ATTENDING OVER SPACE AND OBJECTS

by

Larence Becker

A dissertation submitted to The Johns Hopkins University in conformity with the requirements for the degree of Doctor of Philosophy

Baltimore, Maryland

May, 2001

© Larence Becker. 2001

All Rights Reserved
Abstract

In the research literature, there are widely varying accounts of the effects of spatial separation on attentional tasks. Using different methods, researchers have determined that, performance improves for attentional shifts over shorter distances, performance is unaffected by the distance between attended items, and that performance is impaired when attended objects are near each other. The experiments in the first part of the dissertation addressed this issue using a paradigm in which subjects identified two target letters, while inter-target separation, distractor presence, and distractor category were varied. The major findings of these experiments were 1) performance improved as the separation between target letters increased, and 2) distractors from the target category created additional interference at short target separations. These findings are discussed in terms of models of attentional suppression and imprecise targeting (Bahcall & Kowler, 1999). The final two experiments examined how the spatial relationship between two attended objects influences the interfering effects of nearby distractors. When distractors and target letters were all equidistant from fixation, performance was the same whether the distracting letters were between, or just exterior to the target letter pair. In the final experiment, there were always two distractor-target-distractor groups; these two groups were either arranged to form two parallel groups of three across the display, or the groups were collinear, with the distractors being interior and exterior to the target pair. (Retinal eccentricity of all elements was controlled for between conditions.) Target identification performance was worse for the collinear condition than the parallel condition. It is
hypothesized that attention is oriented along the axis connecting the attended locations and extending beyond them.

Dissertation Advisor: Dr. Howard Egeth
Acknowlegdments

At a graduate school interview at the University of Pennsylvania, Dr. Henry Gleitman asked me with whom I was thinking of working at other schools. I mentioned Dr. Egeth as a likely candidate. In response, Dr. Gleitman said, “Howard Egeth; he’s a good man.”

Dr. Gleitman was right. It has been a total pleasure working with my advisor, Dr. Egeth, and I thank him for making my years at Hopkins as rewarding and fun as they have turned out to be. He always has time for his students, and his mind is wide open to any new ideas they might have. What more could one ask?

I also offer thanks to my other defacto mentor in vision and attention, Dr. Steven Yantis. I loved being a teaching assistant for Dr. Yantis, as I’m sure I learned more than any of the undergrads in the courses he taught. In addition, his intellect has inspired me on many a weary Friday morning.

Thanks are also due to Drs. Trish Van Zandt, Gregory Ball, and Brenda Rapp for their encouragement, friendship, and excellent teaching. In regard to all of the Hopkins faculty, I feel I have been treated as a peer from the moment I arrived. Thanks.

From NYU, I thank Drs. Sherry Busbee, Orn Bragason, Steven Fried, and Stacey Layman, for lighting the fire, and then convincing me I could do it!

For miscellaneous keys and papers and paychecks and friendly reminders, not to mention, general helpfulness, I thank the Hopkins Psychology Department Support staff: Dru Horne, Joan Krach, Anne Daly, Debbie Grover and Terri Dannettel.
During my years here I’ve been fortunate enough to become friends with the cohort of students that preceded me, and those who came years after I started:

Of those who came before me I thank Tom Ghirardelli, Tim Grandison, Sharma Hendel, Dave Scrams, and Barry Vaughan for allowing a mere first year to join you for lunch, and for providing many a laugh when I should have been working.

For many shared gallons of coffee and hours of guy-talk I thank Derek Houston, my first great coffee-buddy, and Joseph Sala, who has ably filled Derek’s mugs- etc. shoes.

And, over the last few years, it just wouldn’t have been lunch without Sarah Shomstein, Elizabeth Johnson and Robert Rauschenberger. You made me hunger for 12:00...and it wasn’t because of the food at Levering! Do not forget the lessons I have taught you.

Thanks to Ruth Tincoff for friendship, movie nights, many a lunch, and for general commiserating. And to Andy Leber, who I’m so grateful I convinced to come to Hopkins; for hours of philosophising, BS’ing, and guitar playing. (I would never have played in public without his inspiration!)

Finally, I am most thankful for the support of my parents, Seymour and Rosalie Becker, and my brother, Jeff Becker, and my sister, Michele Davidson, and their families. I couldn’t have gotten this far without all of them. This is for them as much as for me. Especially dad—the Professor—my first great intellectual mentor.

And, I thank Sunny Sipes. For late night calls, meeting me at the station, and everything in between. SunnySunnySunnySunny. There. I said it.
# Table of Contents

Abstract .......................................................................................................................... ii

Acknowledgments ........................................................................................................ iv

List of Tables ................................................................................................................ viii

List of Figures ............................................................................................................... ix

Introduction .................................................................................................................. 1

Chapter 1: Perceiving a Single Object Among Others .................................................. 3

  Crowding and Masking ............................................................................................... 3

  Kinds of Perceptual Interference .............................................................................. 7

  Attention and Visual Search .................................................................................... 21

  Theories of Visual Attention .................................................................................. 22

  Perceptual Organization and Attention ................................................................. 29

Chapter 2: Processing Around an Attended Object ..................................................... 39

  Scope of the dissertation .......................................................................................... 54

Chapter 3: The Experiments, Part I ........................................................................... 61

  Experiment 1 ............................................................................................................ 61

  Experiment 2 ............................................................................................................ 74

  Experiment 3 ............................................................................................................ 86

Chapter 4: The Experiments, Part II .......................................................................... 94

  Experiment 4 ............................................................................................................ 94

  Experiment 5 ............................................................................................................ 104

General Discussion .................................................................................................... 114
References...................................................................................................................... 123
Appendix: The PEST Procedure.................................................................................... 137
  The PEST Procedure ................................................................................................. 138
  PEST in the Current Experiments ............................................................................ 143
VITA.............................................................................................................................. 146
List of Tables

Table 1. Experiment 1 accuracy, by experiment condition.............................................67

Table 2. Experiment 5 accuracy, by experiment condition.............................................109
List of Figures

Figure 1. Two stimulus displays from study by Banks and Prinzmetal (1976) ..................38
Figure 2. Stimuli based on Enns and Rensink (1990)..................................................38
Figure 3. Stimuli configuration and procedure. Bahecall and Kowler (1999)..............59
Figure 4a. Difference of Gaussians...........................................................................60
Figure 4b. Sum of two differences-of-Gaussians.........................................................60
Figure 5. Stimuli and procedure for Experiment 1 ......................................................71
Figure 6. Experiment 1 data for all subjects combined ..............................................72
Figure 7a. Experiment 1 data: 24-letter (High density) condition, by subject ..........73
Figure 7b. Experiment 1 data: 2-letter (Low density) condition, by subject ..........73
Figure 8. Stimuli and procedure for Experiment 2 .....................................................83
Figure 9. Experiment 2 data for all subjects combined ..............................................84
Figure 10a. Experiment 2 data: letter-distractor condition by subject .....................85
Figure 10b. Experiment 2 data; number-distractor condition, by subject...............85
Figure 11. Experiment 3 stimuli ...............................................................................92
Figure 12a. Experiment 3 data; all subjects combined ............................................93
Figure 12b. Experiment 3 data, by subject.................................................................93
Figure 13. Experiment 4 stimuli ...............................................................................103
Figure 14. Experiment 4 data.....................................................................................103
Figure 15. Experiment 5 stimuli ...............................................................................113
Figure 16. Experiment 5 data, by subject.................................................................113
Figure 17. Example of a PEST run ........................................................................145
Introduction

As with any researcher, I am often asked by friends or family members what it is I study. Regarding this present project, I tell the interested party that I am investigating the question, How is one’s ability to identify two objects in a scene affected by the distance between them? I then explain that in one of my experiments, subjects view briefly presented displays and I measure their ability to identify two target objects as a function of spatial separation. It is at this point that I usually get that certain quizzical look. Not because they don’t understand, but because they do, and in their mind, this is a question whose answer seems obvious. With just slight hesitation they usually then state that, why, of course, the objects are easier to identify when they are closer together. To the naïve observer, the answer seems intuitive and clear.

Common sense tells us one thing about attending to two objects, but is this really the truth? People often (mistakenly) believe that attending to an object implies fixating it. Thus, when I explain my research question, it is often assumed that one, if not both, of the objects is (are) fixated at some point. If one object is fixated, then yes, an object nearer this fixated object will be better seen than a farther one, because visual resolution is better near fixation (Goldstein, 1996). However, attention and fixation can be disentangled. We are all familiar with watching someone “out of the corner of our eye”. This act suggests that we can covertly attend to something, without gazing right at it. In fact, this kind of covert attending is something we do all the time. While driving, one doesn’t look directly at all surrounding cars and signs; gaze is (largely) straight ahead, although we
attend to the car in the neighboring lane, or the stop sign coming up on the right, etc. Thus, not only can we attend covertly, but we can do so to multiple objects at a time. (Whether we attend simultaneously or by shifting back and forth is a question for another dissertation.) When attention is viewed in this light, the original research question becomes a bit more problematic than it may have seemed at first.

As alluded to, this dissertation is a study of attention to two objects. Within this major topic, there are two questions guiding the investigation: How is the ability to identify two attended objects affected by the distance between the objects? And, how is the ability to identify two attended objects affected by other irrelevant objects in the visual scene?

The plan for the dissertation is as follows: To begin, the issue of perceiving two objects in a scene is approached in the literature review from two different perspectives: the first chapter looks at how perception of a single item is influenced by the placement of other items around it; I refer to this as the outside-in view of object perception. The second chapter examines how attending to a single item effects perception of items around it; I refer to this as the inside-out view of object perception. Following the literature review, five experiments are described which investigate the research questions posed above. The findings are further interpreted in the General Discussion. In addition, an appendix is included describing the adaptive procedure, Parameter Estimation by Sequential Testing (or PEST; Taylor & Creelman, 1967) used in a number of the experiments.

(Note that figures appear at the end of the chapter or experiment in which they are cited.)
Chapter 1: Perceiving a Single Object Among Others

Crowding and Masking

First, to clarify some terminology: when identification of a target is impaired by the presence of nearby distractor elements, this is variously called “lateral masking” or “crowding”. Technically, however, these are not the same. Lateral masking is part of the larger class of phenomena called “visual masking”. In his definitive book on the topic, Breitmeyer defines visual masking as “the reduction of the visibility of one stimulus, called the target, by a spatiotemporally overlapping or contiguous second stimulus, called a mask.” (Breitmeyer, 1984, p. 2). This general definition encompasses a range of conditions under which masking has been found to occur: Temporally, the masking element may precede the target (paracontrast, or forward masking), follow the target (metacontrast, or backward masking), or co-occur with the target. Spatially, the mask and target may occupy the same location, or, be proximal to each other. Thus, this dissertation deals with the special case of simultaneous and neighboring presentation of target and mask.

Breitmeyer proposes a model of visual masking based on the working of both center-surround antagonism within channels and mutual and reciprocal inhibition between channels. According to Breitmeyer’s model, masking results from interference occurring at sensory levels of visual processing.

How does lateral masking differ from crowding? Crowding describes a scene in which a number of objects are presented simultaneously with, and in the vicinity of, a
target item. The crowding objects could be near the target—producing lateral masking—but other objects contributing to crowding could be some distance (e.g., a few degrees of visual angle) from the target. Difficulty in identifying the middle letter in the string of letters that follow below would be an example of interference caused by crowding:

Fixate here  $\Rightarrow$  +  d k b h k t l

Compare with  +  b h k

Or,  +  h

(The target is the same distance from fixation in all examples.)

