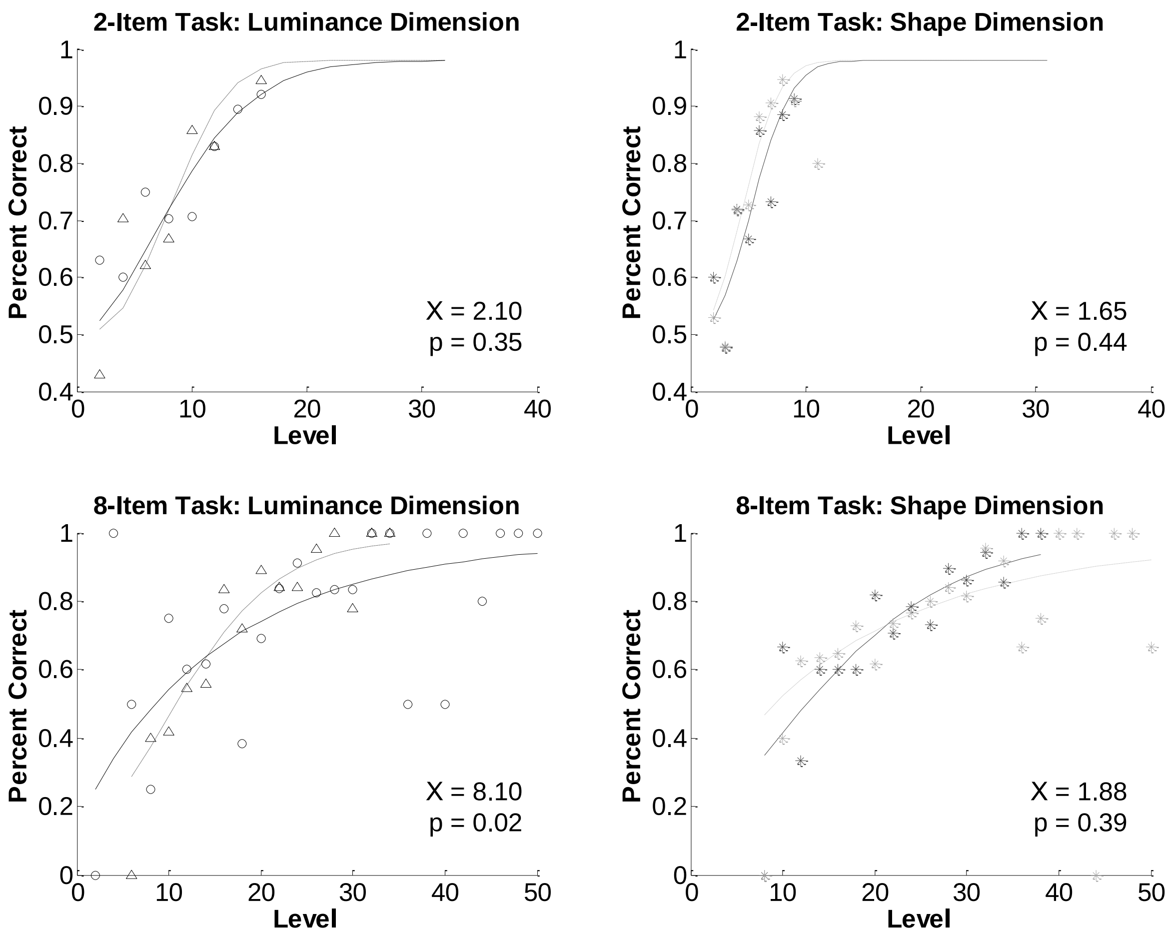


Figure 3.9: Comparison of Exp. 4 Conditions for Participant 2. The plots shows Participant 2's psychometric functions for each condition. In the Luminance dimension, circle markers show the participant's raw data for Condition 1 (circular items) and the solid line the line of best fit; triangle markers show the participant's raw data for Condition 2 (triangular items) and the dashed line the line of best fit. In the Shape Dimension, black markers show the participant's raw data for Condition 1 (darker items) and the solid black line the line of best fit; gray markers show the participant's raw data for Condition 2 (lighter items) and the dashed gray the line of best fit. We performed likelihood ratio tests for all conditions and give the chi-square score (X) for each. For this participant, the difference between conditions in the Luminance dimension of the 8-item task is significant.

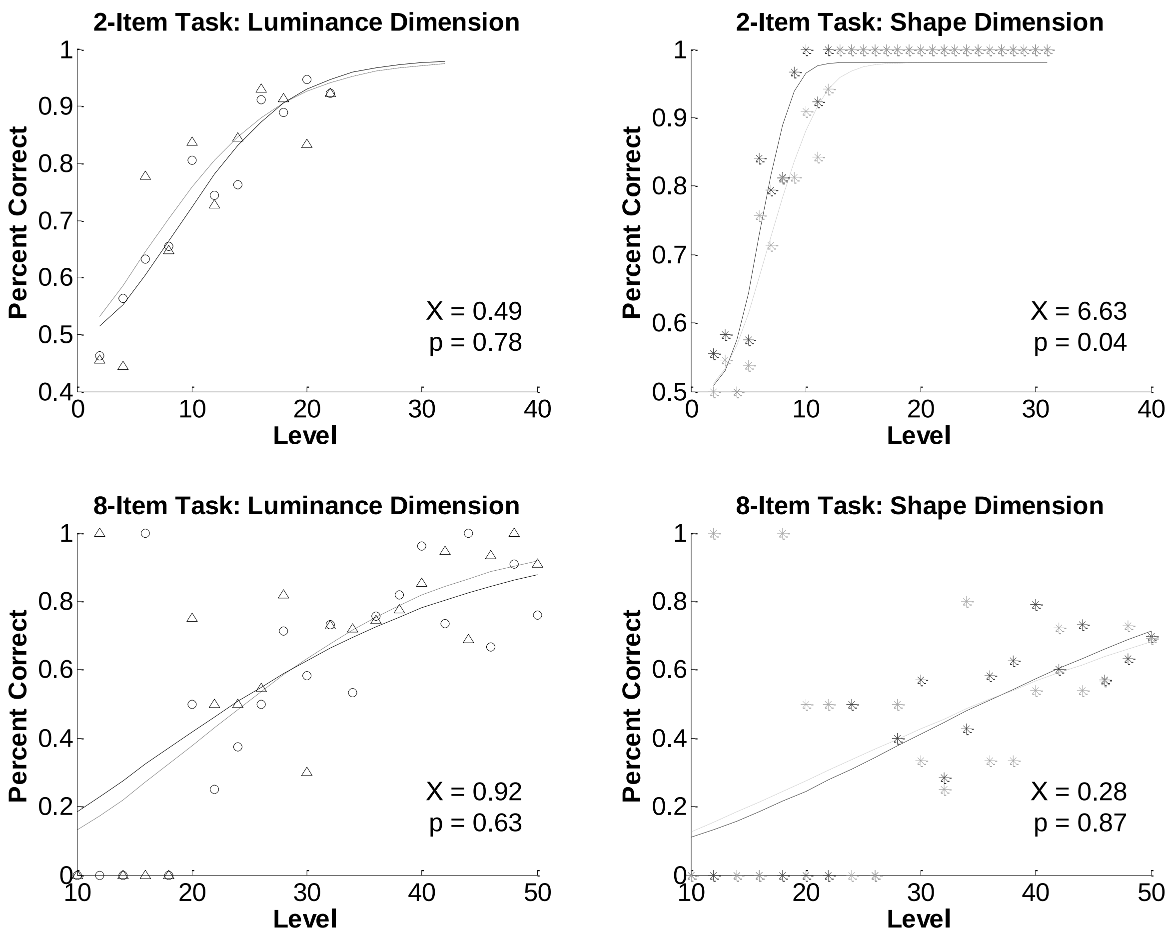


Figure 3.10: Comparison of Exp. 4 Conditions for Participant 3. The plots shows Participant 3's psychometric functions for each condition. In the Luminance dimension, circle markers show the participant's raw data for Condition 1 (circular items) and the solid line the line of best fit; triangle markers show the participant's raw data for Condition 2 (triangular items) and the dashed line the line of best fit. In the Shape Dimension, black markers show the participant's raw data for Condition 1 (darker items) and the solid black line the line of best fit; gray markers show the participant's raw data for Condition 2 (lighter items) and the dashed gray the line of best fit. We performed likelihood ratio tests for all conditions and give the chi-square score (X) for each. For this participant, the difference between conditions in the Shape dimension of the 2-item task is significant.