Note that in the first (or crowding) example, the difficulty in identifying the middle letter is not merely perceiving it, but in individuating, or separating it from the crowd.

In a recent study, He, Cavanaugh, and Intriligator (1997, Exp. 1) looked at crowding and its relation with lower level visual processes. Displays consisted of a round grating shown either alone or as the fourth of five such gratings (counting away from fixation) in a vertical traffic light arrangement placed above or below fixation. The gratings could be oriented at 45° or -45°. Subjects were much better at identifying the orientation of the target grating when it was alone than when it was flanked by other gratings.
In the next part of the study, the authors assessed the level of visual adaptation to a grating; visual adaptation being a characteristic associated with fatigue of low-level visual areas. Adaptation was determined by measuring the degree to which the threshold to identify the orientation of a grating at a particular location was raised following pre-exposure to a grating at that location. There were two pre-exposure conditions: a grating could be shown at the critical location either alone, or along with four other gratings in the arrangement described above. For both the lone grating and the grating in a crowd, the level of adaptation was very similar. He et al.'s finding suggests that under crowding conditions, the visual system is able resolve the grating's orientation (producing similar adaptation effects as for the lone grating conditions), even when attention could not (as revealed by behavioral measures of accuracy which differed between conditions). Thus, a subject's inability to discriminate the grating's orientation in the crowded display was not produced by lateral masking, which, according to Breitmeyer (1984), is a lower level vision phenomena.

Finally, Palomares and Pelli (submitted) explicitly demonstrated that (as they claim) "crowding is unlike ordinary masking". In this study, a single target letter (when present) was displayed in the periphery (4° visual angle from fixation), either along with two other flanking letters, or covered in part or whole by noise. Thresholds for both target detection and identification were determined under various manipulations of stimulus size, target-flanker separation and contrast.1 The letter flankers condition served

---

1 The detection experiment used a 2-interval forced choice paradigm, in which only one interval contained a target.
as a gauge of crowding effects while the noise condition served as a gauge of masking effects. Qualitatively different patterns of data were produced for each of these conditions: under crowding conditions, effects of target-flanker spacing were independent of the size of the signal and mask; i.e., they are determined by stimulus eccentricity; in contrast, spacing effects under masking conditions scaled with the size of the stimuli. In addition, crowding was found to affect letter identification but not detection, whereas masking affected both identification and detection. This latter finding once again points to masking and crowding as resulting from interference at different stages of visual processing, with masking associated with interference at earlier stages and crowding associated with interference at later stages.

The purpose of the preceding discussion was to categorize two different types of interference effects that can occur when objects are placed near each other. To summarize, in lateral masking there is interference in perceiving a target element contiguous with a distractor (e.g., another object or a noise mask), resulting from low-level visual interactions. Whereas, in crowding, individuation and identification of a target element is impaired when the target is in the vicinity of one or more distractor items, sharing some characteristic with the target; crowding is likely modulated by higher level (e.g., attentive) processes. Though it is not always clear how to assign each specific finding into this dichotomy, the terminology does provide us with a consistent vocabulary for describing the effects shown in the studies described below.

Note that this dichotomy is in close agreement with that put forth by Wolford and Chambers (1983, discussed later in dissertation). These authors argued that lateral
masking is composed of two processes: spatial interaction being a low-level or hard-wired process, which is manifest at small separations, and attention, which operates at a later level and at larger separations. For consistency, throughout this dissertation, I will refer to the low-level process as lateral masking and the higher level process as crowding.

Kinds of Perceptual Interference

The following review begins with a discussion of perceptual interference between two simple contours. Moving to more complex objects, the subsequent section examines studies of masking and crowding among letter targets and distractors. Finally, we look at the effect of grouping among objects on the perception of a letter target.

Flom and colleagues (Flom, Weymouth & Kahneman, 1963: Flom, Heath & Takahashi, 1963) performed studies on what they called “contour interaction”. In both studies, subjects were shown a Landolt C and four surrounding bars. If connected, the bars would have formed an upright square around the C (the ends of the bars were not connected, though). The gap forming the C could be at three, six, nine or twelve o’clock—always directly opposite a bar. The width and location of the gap in the C were varied, as was the distance of the bars from the C. Subjects had to determine the orientation of the C (that is, was the gap up or down or left or right?). When the display was shown monocularly, performance was worst when the bars were about two gap lengths from the C. Interference caused by the bars diminished as they were placed farther from the C, with no masking effect past five gap lengths. Flom, Heath & Takahashi, 1963, also used a condition in which the bars were displayed to one eye and
the C was displayed to the other eye. No difference was found in the pattern of performance between the binocular and monocular conditions, suggesting that the crowding effect is supraretinal: it occurs at a point in visual processing after information from the two eyes has been integrated.

Using very simple stimuli, Flom and colleagues had shown that the presence of a thick line could interfere with the visibility of a nearby contour. Similar interference effects have also been found by researchers using other types of stimuli as targets, e.g., tilted lines (Andriessen & Bouma, 1976; Westheimer, Shimamura and McKee, 1976), and vernier-type lines (Westheimer & Hauske, 1975; Levi, Klein & Aitsebamo, 1985). Note that for all the cited studies, except that of Levi, et al., stimuli were presented at fovea. Levi, et al. manipulated the target eccentricity and found that as this factor increased, so did the target-flanker distance required to produce maximum interference. The authors hypothesize that crowding effects result from multiple stimuli impinging on a single, or adjacent, cortical ocular dominance column(s), with interference, hence, scaling with cortical magnification.

What kinds of masking effects occur between contours of more complex stimuli? Bouma (1970) examined the lateral interactions between adjacent letters. On each trial, a target letter was briefly presented (200 ms), flanked by either 0, 1 or 2 distractor letters. Target eccentricity was also varied. At all target eccentricities, performance was best for the 0 distractor condition, and worst for the 2 distractor condition. The distance between target and flankers was then manipulated in the 2-flanker condition. For performance to equal that of the no-flanker condition, a target at $x^\circ$ (visual angle) eccentricity generally
had to be at least x°/2 from its nearest flanker. This scaling factor is again reflective of
cortical magnification of visual receptive fields, which grow larger with peripheral
location. In short, Bouma had found that the more sides of a letter that are flanked with
other letters, the more difficult the letter is to identify; and the effect of these flankers
decreases with distance from the target.

Bouma demonstrated how one or two flanking neighbors could interfere with
letter perception. Taylor and Brown (1972) investigated the visibility of letters in long
strings of other letters. Under binocular viewing conditions, subjects were given
unlimited viewing time to identify letters in strings of 5, 6, 7 or 9 letters, presented to the
right of fixation. Identification accuracy was highest for letters near fixation, and
diminished with letter eccentricity. However, there was always an upswing in
performance for the letter at the far end of the string as compared with the penultimate
letter in the string, resulting in a U-shaped function describing subject identification
accuracy. In one display condition, the 9-letter strings were presented dichoptically, with
alternating letters presented to each eye (e.g., letters 1-3-5-7-9 to the left eye and letters 2-
4-6-8 to the right eye.) Dichoptic presentation was compared with monocular
presentation, in which all stimuli were presented to one eye. Letter-detection accuracy
was similar for dichoptic and monocular presentation. This result concurs with Flom,
Heath, and Takehashi’s (1963) finding that masking (or, more accurately, crowding)
effects may occur at a supraretinal, or more central level of processing.

More recently, Strasburger, Harvey, and Rentschler (1991, Experiment 3)
investigated crowding in a study in which target-flanker separation, target eccentricity,
and target size were varied. Target digits were presented along with two flanking digits along the horizontal meridian for 100 ms. With targets at 4° eccentricity, thresholds to identify the central target digit reached asymptote when flankers were about 1.2° from the target (measured from digit center to digit center). The (approximate) same point of asymptote was found regardless of stimuli size (0.4° to 1.0°). Thus, the degree of interference caused by crowding was determined largely by the absolute separation between target and flankers, not by the relative separation, i.e., as determined by stimulus size (see also Palomares & Pelli (submitted) for a similar finding).

In their Experiment 4, Strasburger et al. varied the number of flankers: there were either two flanking items, as in the experiment already described, or, there were four flankers, with two on each side of the target. The size of the crowding effect was found to increase with the added masks. This again suggests how crowding is different from lateral masking, in that the level of interference was augmented by objects not adjacent to the target. Note that an explanation that lateral masking effects may have propagated from the farther flankers, past the inner flankers to the target does not hold; the experimenters had earlier shown no such effects of flunker interference for single flankers presented at the same separations as these farther flankers. It would appear that merely the increased numerosity, and, hence, greater local confusion in individuating the target from the distractors made the task more difficult.

In the studies of object and letter identification discussed thus far, the factors which were manipulated included target eccentricity, the number of distractors, and target-distractor separation distance. However, the nature of the stimuli themselves has
also been found to modulate interference between neighboring objects. In the following section I will review research that has looked at how masking effects are influenced by characteristics of the stimuli in a display.

Kooi, Toet, Tripathy and Levi (1994) examined how physical differences between targets and flanking objects influence masking. Experimental displays consisted of a target letter T surrounded on four sides by four other T's on a gray background. T-stimuli could be black or white, and were presented at any of four orientations (T rotated by some multiple of 90°). Displays were centered at 10° eccentricity from fixation, and were presented for 150 ms. Subjects' task was to identify the orientation of the target T. Data indicated that as the similarity between the target and distractors increased, performance accuracy decreased. Specifically, performance was most impaired when target and flankers shared features along one of the following dimensions: polarity (i.e., both white or both black), depth, contrast, or shape (in the shape manipulation, flanker T's were slightly altered to produce a different shape). (An increase in interference with target-flanker similarity has also been reported by Polat & Sagi, 1993, who used oriented gabor-patches as targets, and Leat, Li & Epp, 1999, who used rotated C’s and E’s as targets).

Kooi et al. relate their findings to the coding of object position in vision, perceptual grouping and preattentive pop-out (see dissertation section on Theories of Attention). As the target and flankers become more and more alike, they tend to group together. In turn, the greater the degree of target-flanker grouping, the less will the target pop-out from the display, which correspondingly, reduces the ability to physically locate
the target in space. All of which combine to make target processing (identification, etc.) more difficult.

Kooi et al. propose that a single mechanism can explain these phenomena. When similar objects (i.e., those sharing a particular feature) are neighbors, processing of these objects is mutually inhibited by the presence of the others (see also Duncan & Humphreys, 1989). A cluster of elements with a shared feature could then be segregated based on their common inhibition. A target within such a grouping would thus suffer from masking.

Meanwhile, an object that is different from its neighbors is released from this inhibition, and its associated masking. The dissimilar element would thus receive a processing benefit compared with the processing of grouped elements nearby (whose processing is inhibited), and if it is sufficiently different, the target could visually pop-out from the display.

The notion that objects having similar features produce mutual inhibition in their processing was actually put forth (though, in somewhat different terms) as early as the 1970’s. At this time a number of researchers were investigating various aspects of interference, which occurred while processing letters presented together. A classic study of this type was conducted by Bjork and Murray (1977). In their experiments, displays, which were briefly presented (25-50 ms), were a 4 x 4 matrix, consisting of 2 letters and 14 pound signs (#). Only one letter was identified on a trial, and it was specified to the subject via a post stimulus cue (an arrow pointing to one of the columns). The target and mask letter could be related in any of the following ways: they could be: 1) the same
letter, 2) a different letter having a competing response, 3) a different letter, having no associated response (i.e., a letter that never served as a target) or, 4) in a final condition, one letter and 15 pound signs were presented. Accuracy in target-identification was poorest when the noise letter was the same as the target, and best for the single target condition. (Note that Santee and Egeth, 1982, found this “repeated letter inferiority effect” for masked, but not unmasked, displays). Bjork and Murray also found that interference decreased as separation between target letter and noise letter increased.

A number of different models were proposed to explain the interference effects found in letter processing. The models of Estes (1972) and Bjork and Murray (1977) attributed these effects to interactions occurring at the level of visual feature extraction (as opposed to a later cognitive or decision level). Based on their findings, Bjork and Murray developed the feature specific inhibitory channels model. According to the model, signal detection processes (e.g., letter identification) rely on the presence of feature detectors, which are hierarchically organized and have a limited capacity. Feature detectors are hierarchically organized in that they are tuned for various levels of stimulus complexity, with the more complex detectors having larger receptive fields. (The authors draw a comparison to the system of receptive fields in vision described by Hubel and Wiesel, 1965, 1968.) When a particular detector is active (i.e., having received an appropriate matching input), other inputs to the same detector are inhibited—this mechanism prevents a limited capacity feature detector from being overloaded. Interference occurs to the extent that feature detectors are shared by signal and noise items, with sharing being a function of both the similarity between the signal and noise
and the distance between them. Thus, the more similar and proximal one letter is to a second letter, the more perceptual interference there will be between them.

Egeth and Santee (1981) used stimuli similar to those of Bjork and Murray, but also included a condition in which a capital letter (e.g., A) was presented with its lower-case counterpart (e.g., a) as a mask. This same-letter/lower-case mask condition produced poorer performance in target identification than when a different capital letter served as a mask. This result obviously cannot be explained in terms of inhibition at the feature level. Based on ideas from Walley and Weiden (1973), it was hypothesized that letter processing occurs within pattern-recognition networks, proceeding hierarchically from low-level feature detectors through high-level recognition channels. Interference within these higher level “gnostic units” (i.e., as when multiple similarly named stimuli are encoded simultaneously) produces what is referred to as cognitive masking.

Paradoxically, some researchers have found a benefit in processing similar neighboring letters. Eriksen and colleagues (Eriksen & Hoffman, 1972, 1973; Eriksen & St. James, 1986) have consistently found that a speeded response made in a target-identification task is delayed when the target is flanked by letters having a competing response as compared to responses made when the target is flanked by letters having a compatible response (e.g., the same letter). Interference on these tasks is ascribed to decision or response level processes. As these processes occur after stimulus identification, performance decrements arising for identification of a target paired with an incompatible flanker do not result from an actual decrease in letter visibility, but,
rather, because of the response competition between the neighboring letters. (See dissertation Chapter 2 for a further discussion of flanker studies.)

The predictions of the response competition model would seem to directly contradict those of the feature specific inhibitory channels model, which claims that identification of a letter would be impaired by the nearby presence of a similar letter. Each model, however, describes interference occurring at a different level of processing, and therefore, can account for measures that reflect a particular processing level. For example, Bjork and Murray (1977) were able to confirm both models in a single study: compared with the condition in which targets and flanking letters were dissimilar, when targets were the same as (i.e., compatible with) flanking letters, identification accuracy was lower, but corresponding response times were faster. The accuracy data are explained by feature specific interactions early in processing, while the response time data were in accord with response competition occurring at a later decision stage of visual processing.

The studies and models described above describe the general interference effect that occurs when identifying a letter placed nearby other letters. I will next discuss how the spatial arrangement of stimuli influences similar tasks. In a psychophysical study, Toet and Levi (1992) mapped out what they called zones of spatial interaction: these are areas around a target which when occupied, produce equal levels of target masking. Stimuli were three T-shapes—a target and two flanking masks equidistant from and collinear with the target. Four factors were varied: target eccentricity, target polar angle from a central fixation, mask polar angle around the target, and target-mask distance.
Results showed that when the target and masks were arranged along an imaginary radial line emanating from fixation, masks could be placed farther from the target to achieve a given level of masking than when the target and masks were arranged isoeccentrically (all three at the same distance from fixation). At all eccentricities tested (2.5°, 5.0° and 10.0°), the resulting zones of spatial interaction were thus elliptical with the longer axis oriented toward fixation.

In the Toet and Levi (1992) study the distances from a target to each flanker were always equal (i.e., they changed together). As a result, when the trio were arranged radially with fixation, no difference could be distinguished between the effect produced by the more versus less peripheral mask. Chambers and Wolford (1983), fortunately, had explored the issue of relative flanker eccentricity nine years earlier. That study used a briefly presented display (100 ms) of two letters: a target (either F, C, U or T) and a flanking character (always H). The pair could appear at any of the four cardinal compass locations around the display, always separated by .43°, center-to-center. The flanker could be positioned either directly above, below, to the left or to the right of the target letter. The distance from the centerpoint of the pair of letters to fixation was kept constant (5.0°). In agreement with Toet and Levi, Chambers and Wolford had found that a flanker placed along an imaginary radial line passing from fixation through the target had a greater masking effect than a flanker isoeccentric with the target. Regarding the role of relative eccentricity of the flanker, performance in identifying the target letter was worst when the flanker eccentricity was greater than that of the target—that is, when the flanker was on the target’s outside, relative to fixation. This suggests that Toet and
Levi’s elliptical zones of spatial interaction would more accurately be depicted as egg-shaped, with the flattened section toward fixation.

One explanation for this masking asymmetry (i.e., greater interference from a mask with greater eccentricity versus lesser eccentricity than a target) was offered by Wolford’s (1975) feature perturbation model. According to this model, there is a certain probability that the represented features of neighboring letters will perceptually migrate or be mislocated by the observer over time. Interference occurs when features of a noise item combine with features of a target, leading to a failure of identification. The model also asserts that there is a bias for features to drift foveally rather than in any other direction. This final premise explains why masking would be greatest for noise elements placed peripherally from a target, as features from these masking positions have the greatest likelihood of drifting into a target’s perceptual space.

A competing hypothesis was offered by Banks, Larson and Prinzmetal (1979). These researchers explained lateral masking in terms of Gestalt grouping between adjacent letters. The target-mask pair, they claimed, is perceived as a group, from which the identities of the particular elements must later be extracted. To explain why peripheral flankers cause a greater detriment in identification performance than central flankers, they appealed to Beck’s (1972) finding that perceptual grouping is more prone to occur in the periphery than in the center of the visual field. Banks et al. reasoned that for a given target eccentricity, there is a higher likelihood for the target to group with a peripherally placed flanker than with a foveally placed flanker. Thus, the target and flanker are easier to individuate the nearer their centroid is to fixation because grouping
between them is less likely; as the pair move farther from fixation, grouping between them becomes more likely, and identification of each letter becomes more difficult.

Banks, Larson and Prinzmetal (1979) tested the grouping hypothesis using displays in which the target was 5° either to the left or right of fixation, and was 1° from the nearest flanker. There were four conditions: In two of the conditions there was a single flanker letter, placed either on the peripheral or foveal side of the target. In the other two conditions, the flanker (placed as just described) was in the middle of a 5 letter vertical stack of letters (all the same letter, H). Below are examples of these arrangements; note that on half the trials the flankers were closer to fixation than the target, and that stimulus sets could appear on either side of fixation.

**Single Flanker Arrangement**  

\[
\begin{array}{cccc}
 F & T & + \\
 F & F & F \\
\end{array}
\]

**Stacked Flanker Arrangement**  

\[
\begin{array}{cccc}
 F & T & + \\
 F & F & F & F \\
\end{array}
\]

(T = target location,  F = flanker/mask location, ‘+’ = fixation)

Banks et al. predicted that the strong grouping among flanker elements (based on their similarity) should “release” the target from the target-flanker group and increase the target’s perceptibility as compared with the single flanker conditions. They also predicted that the effect of grouping among the flanker items should be greater for the peripheral mask condition, because grouping is stronger at the farther distance (based on Beck’s finding). Their data supported both of these predictions: target identification accuracy was better in the grouped-flanker condition than the single-flanker condition for each
flanker eccentricity (peripheral or foveal, relative to the target). In addition, the improvement in accuracy for the grouped-flanker condition over the single-flanker condition was much larger for peripherally placed flankers.

Banks et al. (1979) used Gestalt grouping to explain the interference between letters, whereas Wolford and Chambers (1983) later used the same experimental paradigm to demonstrate that contour interactions (i.e., low-level visual masking) can also produce interference between a target and nearby flankers. Wolford and Chambers replicated Banks et al. (1979, Experiment 3, described above), except, they used two different target-flanker distances (this is the distance between the target and the nearest flanker letter): Banks et al.’s target-flanker distance was fixed at 1° visual angle (center to center), Wolford and Chambers used distances of .96° and .30°. Only peripheral masks were used. Results for the .96° were similar to those of Banks et al.: identification accuracy was better with five masks than with one mask. This function, however, was reversed for the close target-flanker configuration: identification accuracy was better with one mask than with five masks. In the close target-flanker condition, the target was now impinged upon by lateral masking associated with the flankers, which was compounded in the multiple flanker condition. This low-level interference could not be overcome by grouping among the flankers.

In their final experiment Wolford and Chambers were able to eliminate the effect of grouping among the flankers, even at the .96° target-flanker separation. These authors argued that the grouping benefit that Banks et al. observed was really the result of the additional information that the column of masks provided regarding the location of the
target. Wolford and Chambers tested their "cue-validity hypothesis" as follows: As in their previous experiment (described above), the subject's task was to identify a target—A, Y, U, or T—with the letter H always serving as the flanker letter. On some trials multiple flankers were stacked next to a single target (this was the arrangement used in the earlier experiment). However, to remove information provided by the letter arrangement, in one experimental condition, trials were included in which multiple targets (all the same letter) were stacked next to a single flanker mask:

<table>
<thead>
<tr>
<th>Stacked Flanker Arrangement</th>
<th>Stacked Target Arrangement</th>
</tr>
</thead>
<tbody>
<tr>
<td>F</td>
<td>T</td>
</tr>
<tr>
<td>F</td>
<td>T</td>
</tr>
<tr>
<td>F T</td>
<td>T F +</td>
</tr>
<tr>
<td>F</td>
<td>T</td>
</tr>
<tr>
<td>F</td>
<td>T</td>
</tr>
</tbody>
</table>

(T = target location, F = flanker/mask location, ‘+’ = fixation)

Under these revised grouping contingencies, subjects could no longer simply look to the ungrouped single letter to find the target. As predicted, when target location information was no longer provided by the stimulus configuration, grouping among distractor letters no longer facilitated target identification.

Wolford and Chambers concluded that at small target-flanker separations, feature-level interactions create interference between adjacent letters. At wider separations, however, interference is determined by the distribution of attention. Note, however, that Leat, Li & Epp (1999), observed both attentional effects, and contour interactions, at relatively short target-flanker distances. These findings remind us that as both low-level
and high-level visual processes contribute to perception, so may they separately succumb
to interference under the right circumstances.