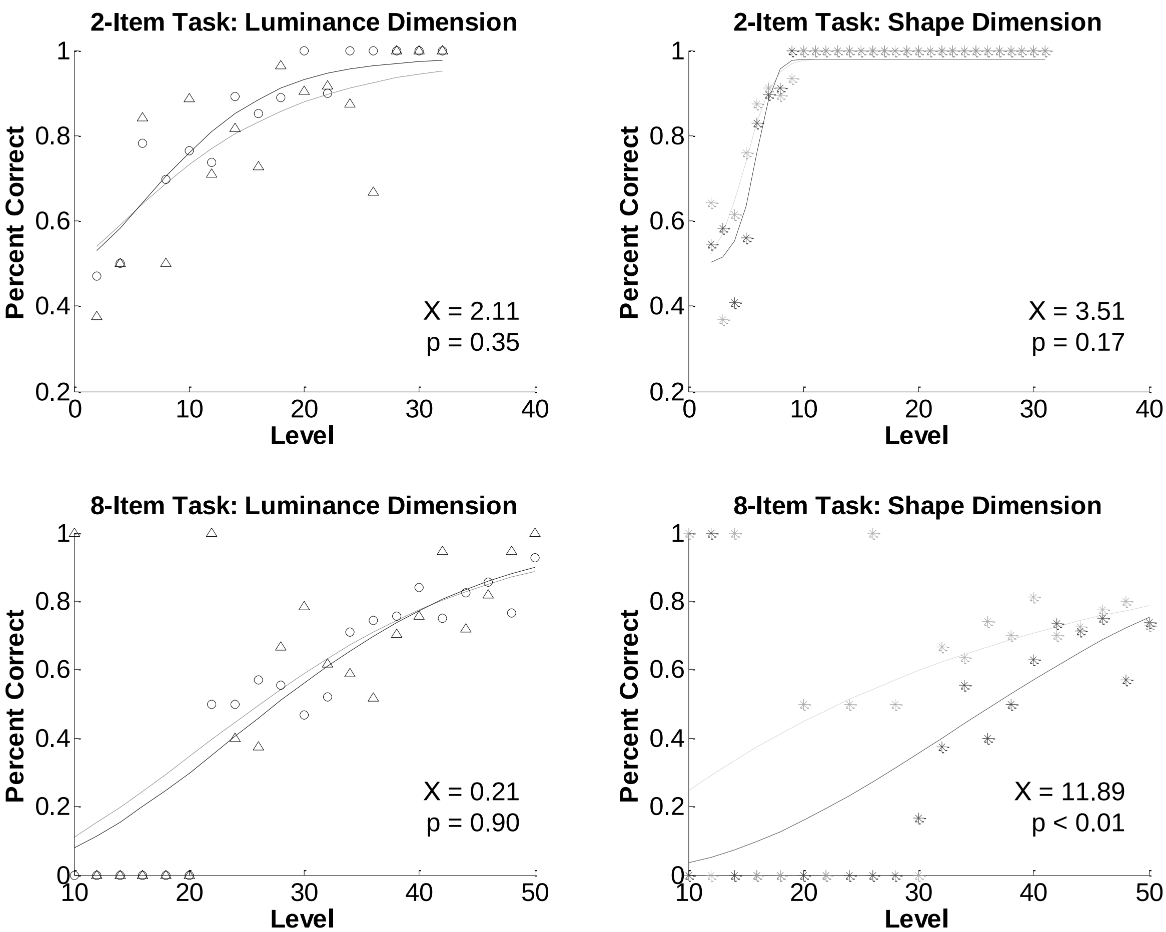


Figure 3.11: Comparison of Exp. 4 Conditions for Participant 4. The plots shows Participant 4's psychometric functions for each condition. In the Luminance dimension, circle markers show the participant's raw data for Condition 1 (circular items) and the solid line the line of best fit; triangle markers show the participant's raw data for Condition 2 (triangular items) and the dashed line the line of best fit. In the Shape Dimension, black markers show the participant's raw data for Condition 1 (darker items) and the solid black line the line of best fit; gray markers show the participant's raw data for Condition 2 (lighter items) and the dashed gray the line of best fit. We performed likelihood ratio tests for all conditions and give the chi-square score (X) for each. For this participant, the difference between conditions in the Shape dimension of the 8-item task is significant.

same parameters for some participants in some conditions; however, this was only the case for three comparisons (one each for three participants) out of the sixteen total comparisons (four participants x two tasks x two feature dimensions) and these results were idiosyncratic across participants.

Having established that there were no consistent differences across participants, we aggregated the data for each irrelevant feature dimension condition for our equisalience analysis. For each participant, we used an MCMC procedure to fit a Weibull function to the aggregated data. This produced 10,000 samples each of the α_L, β_L, α_S, and β_S parameters for each task. We cleaned the samples by discarding the first 2,000 and then keeping only every fifth sample to correct for a small degree of autocorrelation. From these samples, we used the median parameter values to fit an equisalience function for each task. We also estimated the credible interval for each task by fitting the equisalience function to the paired α and β parameters that fell between 2.5% and 97.5% of the distribution of the cleaned samples. Figure 3.12 shows each participants equisalience function for the 2-item task (solid line) and the 8-item task (dotted line), with luminance level presented on the *x*-axis and shape level on the *y*-axis.

Each point along the equisalience function corresponds to a stimulus level at which task performance is matched for two feature dimensions. Since the target level was fixed, stimulus level always refers to the level of the distractor. For both feature dimensions, low levels indicate more similarity between targets and distractors while high levels indicate less similarity between the two. Another way to think of the stimulus levels is in terms of the difference between the target(s) and distractor(s). For each participant, the solid line