Heretofore, most studies that I have described have involved a small number of
stimuli within a display. My aim in this literature review, however, is to examine the
perception of an object within both sparse and cluttered scenes. As more and more
elements are added to a visual display, the role of attentional selectivity becomes much
more important, nay, necessary. I will therefore take a short detour and discuss theories
of visual attention in general, after which, I will review attentional theories which address
issues of perceptual grouping in visual search.

Attention and Visual Search

Psychophysics vs. Visual Search: The experiments reviewed thus far can virtually
all be described as psychophysical studies. When stimuli were presented, subjects knew
where the target was located; their difficulty in identifying it arose either from the various
impediments placed nearer or farther from the target, or from some manipulation of the
target itself (eccentricity, size, etc.). In psychophysical studies, subject performance is
assessed over a range of stimuli, producing some measure of threshold for a particular
task (i.e., level at which subjects perform at 75% accuracy).

If a target's location is not known, and the target must instead be found in a
display, this becomes a visual search task. Duncan (1985) distinguishes between the
target's defining attribute, which tells the subject which element is the target, and the
target's reported attribute, which identifies the subject's response, and which may be
totally different from the defining attribute. Thus, in a task to identify a red letter, "red" is the defining feature, and the letter identity is the reported attribute. In visual search, the target may be clearly visible; the question is, How quickly can it be found when placed in a field of distractor elements? As such, response time is usually recorded in visual search experiments. (Note, however, that an experiment using a visual search task can also incorporate a psychophysical manipulation.)

Over the last three decades, the visual search paradigm has proven to be a valuable tool in the study of visual attention. The reason for this is that visual search provides a measure of the selectivity of visual processing—selectivity being perhaps the essence of what we call attention. If we assume that under some circumstances not all stimuli in the visual field can be fully processed at once, the ease or difficulty with which a particular target can be found provides a gauge of the processing priorities of the visual system. Thus, if the response to a target is fast, we can assume that, for whatever reason, the target stimulus was operated upon earlier in the visual processing-response stream than a target to which a response is made more slowly. By manipulating various aspects of the scene being searched, attention researchers attempt to come up with that "whatever reason", to finally arrive at a theory of both visual processing and attention. The next section introduces some of the more influential theories of visual attention.

Theories of Visual Attention

Since the time of Neisser (1967), many attention theorists (e.g., Marr, 1980; Julesz & Bergen, 1983; Treisman & Gelade, 1980; Wolfe & Cave, 1989); have concurred
that visual perception occurs in a series of stages. Neisser had pointed out that a detailed analysis of the entire visual scene is beyond the capacity of the visual system; therefore, to get around this limitation, some preliminary segregation is required to break the scene up into smaller units. The units produced by this early "preattentive" stage serve as candidates for further analysis by a later "attentive" stage of processing.

Preattentive processing is carried out automatically and in parallel, i.e., simultaneously, across the visual field. For processing to be automatic, this implies that the visual system is tuned to process certain elementary components or primitives. Julesz (1981; Julesz & Bergen, 1983) labeled these components "textons" and claimed that the visual system's differential sensitivity to these elements led to the instantaneous perception of various textures across differing texton densities. Based on his studies of texture discrimination Julesz included "elongated blobs" (e.g., line segments), line endings and crossings of lines as likely textons. (Note that not all theorists agree as to the particular features that make-up these basic perceptual units.)

The general view of preattentive processing described above is consistent (at least metaphorically) with low-level vision physiology: a number of different basic visual properties (e.g., color, orientation) are known to be registered in parallel early in visual processing (i.e., area V1) and analyzed within distinct subsystems (Hubel & Wiesel, 1968).

The visual primitives of the preattentive stage are next operated upon during the attentive stage of processing. The function of attentive processing is to combine these independent elementary components in their place, to create complete objects from vague
figures. Unlike the preattentive stage, which operates over the entire visual field, the attentive stage of processing has a limited capacity and is restricted to only a portion of the visual scene at any one time. As a result, the "spotlight" of attention (Posner, Snyder & Davidson, 1980) is allocated serially, object by object (or region by region). The product of this synthesis (Neisser, 1967) by attentive operations yields a fully realized percept of a three-dimensional object. The representation thus constructed may then be matched to stored representations (that is, memory) resulting in some level of recognition of the viewed object.

A major problem for attentional theories has been to describe how the process of "binding" together separate properties to form objects is performed in the visual system. Feature Integration Theory (FIT) as proposed by Treisman (Treisman & Gelade, 1980; Treisman & Gormican, 1992) was an early, and as it has turned out, enduring explanation of the binding problem in particular, and attentional allocation in general.

An empirical basis for FIT came from comparing the results for two types of visual search studies: feature searches and conjunction searches. In a feature search, the target is defined by a unique feature value (e.g., a red target letter in a field of green letters). In a conjunction search the target shares at least one feature with all distractors, but is defined by a unique combination of features (e.g., a pink O target in a field of green O's and pink N's). Treisman (Treisman, Sykes & Gelade, 1977; Treisman & Gelade, 1980) found that feature searches (such as the one just described) tended to yield flat search slopes: the mean response time (RT) to detect the presence or absence of such a target was independent of the number of distractor letters in the display. These targets
were often said to “pop-out” from the display. The results suggested that feature
differences are detected via a parallel search of display stimuli. However, in the
conjunction search task, mean RTs tended to increase monotonically with distractor
number (Treisman & Gelade, 1980), with the ratio of target absent slope to target present
slope being approximately 2:1. Treisman interpreted these results as indicating that the
items were inspected serially to perform the conjunction task. (Specifically, the 2:1 ratio
suggests a serial self-terminating search in which all distractors are inspected in the
target-absent condition, and one half, on average, are inspected in the target-present
condition (Sternberg, 1969)).

Feature Integration Theory starts with the premise that basic stimulus properties
(color, orientation, etc.) are initially registered in parallel in retinotopically organized
“feature maps”. A different feature map exists for each dimension of a feature: for
instance, there are separate maps for the colors red, green, and blue, as there are for the
orientations of horizontal and vertical (there is some debate as to what are the identities of
all the feature maps). The presence of a stimulus with a particular feature value produces
activation in the corresponding feature map. However, two important restrictions are that
1) activation within one feature map is separate from that of all other feature maps, and 2)
the location of activation within a map is not accessible by the observer (at least not
preattentively). Object properties, at this stage are described as being free-floating.

Here is where attention comes in: “focal attention provides the glue which
integrates the initially separable features into unitary objects” (Treisman & Gelade,
1980, p. 98). According to FIT, when called upon, attentional mechanisms act on a
“master map” of locations. The master map has access to the myriad feature maps and
their associated activations for each display location. As attention is focused on a
location on the master map, the different properties across the feature maps are bound
together into a single entity. In a conjunction search, for example, individual locations on
the master map are inspected serially until, for example, a target is found, terminating the
search—hence, this is referred to as a “serial, self-terminating search”.

Treisman and her colleagues’ (Treisman, Sykes, & Gelade, 1977; Treisman &
Gelade, 1980) findings in regard to feature and conjunction search are predicted by FIT:
when searching for a red X among green X’s, activation within the red feature map alone
is sufficient evidence by which to make a present response, and this task will be
unaffected by the number of green distractor X’s, yielding a flat search slope. However,
in a conjunction search (red X among red O’s and green X’s), evidence for a target’s
presence (green + X) cannot be found in any one feature map. FIT predicts that only by
serial inspection of the stimuli, and the conjoining of each item’s respective features
could the target be found: the steep slopes found in studies of conjunction searches
support this prediction. (See also Treisman & Schmidt, 1982, and Treisman & Souther,
1985, for additional studies in support of FIT.)

Since its advent two decades ago, FIT has provided researchers with a highly
effective framework by which to interpret many of the findings in attention research.
However, as with any theory, over time some of its premises have come under fire.
Among the empirical findings unexplained by FIT are the following:
In a visual search study by Nakayama and Silverman (1986), targets were defined by a conjunction of color and depth. Search slopes for this task were flat (i.e., independent of distractor number), suggesting that subjects were able to perform this conjunction search in parallel. Similarly, Wolfe, Cave, and Franzel (1989) found that search for a target defined by a triple conjunction of features (size X color X orientation) could also be performed very efficiently. Both these findings are in direct opposition to the serial self-terminating search predicted by FIT. In each case, the salience of the target facilitated search, obviating the need for serial inspection of the stimuli.

Finally, Egeth, Virzi, and Garbart (1984) showed that when performing a conjunction search, inspection could be restricted to a subset of display elements sharing one particular feature with the target. In this study, the target was a red O and distractors were black O’s and red N’s. When subjects were instructed to attend to the red stimuli, search times were independent of the number of black distractors. Apparently, the viewer’s intentions play a role in determining which stimuli are inspected. This finding is inconsistent with the notion of random selection of master map locations. Subjects can selectively search just those locations containing, say, red objects.

In contrast with FIT are theories that can account for the findings just recounted by claiming that attention can be guided by low-level visual features (Hoffman, 1978, 1979) and stimulus salience (see the revised FIT, Treisman & Sato, 1992). Wolfe’s Guided Search Model (Wolfe, Cave, & Franzel, 1989; and Guided Search 2.0, Wolfe, 1994) incorporates both these attributes, while also sharing many features with FIT.
The centerpiece of Guided Search is the activation map: this retinotopic map (analogous to the master map in FIT) represents the likelihood of a target being at a particular location, based on the sum of bottom-up and top-down components. (In the language of perception, "bottom-up" refers to the physical characteristics of the stimulus itself. "Top-down" refers to factors resulting from the observers state, task or intention.) Attention is directed to the activation map location having the highest activation, and then to other locations in decreasing order of activation.

The bottom-up component of activation represents an item's salience, that is, the difference between the item and all others in the display. Salience is measured separately within feature types, so that a single horizontal red line amidst red vertical lines would have low color salience, but high orientation salience and hence, high activation. (Note that low-level maps in Guided Search are by feature type (color, orientation, e.g.) as opposed to feature dimension (red, horizontal, e.g.) as is the case for FIT). In addition, the salience of an item in a scene is inversely related to its separation from other items, with proximal items contributing to an item's salience-based activation more than those farther away (Nothdurft, 1993). These salience measures are then summed across the feature types to produce an item's bottom-up activation.

The top-down component of activation is determined by the number and similarity of features shared between a stimulus element and the target. The greater degree to which an item matches a target, the greater its top-down activation. (Conversely, in the revised FIT, a similar trick is accomplished by inhibiting activation at items dissimilar to the target (Treisman & Sato, 1992)).
It is through the influence of this top-down component that properties of the target can be said to guide search, permitting, for example, only items of a particular color to be searched (Egeth, Virzi and Garbart, 1984). Likewise, bottom-up activation of a sufficiently salient stimulus can make a target defined by a triple conjunction of features pop-out from a display (Wolfe, Cave & Franzel, 1989).

Perceptual Organization and Attention

I have discussed both Feature Integration Theory and Guided Search at length because together they provide a good foundation for understanding theories of visual attention and processing. However, as resilient as these theories have been over time, they do share one shortcoming: Neither theory addresses the role of grouping and perceptual organization in visual processing (though, see Treisman & Sato, 1992 for the revised FIT). In the following section I review studies that have explored this issue, and describe the resulting theories which incorporate these findings.

Banks and colleagues (Banks, Bodinger, & Illige, 1974; Banks & Prinzmetal, 1976) employed the visual search paradigm to examine the effects of grouping among stimuli on target identification. In these studies, a target letter (a T or an F) was presented in an array with a number of letter-like distractors, and/or noise patches (depending on the study; see Figure 1). The task in each study was to determine which possible target, T or F, was present in the array. The primary manipulation in the studies was the degree to which the target letter was integrated among a set of distractor elements; the target could be in the middle of a set of distractors, it could totally be set apart from them, or its
placement could be at some level of grouping in between these. In both studies it was found that the target became harder to identify (as reflected in elevated response times) as the degree of grouping between target and distractors increased.

Banks and Prinzmetal (1976) proposed a two-stage model to explain their findings: in the first stage, a perceptual parser organizes the visual field, and in the second stage an analyzer operates on groups output by the first stage. These ideas were later adapted and expanded on by both Kahneman and Henik (1981) and Duncan and Humphreys (1989), both of whose theories are described below. Note that these theories share the idea that attention operates on highly processed (i.e., grouped) units, rather than simple features or primitives (Julesz & Bergen, 1983; Treisman & Gelade, 1980).

Kahneman and Henik (1981) investigated grouping and perceptual organization in attentional tasks (including visual search) and based on their findings developed a model of attentional selection. In one experiment, twelve letters were presented in two rows of six for 200 ms. Six letters were red and six were blue, and they were shown in various groupings of red and blue. The participants’ task was to identify all the letters of a specified target color. Performance was best (i.e., the most letters of a target color were identified) when letters of the target color were blocked in the display; i.e., as a single row of six, or two sets of three. Performance was worst when different color letters were alternated in the display. These results suggest 1) that items are selected as groups, and 2) that coding is inversely related to the number of groups: selection is facilitated when there are fewer perceptual units.
Kahneman and Henik’s model of attentional selection emphasizes the role of the perceptual object. They state, “resources (or attention) are allocated to groups (or perceptual units) that are formed by an early preattentive process. We assumed that these resources are limited and that they are allocated hierarchically—first to groups, then to individual items within a group” (p. 197). As with the other attentional theorists discussed thus far, Kahneman and Henik claimed that preattentive processing operates in parallel, and that its output can serve to orient and direct attentional processes. Finally, preattentive perceptual organization would likely follow the Gestalt principles of grouping—similarity, proximity, continuation, etc.

What factors related to grouping make a particular visual search more or less efficient? Let’s look at an experiment performed by Bundesen and Pedersen (1983). In this visual search study, subjects identified the alphanumeric class of a target of a known color among a number of letters and numbers. Nontargets could be any one of 2, 4 or 6 different colors, and there could be 8, 16 or 24 nontargets in a display. In one condition, all nontargets were presented in clusters based on color. In this grouped-color condition, adding a single nontarget had a much smaller effect on response time (about 2 ms/item) than adding an additional nontarget color group (about 12 ms/group; note that this increase is based on a comparison of two displays having the same number of nontargets, but a different number of colors). Based on these findings, the authors state that search comprised sequential evaluation of groups of same-color nontargets until the target was found.
Citing Bundesen and Pedersen (1983), as well as their own work, Duncan and Humphreys (1989) proposed a model of grouping in visual search that emphasized the role of similarity between targets and nontargets (T-N), and also the similarity between nontargets (N-N) themselves. They claimed that visual search became more efficient as 1) T-N similarity decreased, and 2) N-N similarity increased (see also Banks & Prinzmetal, 1976, who posed a similar hypothesis).

According to this model, processing in visual search begins with a parallel stage of perceptual grouping the authors label description. The scene is coded and segmented (according to Gestalt principles, etc.) producing a set of hierarchical structural units, each defined by a set of properties (color, size, shape, etc.). The representation of these units, though it is highly processed, is outside awareness. As such, the structural units compete for access to visual short-term memory (VSTM), where they can serve as the focus for action (e.g., response to a specified target). The better that nontargets can compete with targets for access to VSTM the less efficient will be the search.

It is at the second stage, selection, in which T-N and N-N relations become paramount: access to VSTM is determined by allocation of a limited capacity resource which is assigned to the structural units created in the first stage. The greater the allocation of this resource to any unit, the higher the probability that the unit will gain access to VSTM. Each unit competes for the resource in proportion to a weight assigned to it. In the case of visual search, units are matched to a target template, which is a description of the object of the search (i.e., based on its defining feature): the greater the degree to which a unit matches the template, the greater its assigned weight. Hence, a
nontarget similar to a target attracts more of this resource than a less similar nontarget, and is more likely to gain access to VSTM in lieu of the actual target. (Note that this function resembles the top-down activation component in Wolfe’s Guided Search Model.)

Another major claim of this model is that the weights assigned to structural units covary with the level of grouping between the units. That is, the weighting of one unit is linked to that of other units in proportion to the degree to which the units are grouped together. Weighting is thus uniformly reduced across a group of similar units (e.g., nontarget distractors) through what Duncan and Humphreys call spreading suppression. Spreading suppression will facilitate visual search when N-N similarity is high: nontargets are mutually assigned low competitive weights and are then rejected from search en masse. Conversely, when N-N similarity is low, there will be less opportunity for spreading suppression among them, and search will be less efficient. (See also Humphreys and Muller (1993), whose connectionist model incorporated many of Duncan and Humphreys’ (1989) assumptions regarding the role of T-N and N-N relations in grouping and visual search.)

Both these models can nicely account for the findings of Bundesen and Pedersen (1983). There was little effect on RT when nontargets were added to an existing color-group because the added elements were subsumed within the wave of spreading suppression due to their similarity to other nearby nontargets (N-N similarity). However, a much larger detriment to performance occurred when a new nontarget color was added because this additional group could not benefit from suppression spreading across any
existing color-group (due to low N-N similarity); the new color-group had to be processed and suppressed separately. ²

As noted earlier in this section, beginning with Neisser (1967), a number of theories of visual processing and visual search have described a preattentive processing stage in which simple feature values (color, orientation, etc.) are encoded in parallel across a visual scene (Julesz, 1983; Treisman and Gelade, 1980; Wolfe, 1994). These retinotopic feature maps (i.e., Treisman & Gelade, 1980) form the basis for attentional selection and are assumed to be the product of a minimal amount of processing; when the observer needs more information than is held within these simple representations, attention must be directed to local objects in a serial fashion.

There is growing evidence, however, that attention is guided by perceptual units that represent a high level of information from the visual field: For example, Enns and Rensink (1990), showed that attention can be directed by 3-D surface and shading information. In that study, displays consisted of many items, each formed by an arrangement of three quadrilaterals. In one condition, each set of quadrilaterals always formed a cube; in the other condition the three quadrilaterals formed a flat figure (see Figure 2). Notice that when the stimuli are cubes, if one cube is oriented differently from the other cubes, it appears to pop out from the display (and search is highly efficient). In the flat-figure condition, a differently oriented arrangement does not pop out. In the cube condition, the shading of the pop-out cube is perceived as being inconsistent with the

² Details of this description would not be the same for Humphreys and Muller (1993), though the thrust of the argument would be the same.
apparent lighting of the scene (as reflected by the other cubes); whereas, in the flat-figure condition, this 3-D shading cue is lacking. (See also He and Nakayama (1992), and Ramachandran (1988) for similar findings of attentional guidance from 3-D surfaces and shading.)

In Grossberg, Mingolla, and Ross' (1994) neural theory of visual search these highly processed representations play a key role in the allocation of attention. According to this theory, the output of an early parallel processing stage serves as input to grouping mechanisms, whose resulting representations serve as input to attentional mechanisms. As with other such theories already described, in the initial processing step, simple features are registered retinotopically. In the second step of processing, neighboring objects sharing various features (e.g., depth and color) are grouped together; the contours between these grouped regions are delineated and eventually bounded surfaces emerge. These surfaces ultimately aid perception by separating figure from ground in the visual scene. In a visual search task, these regions are inspected by object recognition mechanisms, based on salience and top down priming from stored target representations. This segmentation and inspection process repeats iteratively within and across groupings depending on how closely group members match a stored target representation.

The theories of Duncan and Humphreys (1989) and Grossberg, Mingolla, and Ross (1994) make it clear that the line between preattentive and attentive processes is much fuzzier than originally conceived by Neisser. Recall that as instituted by Treisman in FIT and Julesz’s Texton Theory, the primitives of vision were described as limited in both number and form. Over time, however, the lexicon of so-called preattentive features
has grown to include highly processed representations (i.e., 3-D forms, Enns & Rensink, 1990) with no simple analog in early vision. As such, Joseph and Nakayama (1998) propose that the preattentive/attentive dichotomy is no longer useful and should be rejected. Rather, they propose that all conscious vision requires attention, and that easy and difficult tasks differ in the spatial scale and resolution of attention required to perform them. Thus, Joseph and Nakayama claim that even the simplest of feature detection tasks (i.e., a “pop-out”) is really an act of global pattern recognition accomplished with a wide distribution of attention. The most difficult of conjunction searches, on the other hand requires a much finer allocation of attention.

Results of a number of recent studies provide evidence that indeed, attention may be needed even for visual tasks long regarded as being preattentive and performed in parallel. Joseph, Chun and Nakayama (1997) found that an otherwise easy orientation detection task was severely impaired when subjects had to concurrently perform a separate attentionally demanding task (but see Braun, 1998). And in a study by Mack and Rock (1998) subjects often claimed to not even see a stimulus shown at fixation when they were attending to an object in the periphery. Mack and Rock have documented many similar cases in which subjects could not detect a suprathreshold stimulus while engaged in a second task, naming this unusual phenomenon “inattentional blindness”. The findings of both Joseph, Chun and Nakayama (1997), and Mack and Rock (1998), strongly suggest that conscious vision is not possible without attention (but see Moore & Egeth, 1997).
The four-decade-old preattentive/attentive distinction is clearly on shaky ground. Joseph and Nakayama claim that not only are we constrained by the use of terms such as "preattentive", "parallel" and "pop-out", but that we are in fact misled by them, in that their underlying premise is faulty. On the plus side, theorists such as Neisser and Treisman have provided perceptual psychologists with a language with which to speak about attention and a basic framework which has proven to be adaptable as new findings are uncovered. Old habits die hard, however. What may be needed is for a radical theory such as that of Joseph and Nakayama to be more fully fleshed out, with an accompanying terminology of its own. Until such time, we'll likely continue to read about preattentive and attentive processes.

The aim of this review is to establish the theoretical and empirical framework by which to understand how spatial separation affects attentional tasks. With that in mind, this first chapter took what might be described as the outside-in view of this issue, examining how the presence of surrounding objects affect the processing of a particular object. I next examine the inside-out view, and ask, How does attending to a particular object affect the processing of surrounding elements?"
Figure 1. Two stimulus displays from study by Banks and Prinzmetal (1976). The task was to detect whether a "T" or and "F" is present. At left, target, T, is not a part of distractor group. At right, target, T, is grouped with distractors.