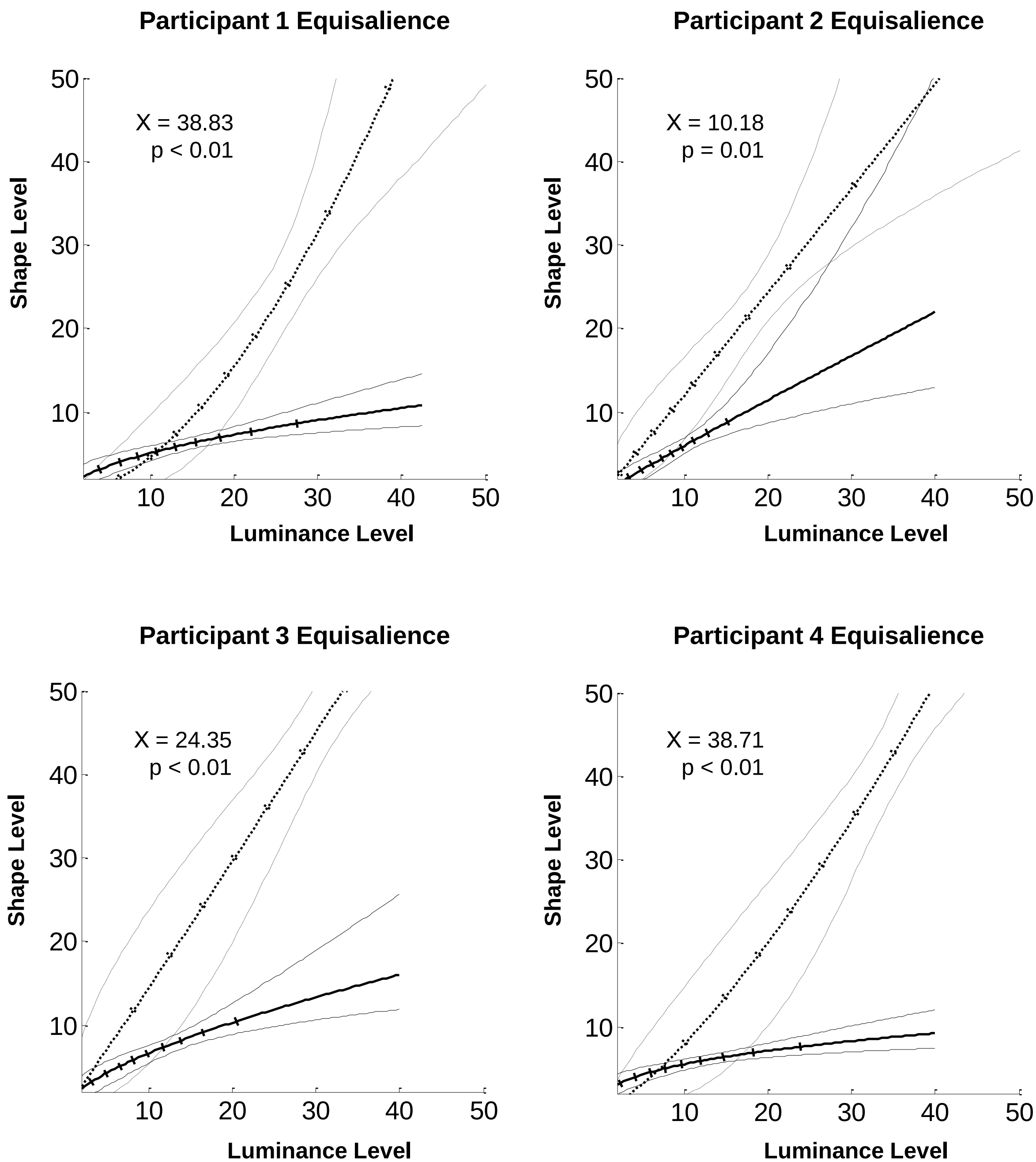


Figure 3.12: Exp. 4 Equisalience Plots. The Luminance-to-Shape equisalience plots are shown for each participant. The solid line represents the equisalience function for the 2-item task and the dotted line represents the equisalience function for the 8-item task. The tick marks indicate levels of performance, and the outer lines represent 95% confidence intervals. For each participant, the lines diverge, depicting statistically significant differences (X) in the Luminance-to-Shape equisalience of the two tasks.

(2-item task) and the dotted line (8-item task) lines diverge, indicating that the luminance-to-shape equisalience functions of the tasks are indeed different. In each case, the 8-item task required larger shape differences than did the 2-item task to match the performance for a given level of luminance. A likelihood ratio test confirms that these differences are highly significant.

3.4 DISCUSSION AND CONCLUSIONS

The different luminance-to-shape equisalience functions suggest that the two tasks either have access to different kinds of information or access the same information in different ways. Specifically, the equisalience functions suggest that shape differences—relative to luminance differences—have less salience in the 8-item task than they do in the 2-item task. This is consistent with the informal observation that motivated our present study: that the 2-item task failed to produce perceptually matched luminance and shape levels for an 8-item centroid task, and that the resulting centroid shape discriminations were much more difficult that the corresponding centroid luminance discriminations.

Given the similarities between the two tasks, their different equisalience functions are surprising indeed. We used the same procedure to assign item locations in both tasks; we displayed the stimulus clouds for the same amount of time in both tasks; participants tracked item location information and entered their responses via mouse click for both tasks. (It should be noted, however, that none of these task similarities are required to perform an equisalience analysis between tasks.) A priori, it would seem that the 2-item task and the 8-item should have access to exactly the same information in exactly the same way. And yet their strikingly different equisalience functions suggest this is not the case.

One possible explanation for our results is that crowding reduces the shape salience relative to the luminance salience in the 8-item task compared to the 2-item task. While plausible, more research is needed in order to determine how much of a role, if any, crowding plays in these differences. This could be tested using an equisalience analysis of two centroid tasks, each with a different number of items.

We believe it is an uncontroversial assertion that our 2-item task and 8-item task are much more similar than are a typical search task and centroid task. If two highly-similar tasks (2-item and 8-item) produce markedly different equisalience functions, we would predict that less similar tasks (search and centroid) would produce even more markedly different equisalience functions. Of course, this prediction needs to be tested, but we believe this will be another fruitful line of investigation. If the centroid task does, in fact, have access to different sorts of information than other feature-based attention tasks, that would help explain why centroid results can diverge so dramatically from typical search results.