Figure 2. Stimuli based on Enns and Rensink (1990). The task in this visual search experiment was to detect the presence of the odd figure in the display. Three-dimensional stimuli at left, and "flat" stimuli at right.
Chapter 2: Processing Around an Attended Object

Intuition tells us that if one is attending to a particular item (i.e., an object or a spatial location), targets near the item are better perceived than those farther from it. Indeed, researchers have reported many findings to this effect (for review see Yantis, 1998). For example, Hoffman and Nelson (1981) demonstrated that successfully performing a visual search depends on attention being allocated to a target’s spatial position. In this study, displays consisted of four letters, one at each cardinal compass point. Subjects performed a dual task: they searched for a target letter and had to make an orientation judgment about a U-shaped figure that appeared next to one of the letters. In different blocks of trials, subjects were instructed to allocate various proportions of attention to each task. Hoffman and Nelson found that focusing on one task produced a decline in performing the other; however, the trade-off in performance was reduced when the target letter and U-shape were in adjacent positions. The authors surmised that when attention is directed to a location there is a benefit in processing information nearby.

Using a flanker paradigm, Eriksen and colleagues (Eriksen & Hoffman, 1972, 1973; Eriksen & Yeh, 1985; Eriksen & St. James, 1986) have also shown facilitated processing near an attended location. Recall that flanker effects refer to the high-level interference (i.e., at the decision or response level) caused by items near a target letter. In one study (Eriksen & Hoffman, 1973) there were two sets of targets (A and U versus H and M) each set requiring a different—that is, competing—response. Displays consisted of twelve letters arranged around the edge of an imaginary circle and the target was
indicated by a bar perpendicular to the circle, adjacent to where the target letter would appear. The distance between the target and the incompatible elements was varied, as was the time by which the cue preceded the display (SOA ranged from 0 to 350 ms).

The major findings of this study were 1) the nearer an incompatible flanking item was to the target item, the more response times were slowed, and 2) flanker interference diminished with increasing cue-display SOA. This latter finding suggests that the attentional aperture, so to speak, can be narrowed over time. However, even at the maximum SOA, the effect of an adjacent flanking letter was not totally eliminated, indicating that the aperture did not shrink to less than 1° of visual angle, a figure the authors labeled as the minimum resolution of attentional focus. Eriksen and St. James (1986) hypothesized that attention operates like a zoom lens, capable of being widely distributed across the visual field or narrowly focused on a small region, with resolution improving as the size of the attended region decreases.

The studies of Eriksen and colleagues describe scenarios in which attentional selectivity fails: flanker compatibility effects are only manifest when an object is inadvertently perceived. In a study in which subjects willfully attended to two items, a processing advantage for regions near an attended location was also reported by LaBerge and Brown (1986). Using a dual task paradigm, two target displays were shown in succession, with the second superimposed on the first. First, five letters were shown and subjects had to detect either a target word or a centrally located letter (these conditions were blocked). The second display consisted of a character and two flanking elements: subjects only responded if the character was an R, and had been preceded by a particular
target (letter or word, depending on the block). When a letter served as the first target, RTs were speediest for targets nearest the central (preceding) target, and grew slower with distance from the center. When a word served as the first target and attention was widely distributed, response times did not vary with eccentricity, and were uniformly faster than in the letter-target condition. This latter finding affirms that the processing advantage for shorter target separations was not merely the result of increased eccentricity and reduced sensitivity.

If there are processing benefits in the vicinity of the attentional locus, what is the structure of this surrounding attentional field? An answer to this question was provided by Downing and Pinker (1985) who conducted a parametric investigation into the effect of distance from an attended location on target detection. Displays comprised 10 small boxes aligned horizontally, and numbered above from 1 to 10. On each trial, a number appeared at fixation indicating the likely box location (70% valid) of a subsequent target (a spot of light). Subjects were instructed to attend to the location of the cued box, while maintaining central fixation. In this speeded task, subjects pushed a button when the spot of light appeared. In agreement with the studies cited above, response times were fastest to targets at the cued location, and increased with target distance from the cue (though, this effect was somewhat attenuated for targets in a different hemifield from the cued location).

Downing and Pinker (1985), and LaBerge and Brown (1986, 1989) explained their findings by positing a gradient of attention (see also Cheal, Lyon & Gottlob, 1994; and Henderson & MacQuistan, 1993). According to this view, at the focus of attention
lies the apex of processing resources—objects here receive the greatest processing benefit—with available resources diminishing with distance from the focus. By comparison, recall that Eriksen and colleagues presented a competing view, describing attention in terms of a “zoom lens”. Similarly, attention has also been compared to a “moving spotlight” (Posner, Snyder & Davidson, 1980; Shulman, Remington and McLean, 1979; Tsal, 1983). Advocates of the moving spotlight metaphor claim that attention sweeps from location to location at (approximately) a constant velocity. (According to different versions, the spotlight may either traverse all intervening locations, or be shut off between them.) The moving spotlight model thereby predicts that the time to make a shift of attention will be proportional to the distance between attended items. Note that various attention-like operations such as mental rotation (Shepard & Metzler, 1971) and mental curve tracing (Pringle & Egeth, 1988) have also shown this analog characteristic of distance-dependent performance (see also Finke & Pinker, 1982).

The results of a study by Tsal (1983) support the moving spotlight (or analog) model of attentional shifting (see also Shulman, Remington, & McLean, 1979). The displays used by Tsal had six possible target locations, three on either side of fixation along the horizontal meridian. The eccentricities of these locations were 4°, 8°, and 12° visual angle from fixation. Subjects began each trial attending to a central fixation crosshair. On each trial, a cue appeared at one of the possible target locations, followed by a target (O or X) which always appeared at the cued location (100% valid cue). Cue-target onset asynchrony (CTOA) varied from 50 to 183 ms. For each target eccentricity,
response times to identify the target were longest for short CTOAs, and decreased to an asymptote with increased CTOAs. The key finding was that the asymptote occurred earlier (i.e., for shorter CTOAs) for the less eccentric locations than for more distant locations. Tsal interpreted the time to asymptote to be a measure of the time needed for attention to arrive at the particular location, meaning that attention took longer to travel a longer distance than a shorter one; the hallmark of an analog shift of attention.

However, Yantis (1988) offers an alternative, perceptual explanation: Assume that movements of attention are abrupt, but that their speed varies stochastically. Also, assume that responding is fastest when attention is allocated to the target location. Yantis hypothesizes that the time required to make a shift of attention may be a function of the perceptibility of the attended location. Thus, the more peripheral target locations may reach asymptote after the less eccentric ones because peripheral locations are not perceived as well, reducing the likelihood of an early strike by an abrupt attentional shift. Probabilistically, there would be fewer (relatively) fast shifts made to the more peripheral locations than the less peripheral locations, producing the pattern of results found by Tsal (1983). (Also see Yantis, 1988, for a critique of Shulman, Remington & McLean, 1979.)

In a recent study, the spotlight and gradient hypotheses were directly pitted against each other (LaBerge, Carlson, Williams, & Bunney, 1997). In this "triadic"-task subjects had to identify three characters, one each from three different displays superimposed upon each other in sequence. The first target was always centrally located, whereas the locations of the second and third targets varied; all locations were within 1.7° visual angle. If the spotlight hypotheses held true, then the time to respond to the third character
should have increased with distance from the second target, the presumptive previous focus of attention. However, response times were in fact a function of distance from the first, and not the second target. LaBerge et al. attribute this finding to the presence of an area of preparatory attentional activity (i.e., a gradient) distributed near the first target; even after the second target was identified, a large residue of activation remained at the originally fixated and identified location, creating a processing advantage for nearby locations.

Regardless of which metaphor holds forth—spotlight, zoom lens or gradient—are we to conclude that performance in an attentional task always varies inversely with distance from a selected location? Unfortunately, the issue is not so simple. Over the past two decades, evidence has also been accumulating to support the hypothesis that there is no effect of separation on shifts of attention (Eriksen & Webb, 1989; Kröse & Julesz, 1989; Remington & Pierce, 1984). Sagi and Julesz (1985) used a difficult discrimination task to demonstrate that attention moves in an abrupt, discrete manner, independent of distance. In this study, only two letters were presented on each trial. Each letter appeared along the circumference of an imaginary circle (both letters at 4°, or both letters at 8°, eccentricity), in one of 18 possible locations. Letter displays were visible for 10 ms and then masked. An accuracy measure was used: the two observers assessed whether the two letters were the same or different, as each letter could be either a rotated T or L. Results of the study showed that performance was virtually unaffected by the distance between the letters, leading the authors to conclude that shifts of attention were what they called “noninertial.”
A study corroborating the findings of Sagi and Julesz (1985) was performed by Kwak, Dagenbach, and Egeth (1991). These authors measured response time rather than accuracy, and in different experiments used upright T’s and L’s (Experiment 1) and rotated T’s and L’s (Experiment 2) as stimuli. In each case, Kwak et al. found that response time was unaffected by distance between target objects. A diagnostic procedure (Experiment 3) was also performed as part of this study to examine whether the letters were identified serially or in parallel. The results of this procedure supported the claim that observers processed the rotated letters serially. This finding strengthened the assertion that observers really were shifting attention in a discrete manner, rather than processing letters in parallel across the visual field.

Studies by Sperling and colleagues using rapid serial visual presentation (RSVP) have also demonstrated the discrete character of attentional shifts, leading to the development of the Episodic Theory of Attentional Facilitation. In one experiment (Sperling & Weichselgartner, 1995) displays consisted of two separate streams of characters. At each stream location a character was presented and quickly replaced by another character, and then this one was replaced by another letter, and then this one was replaced by another letter, etc. (hence, rapid serial visual presentation). There was one stream of letters and one of numerals. Subjects fixated the numeral stream while monitoring the letter stream for a target character, C. The subjects' task was to switch attention to the digit stream when the target letter appeared, and report the first digits that they saw. The latency between the target letter's appearance and the first reported digit served as a measure of the speed of the attentional shift. In Experiment 2A there were
two manipulations: the distance between the streams could be 2°, 4°, or 7° visual angle. In addition, a visual obstacle, in the form of an RSVP stream of noise elements (#, ?, +, etc.) was sometimes inserted between the streams when the separation was 7°. Using an analytical method that took differences in peripheral resolution into consideration, Sperling and Weichselgartner found 1) no effect of separation between target streams, and 2) no effect of the obstacle (noise stream) on the time to shift attention.

Based on their own and other similar findings, Sperling and Weichselgartner (1995) presented the Episodic Theory of Attentional Facilitation. According to this theory, a shift of visual attention occurs as a quantized jump, whose time is independent of distance. As with a saccadic eye movement, locations between attended objects are not attended. Rather than a moving spotlight, attention is compared to a fixed spotlight, or, actually, a set of fixed spotlights. As the authors describe it, the visual field is like a theater stage that may be illuminated by any one of numerous possible spotlights. A spotlight can be aimed at any location, but may not be moved once so directed. These spotlights can assume any size or shape. When attention shifts, a spotlight switches off at the presently attended location and another spotlight switches on at the to-be-attended location. The switching process is not instantaneous though, taking tenths of seconds for attention to fully power down and power up at the respective locations.

Note that the Episodic Theory of Attentional Facilitation describes the dynamics of attentional movement. As such, it is able to account for the studies discussed in this section (Kwak, Dagenbach, & Egeth, 1991; Sagi & Julesz, 1985; and Sperling & Weichselgartner, 1995) all of which found no effect of distance between serially attended
locations. Conversely, this theory contradicts the moving spotlight metaphor and earlier findings that attention shifts in an analog fashion. The episodic theory, however, may not be inconsistent with a gradient theory of attention. One simply needs to assume that the illumination cast by a fixed spotlight is not of uniform intensity, but is strongest at its center, and diminishes with distance to the light's edge. This conception allows for attentional gradients to arise within circumscribed regions, and quantum attentional shifts to occur across the larger visual field. Perhaps gradients are local, and quantum movements are global.

Thus far, evidence has been presented suggesting that a processing advantage is conferred upon locations nearer attended locations, and that shifts of attention are unaffected by the distance between target elements. To further muddle the situation, there are many recent studies showing performance to be positively correlated with distance—that is, as separation between attended items increased, performance improved (Bahcall & Kowler, 1999; Caputo & Guerra, 1998; Cave & Zimmerman, 1997; Mounts, 2000a; Mounts, 2000b; Turatto & Gandalf, 1999). Cave and Zimmerman (1997, Experiment 2) found this pattern of results using a post-stimulus probe paradigm. A circular array of eight letters was presented briefly (75 ms). The array was removed, and then followed by a probe (a small red square) which appeared at the location of one of the letters. There were two tasks: one task was to detect the presence of a target letter (F for some subjects, Y for others) in the letter array. The critical task was to make a keypress when the probe appeared (which was on 50% of trials). Response times to detect the probe were fastest when it appeared at the target letter location. However, for probes
occurring at non-target locations, RTs generally decreased (i.e., performance improved) with increased distance from the target letter. (This effect was strongest after thousands of trials of practice.) Cave and Zimmerman attribute this effect to a form of flanking inhibition for distractors surrounding the attended target.

Cepeda, Cave, Bichot, and Kim (1998) adapted the post-stimulus probe paradigm to more fully flesh out the nature of flanking inhibition. Each display consisted of four digits presented at the corners of a slightly tilted imaginary diamond shape. After removal of the initial display, a subsequent probe could appear at either a formerly occupied or unoccupied location. The subjects’ task was to identify a uniquely colored digit in the initial display, and respond to a probe, when it appeared. Cepeda et al. found that response times to probes at distractor locations were inhibited relative to probes at the target location. Unexpectedly, there was no inhibition for probes at empty locations, even for those locations in the same quadrant as the target location.

As these results prescribe, Cave and colleagues (Cave, 1998; Cave, Kim, Bichot & Sobel, 1998; Cepeda, Cave, Bichot, & Kim, 1998) proposed a mechanism that provides for feature-driven inhibition of distractor locations. According to their FeatureGate model of attentional selection, objects not sharing (known) features with a target are closed (that is, “gated”) from higher-level processing; vacant, intervening space, however, is not so inhibited. Distance effects (such as found by Cave & Zimmerman, 1997) emerge from FeatureGate’s hierarchical structure: Objects compete for selection within separate neighborhoods of the visual field, with local “winners” passed to (later) intermediate levels of processing. In the region containing the target, there is a strong
competitive bias toward this item that quickly knocks nearby distractors out of contention for selection. More distant distractors don't compete with the target until they've reached more intermediate processing levels. Based on the empirical findings, Cave (1998) asserts that inhibition of distractors must be stronger at the lower than at the higher levels. Adding this assumption to the model yields the requisite distance effects.

An alternative to Cave and colleagues' conception of feature-driven inhibition of distractors is the view that a region of suppressed processing surrounds an attended object. Evidence for such an inhibitory surround comes from a number of studies that used an attentional capture paradigm (Caputo and Guerra, 1998; Mounts, 2000a; 2000b; and Turatto & Gandalfo, 1999; see Theeuwes, 1991; 1992 on Attentional Capture). In the cited studies, subjects searched for a target presented in the company of one or more highly salient distractors. In the Mounts (2000a) study, for example, a single uniquely colored item (the color singleton) was always present in briefly presented test displays of 72 items. The subject's task was to make a discrimination regarding a uniquely shaped item (a form singleton) in the display: an eight, as on a digital watch, which had either its top or bottom line removed (the remaining distractors were all complete figure eights). The color singleton was to be ignored (unless it was also the form singleton).

Discrimination performance, as measured by d', was best when the target (defined as the form singleton) was also the color singleton. Otherwise, performance was worst for targets nearest to the color singleton, and improved as the target-to-color-singleton distance increased. Thus, even though the target was defined by form, and color was totally irrelevant to the task, the data suggest that attention was captured by the color
singleton: processing was facilitated at the color singleton's location (the focus of attention), and processing was suppressed in the nearby surrounding region.

To explain attentional capture by an irrelevant color singleton Mounts (2000b) appealed to Bacon and Ergth's (1994) notion of Singleton Detection Mode in search: since the target was a form singleton, subjects may have been set to select a unique item—and a saliently colored element qualified as unique. Mounts tested this hypothesis in an experiment in which the target was defined by a conjunction of features, and hence, no features were unique to the target (Mounts, 2000b, Experiment 1). Under this condition, attentional capture by the color singleton was eliminated as was the accompanying inhibitory surround.

Mounts discusses what may be a pertinent difference between studies that have shown attention as a gradient versus those that have shown the inhibitory surround. Experiments that have shown attention as a gradient, (Downing & Pinker, 1985; LaBerge & Brown, 1986, 1989) included a considerable period during which subjects focused attention at a possible target location, followed by the appearance of target stimuli. In contrast, experiments that have shown an inhibitory surround (Caputo and Guerra, 1998; Cave & Zimmerman, 1997; Mounts, 2000a; 2000b; and Turatto & Gandalfi, 1999) did not allow for any preparation other than fixation at the display center; subjects simply responded to a briefly presented display. As Mounts point out, these differences are in accord with LaBerge et al.'s (1997) distinction between preparatory and selective attention: Preparatory attention manifests when subjects have time to prepare for a target, i.e., at an expected location. Under these conditions, processing is facilitated at the
attended location and diminishes with distance. Selective attention refers to the perceptual filtering of distracting sensory information. Mounts asserts that under conditions in which selective mechanisms dominate (i.e., there is no chance for a preparatory response) not only is the attended area facilitated, but nearby locations may also be inhibited.

The studies just cited provide evidence that visual processing is reduced near a single target element (Cave & Zimmerman, 1997; Cepeda, Cave, Bichot, and Kim 1998), or near an attention capturing element (Caputo and Guerra, 1998; Mounts, 2000a; 2000b; and Turatto & Gandallo, 1999). Does the same pattern of performance hold when subjects must identify two items displayed simultaneously? Two studies already discussed seem to have addressed this issue: recall that both Sagi and Julesz (1985) and Kwak, Dagenbach and Egeth (1991) presented displays of two letters and found no effect of distance on identification performance. Bahcall & Kowler (1999), however, used densely packed displays of letters and found that identification of two target letters improved as the separation between the targets increased.

The first experiments in this dissertation adapt the paradigms used by Sagi and Julesz (1985) and Bahcall and Kowler (1999) to explore the causes of separation effects in attentional tasks. As I have already discussed the former study, I will now review the procedure and findings of Bahcall and Kowler (1999): Displays consisted of 24 letters arranged equally spaced around an imaginary circle (4° visual angle radius; see Figure 3). Subjects identified two target letters on each trial. A number of different methods were used to indicate the target-letter locations—these methods were blocked in separate
sessions. In some experimental sessions, target location was indicated by cues in a 24 character preview display (i.e., two letters differently colored from the rest); this display was then followed by a 24 letter target display. In other sessions, no pre-display frame was used, and subjects identified 2 red letters among green distractors (these colors also reversed roles). The target display was on for 100, 200 or 300 ms (display times were mixed within blocks). The main finding of this study was that identification of target letters improved as the separation between targets increased. This pattern held regardless of the cueing method used. A review of the data also indicates that performance was frequently better at the shortest separation tested (when letters were at adjacent positions) than at the next largest separation (when one letter separated them). In general though, it is clear that perceptibility of target letters was better when they were far apart than when they were close together.

Bahcall and Kowler (1999) offer two different models of their data, each with a vastly different assumption as to the mechanism responsible for the effects of target separation. One model asserts that because processing resources are limited, adding resources to an attended region draws (and thereby diminishes) resources from the surrounding region. The result is that attending to a specific object facilitates processing at the attended location, while suppressing processing in the surrounding region. Each attended object has an associated distribution function (defined mathematically by a difference-of-Gaussians; see Figure 4a) describing how the quantity of processing resources varies with distance from the object. These separate functions are summed to produce the processing level at each location (see Figure 4b).
This model predicts poorer performance at small separations by the following logic: When two attended objects are located near each other, they each sit within the range of the inhibitory surround of the other. Thus, summing at each location, the level of attentional facilitation at each attended location is reduced by the inhibition associated with the other object. In addition, according to this model, the level of inhibition produced around an attended object diminishes with distance from the object. Thus, as the distance between attended objects increases, they each become less and less vulnerable to the suppressive effects of the other.

A second model proposed by Bahcall and Kowler posits that performance decrements at small target separations are the result of imprecise target selection. Part of the impetus for this model comes from a study in which a decline in targeting precision was found for saccadic eye movements made near an attended location (Kowler, Anderson, Dosher, & Blaser, 1995). In this second model, the processing strength at each target location is described by a single (Gaussian) function: strength is greatest at the attended location and diminishes rapidly with distance. The key assumption under this model is that as the separation between the targets decreases, the variance of each distribution function increases. If the area under the function is fixed, then as target separation decreases (and variance increases), the distribution becomes flatter, and there is a net loss in resources at the attended location accompanied by an increase in strength at neighboring locations. As a result, this model predicts that 1) target identification performance will decline for close targets, and 2) that distractors near targets will undergo further processing when the targets are closer together than when they are far
apart (i.e., because of the wider distribution of processing strength when targets are close). This, in fact, is what Bahcall and Kowler found: subjects were much more likely to misidentify a neighboring letter for a target when there were small separations between targets than when the targets were spread far apart.

Scope of the dissertation

The effect of separation in attentional tasks obviously embodies a complex issue: In the course of this review I have described studies in which performance a) improved, b) showed no change, or c) worsened, as the distance between target objects grew smaller. If one were to perform a parametric manipulation of the possibly relevant factors (display exposure times, response vs. accuracy measures, stimuli configurations, etc.) the number of different conditions would be staggering. The procedures used by Sagi and Julesz (1985) and Bahcall and Kowler (1999) are similar enough, however, that that they allow for a reasonable comparison (see above for a detailed description of each procedure). In each study, stimuli were presented briefly and subjects made a judgment about two characters. In addition, all stimuli were presented along the rim of an imaginary circle so they were all equidistant from fixation. How is it that these studies came to find two qualitatively different patterns of responding? Using displays of 24 letters (2 targets and 22 distractors), Bahcall and Kowler (1999) found that target identification improved as the separation between them increased; whereas, in a difficult same/different task using two-letter displays, Sagi and Julesz (1985) found no effect of distance. Display density would seem to play a crucial role in separation effects.
The purpose of the first set of experiments in this dissertation was to examine whether the difference in display density was responsible for the disparate findings of Bahcall and Kowler (1999) versus Sagi and Julesz (1985). The procedure used in Experiment 1 was adapted from that of Bahcall and Kowler. However, displays were of two possible densities\(^3\): either 24 letters, as in the original experiment, or, only 2 letters, as in the Sagi and Julesz experiments. The effect of inter-target distance on letter identification was assessed for each density condition. In both the low- and high-density conditions, identification performance improved as target separation increased. The 2-letter condition produced a monotonic performance function, while in the 24-letter condition, performance was worst at the second smallest target separation. This difference in performance functions was examined in Experiment 2. Otherwise, performance for both density conditions was virtually the same at the smallest and largest inter-target separations. This finding was further investigated in Experiment 3.