CONCLUSION

Experiments 1 and 2 both demonstrate that centroid and search tasks can produce different patterns of results. In the search task, performance in all constituent-feature target conditions is better than performance in conjunctive target conditions. These search results can be summarized *Color* > *Size* > *Conjunction* (Experiment 1) and *Luminance* > *Shape* > *Conjunction* (Experiment 2). In the centroid task, however, performance in the conjunctive target conditions is intermediate between the constituent-feature target conditions. These centroid results can be summarized *Color* > *Conjunction* > *Size* (Experiment 1) and *Luminance* > *Conjunction* > *Shape* (Experiment 2). We believe these results provide two important, novel contributions to the literature. First, our centroid findings show that conjunctive targets can actually outperform constituent-feature targets. While there already exists evidence of efficient processing of conjunctive targets in some circumstances (e.g., Houck & Hoffman, 1986; Nakayama & Silverman, 1986; and Theeuwes & Kooi, 1994), there are no reported cases—to our knowledge—of conjunctive targets producing *better* task performance compared to constituent-feature targets. This finding becomes more surprising still when one considers that the conjunctive target displays had twice as many items the feature target displays, making them perceptually more complicated. Second, the differences in the centroid and search results provide compelling evidence that visual search is not, on its own, sufficient for investigating feature-based attention. A variety of tasks are needed in order to develop a richer understanding feature-based attention in a broad range of situations.

Experiment 3 shows that centroid performance is modulated by target-distractor similarity and indicates that, under some conditions, there may be no performance cost at all for conjunctive targets. We can borrow the inequality reporting convention used above to describe the pattern of results in Experiment 3. For consistency and ease of comparison to Experiment 2, we refer to the Targets=Darkest condition as *Luminance*, the Targets=Circles condition as *Shape*, and the Targets=DarkestCircles condition as *Conjunction*. We also introduce the "≥" symbol to denote performance differences that do not reach statistical significance. The results of Experiment 3 can then be characterized:

(i) *Luminance* ≥ *Conjunction* > *Shape* for ${\scriptstyle LS}$ display conditions,

(ii) *Conjunction* ≥ *Shape* ≥ *Luminance* for ${\scriptstyle L}S$ display conditions,

(iii) *Luminance* > *Conjunction* > *Shape* for $L{\scriptstyle S}$ display conditions, and

(iv) *Conjunction* ≥ *Luminance* > *Shape* for LS display conditions.

These results replicate and extend those of Experiment 2 by demonstrating that conjunctive centroid judgments can be better than or equal to *both* constituent-feature centroid judgments. They also suggest that targets defined by one feature dimension are not necessarily easier than targets defined by another because of the particular feature dimensions themselves; for example, attending to luminance targets is not always easier than attending to shape targets, as evidenced by the improved performance for shape target condition relative to luminance target condition summarized in (ii). Rather, these results suggest that task performance is influenced by the salience of the contrast between the levels of each feature dimension.

Experiment 4 reveals different luminance-to-shape equisalience functions for a centroid (8-item) task compared to another, highly similar (2-item) task. This suggests that the two tasks access information differently, thus lending support to the idea that there may be something special about the way in which centroids are processed. Specifically, it seems the centroid task requires greater shape salience (relative to luminance salience) compared to the analogous task.

Together, the results of these four experiments suggest that centroids may be processed differently than other sorts of visual stimuli. However, this line of investigation is still in its infancy, so we think it prudent to avoid making strong claims at this time about how centroids might be processed. These results also motivate many exciting future areas of research. In particular, it will be interesting to see whether or not other visual search results, such as search asymmetries, replicate in the context of the centroid paradigm. In addition, we anticipate the equisalience analysis will prove to be a useful tool for comparing the centroid task to other feature-based attention tasks—mostly notably, visual search. We also reaffirm the importance of studying feature-based attention in a range of different tasks in order to gain a fuller understanding of how features guide visual attention.

REFERENCES

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

Becker, S. I., Folk, C. L., & Remington, R. W. (2010). The role of relational information in contingent capture. *Journal of Experimental Psychology: Human Perception and Performance*, *36*(6), 1460.

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

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

Derrington, A. M., Krauskopf, J., & Lennie, P. (1984). Chromatic mechanisms in lateral geniculate nucleus of macaque. *The Journal of Physiology*, *357*(1), 241-265.

Egeth, H. E., Virzi, R. A., & Garbart, H. (1984). Searching for conjunctively defined targets. *Journal of Experimental Psychology: Human Perception and Performance*, *10*(1), 32.

Foster, D. H., & Ward, P. A. (1991). Asymmetries in oriented-line detection indicate two orthogonal filters in early vision. *Proceedings of the Royal Society of London B: Biological Sciences*, *243*(1306), 75-81.

Herrera, C. (2016). Psychophysics of Color Vison.

Houck, M. R., & Hoffman, J. E. (1986). Conjunction of color and form without attention: Evidence from an orientation-contingent color aftereffect. *Journal of Experimental Psychology: Human Perception and Performance*, *12*(2), 186.

Inverso, M., Chubb, C., Wright, C., Shiffrin, R., & Sperling, G. (2016). Comparing Efficiencies in Estimating Centroids and Judging Numerosity. *Journal of Vision*, *16*(12), 1308-1308.

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.

Kahneman, D., Treisman, A., & Gibbs, B. J. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive psychology*, *24*(2), 175-219.

Müller, H. J., & Rabbitt, P. M. (1989). Reflexive and voluntary orienting of visual attention: time course of activation and resistance to interruption. *Journal of Experimental psychology: Human perception and performance*, *15*(2), 315.

Nakayama, K., & Silverman, G. H. (1986). Serial and parallel processing of visual feature conjunctions. *Nature*, *320*(6059), 264-265.

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

Quinlan, P. T. (2003). Visual feature integration theory: past, present, and future. *Psychological bulletin*, *129*(5), 643.

Stockman, A., & Sharpe, L. T. (1999). Cone spectral sensitivities and color matching. *Color vision: From genes to perception*, 53-88.

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.

Theeuwes, J., & Kooi, F. L. (1994). Parallel search for a conjunction of contrast polarity and shape. *Vision Research*, *34*(22), 3013-3016.

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive psychology*, *12*(1), 97-136.

Treisman, A. (1982). Perceptual grouping and attention in visual search for features and for objects. *Journal of Experimental Psychology: Human Perception and Performance*, *8*(2), 194.

Treisman, A. (1985). Preattentive processing in vision. *Computer vision, graphics, and image processing*, *31*(2), 156-177.

Treisman, A. (1988). Features and objects: The fourteenth Bartlett memorial lecture. *The quarterly journal of experimental psychology*, *40*(2), 201-237.

Treisman, A., & Sato, S. (1990). Conjunction search revisited. *Journal of experimental psychology: human perception and performance*, *16*(3), 459.

Winter, A. N., Wright, C., Chubb, C., & Sperling, G. (2016). Conjunctive Targets are Hard in Visual Search but Easy in Centroid Judgments. *Journal of Vision*, *16*(12), 750-750.Wolfe, J. M. (2007). Guided search 4.0. *Integrated models of cognitive systems*, 99-119.

Wolfe, J. M., Cave, K. R., & Franzel, S. L. (1989). Guided search: an alternative to the feature integration model for visual search. *Journal of Experimental Psychology: Human perception and performance*, *15*(3), 419.

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

Wolfe, J. M., & Gancarz, G. (1997). Guided Search 3.0. In *Basic and clinical applications of vision science* (pp. 189-192). Springer Netherlands.

Wolfe, J. M., & Gray, W. (2007). Guided search 4.0. *Integrated models of cognitive systems*, 99-119.

Wright, C. E., Chubb, C., Winkler, A., & Stern, H. S. (2013). Equisalience analysis: A new window into the functional architecture of human cognition.

Zhou, X., Chu, H., Li, X., & Zhan, Y. (2006). Center of mass attracts attention. *Neuroreport*, *17*(1), 85-88.