After finding in the first experiment that display density is indeed a factor in separation effects, I next asked the question, At what level of processing do distractors interfere with dual-target identification? Specifically, Experiment 2 focused on whether the separation effects caused by increased distractor density are attributable to the mere physical presence of distractors (the perceptual interference hypothesis), or to a higher-level characteristic of distractors, i.e., their infomational content (the response

---

\(^3\) In the first experiment, display density is conflated with stimuli numerosity: that is, the 24-letter display has both more numerous stimuli and (generally) higher visual density than the 2-letter condition. For consistency, the term “density” will be used throughout the dissertation to describe the difference between the 2- and 24-letter conditions.
interference hypothesis). These hypotheses were pitted against each other by including a condition in which the target letters were presented among number distractors (as opposed to letter distractors which were used in Experiment 1). The logic behind this manipulation was that if low-level characteristics are sufficient to evoke the density effects found in Experiment 1, then number distractors should yield performance similar to that obtained with letter distractors. The results of Experiment 2 showed that the presence of number distractors produced performance functions that resembled those of the no distractor condition of Experiment, supporting the response interference hypothesis.

The third experiment was an attempt to explain the unexpectedly good overall performance in the dense display condition in Experiment 1. Many studies using the visual search task have shown response times to increase with the number of distractor elements (e.g., Treisman & Gelade, 1980), so the efficiency of target identification for the high density condition in the first experiment (which was similar to the performance at the low density condition), was a curious finding. In Experiment 3, targets were either presented in a circular array of 22 distractors (as in Experiment 1), or among only four distractors, with each target between a pair of flanking letters. The results of Experiment 3 support the hypothesis that masking of targets by distractors in Experiment 1 was reduced due to grouping or collateral masking among the distractors (Duncan and Humphreys, 1989).

The second set of experiments (4 and 5) investigated the effect of distractor placement around two attended locations. If display density modulated separation effects
(compare Bahcall & Kowler, 1999 with Sagi & Julesz, 1985). was this the direct result of distracting letters being placed in between the target locations, or was local crowding by distractors, regardless of their arrangement, sufficient to cause the added interference? In Experiment 4, displays consisted of 20 letters on the circumference of an imaginary circle. In one condition, two distractor letters were between the targets while blank spaces flanked the targets; in the other condition, the distance between the targets was the same, but, distractors flanked the target pair, and the space between them was empty. (Only one inter-target-separation was used). Performance was the same whether or not the space between targets was occupied, supporting theories that assert that locations between attended locations are no more attended than locations external to the attended locations.

The fifth and final experiment further explored the issue addressed in Experiment 4; specifically, this experiment looked at whether attention is allocated along the axis connecting attended locations. In two experimental conditions, flanking letters were either arranged on either side of an imaginary line connecting the two targets (distractors in this condition were placed both between targets and external to them), or the flankers were placed orthogonally to this (imaginary) line. In contrast with the previous experiment, in Experiment 5, it was found that distractor interference was determined by the spatial relationship between attended objects. An explanation for Experiments 4 and 5 is proposed asserting that attention is allocated in an elliptical shape, with its major axis determined by the location of the attended objects.
To summarize, the first set of experiments explored the effect of separation in an attentional task, and the role of display density in producing such effects. These results were then considered in light of the theories and models discussed in the previous sections. The second set of experiments investigated how the spatial relationship between attended locations influences interference brought about by nearby distractor items. Theories of attentional allocation are used to frame these findings. These results are then related to recent findings in neurophysiology.
Figure 3. Stimuli configuration and procedure, Bahcall and Kowler (1999). (Bold letters represent target locations. Targets for the trial depicted were J and P.)
Figure 4a. Difference of Gaussians. The bold curve plots the difference between the two Gaussian curves for each point along the abscissa. For the plot shown, the broad curve is subtracted from the tall curve.

Figure 4b. Sum of two differences-of-Gaussians. This curve plots the sum of two difference-of-Gaussian curves, such as that shown in Figure 4a. The distance between the two peaks indicates the approximate displacement between the individual curves.
Chapter 3: The Experiments, Part I

Experiment 1

To account for their data, Bahcall and Kowler (1999) proposed a pair of models: one posited that there is an area of facilitation at an attended location, surrounded by an area of inhibited processing; the second model asserts that as targets get closer to each other performance degrades due to imprecise target selection. In both models, the processing strength across display locations is defined by a Gaussian-like function, (or the sum of two such functions in the case of the inhibitory-surround model). What factors determine the particular size of the region of interest—that is, the spread of the function? Bahcall and Kowler, for instance, found that the method by which the targets were cued modulated the variance of these functions. The qualitative difference between the findings of Bahcall and Kowler (1999: 24-letter displays) versus Sagi and Julesz (1985: 2-letter displays) suggests that the dispersion of the processing function(s) may also depend on the number of elements in the display.

The first experiment tests the hypothesis that display density plays a role in separation effects in an attentional task. In addition, since the categorical effects of separation have not been reliably established empirically (no effect of separation? worse performance at small separations?) another aim of the experiment was to discern what is the qualitative effect of separation in a dual target identification task. As Bahcall and Kowler’s finding—that performance improves as target separation increases—was somewhat surprising, it was important to replicate that study (the present procedure is
nearly identical to theirs, see Methods). In addition, by using a similar procedure, but
with 2-letter as opposed to 24-letter displays, the influence of density could also be
assessed by comparing the results for the two conditions.

Based on the results of the two pertinent studies cited above, a reasonable
prediction was that 1) for the 24-letter condition, identification performance would be
substantially worse for short rather that longer inter-target separations; as found by
Bahcall & Kowler (1999), whereas 2) for the 2-letter condition, it was uncertain whether
performance would vary at all with inter-target separation, as Sagi and Julesz (1985)
reported no such distance effect. However, if the inhibitory surround is impacted by
display density (e.g., if the magnitude or range of suppression is greater for denser
displays) then one could conservatively predict a response function similar to the 24-letter
condition, though with a smaller net effect of target separation.

A note on methods. Throughout the course of the dissertation experiments, two
different measures are used. In Experiments 1 and 5, an adaptive staircase method was
used in which display-to-mask SOA was varied, while in Experiments 2, 3 and 4,
accuracy was measured, with display-to-mask SOA fixed. Though the general procedures
are very similar, varying the measure allows for some degree of generalization of the
results, rather than attributing any one finding to a particular methodology. It is
especially convincing when a finding can be replicated across experiments using slightly
different measures—as is done in Experiments 1 and 2.

Method
Participants. Four Johns Hopkins University students participated in this experiment. They received class credit for taking part.

Apparatus and stimuli. Experimental programs were run using Micro Experimental Laboratory software (Schneider, 1988) on a Dell Dimension XPS D233 microcomputer, which also recorded responses. Displays were presented on a 17" (diagonal) SVGA color monitor driven by a Matrox Millennium II color graphics card. Viewing distance was 94 cm with head position maintained in a chinrest. The configuration described here was used in all experiments.

Stimuli were displayed on a black screen. The fixation crosshair that began each trial was light gray. The stimulus letters were equally spaced around the rim of an imaginary circle about 8.5° visual angle in diameter, centered at fixation. There were either 2 or 24 letters present in a display. The density condition varied by session. In degrees of visual angle, the dimensions of the letters were .56 x .38. The measure from the center of one letter to the next was 1.1° visual angle. In the 2-letter condition, both letters were the same color (red or green). In the 24-letter condition the two target letters were one color (green or red), and the 22 distractor letters were the other color (red or green). In all experiments, the luminance of the red figures was set at 11.77 cd/m²; the green luminance was then set via flicker photometry to match this value.

There were five different inter-target distances presented (the same distances were used in each density condition). In polar coordinates around fixation, the separations were 15°, 30°, 60°, 120°, and 180°; measured "as the crow flies" across the display, the separations were: 1.1, 2.2, 4.3, 7.4, and 8.5 degrees visual angle. Note that in the circle
of 24-letters, adjacent letters are separated by 15°, polar coordinates (24 \times 15° = 360°); hence, a 30° separation indicates targets with one distractor letter between them. 60° indicates targets with three distractor letters between them, etc. Target locations were randomized across trials; i.e., on each trial, one location of the 24 possible locations was chosen at random. The second target location was obtained by moving the appropriate number of steps around the circle based on the prescribed distance condition for the trial.

**Procedure.** Trial events are depicted in Figure 5. Each trial was initiated with a keypress, causing the fixation crosshair to be displayed at the center of the screen. Participants were instructed to fixate the crosshair throughout each trial. (This instruction was emphasized throughout each session to limit the occurrence of eye movements.) After 1000 ms, the critical display appeared and remained on screen for a period determined by the PEST algorithm (Parameter Estimation by Sequential Testing, Taylor & Creelman. 1967; see Appendix for a full description of the PEST procedure used in these experiments). The displays were then masked: each letter was replaced by a different letter of the same color. The masking display remained on screen for 500 ms and was replaced by a pair of Q’s, one at each of the target letter locations. One Q was red and the second Q was green. This signaled participants that they could now respond. To respond, participants typed one target letter on the keyboard. After the first response was entered, the two Q’s switched colors, to indicate that the first response had been successfully recorded. Participants then entered their second response. The two target
identities could be entered in any order. Feedback was then provided by reshowng the original test display. The feedback display remained on screen until the participant pressed a key, which then began the next trial.

**Design.** The number of letters present in a display (2 or 24) was constant within a session. Each participant performed five sessions of experiments over thirteen days (or fewer). The first session was always the 24-letter condition, and was regarded as practice. The four experimental sessions were divided into two 2-letter and two 24-letter conditions, and were performed in the following order: Subjects 1 and 2: 2-24-24-2 Subject 3 and 4: 24-2-2-24. Specific colors were assigned to targets and distractors for an entire session (e.g., targets were always green and distractors always red for the session). Target color was counterbalanced across sessions (within subjects) and this order was counterbalanced across subjects.

Each session consisted of five PEST runs, one run per inter-target separation (a run being defined as a sequence of adjusted trials for a particular inter-target distance condition). Trials in each condition were interleaved with each other, in a quasi-random order. There were 96 trials in each distance condition, or 480 trials in a session. Each run continued even if it had reached a convergence level (SOA) so there were always 96 trials per run. As per the multiple run PEST procedure described by Kreeiman and

---

4 Bahcall and Kowler (1999) had subjects enter target identities in a specified order, however, in pilot studies, it was found here that it was especially difficult to identify specific targets when they were close together. The targets could be identified, but their locations were unclear. Thus, throughout the current study, targets could be identified in any order. (See also, footnote 5.)
Kaplan (1973, see Appendix) the data for a particular run consisted of the resulting adjusted SOAs at trials 48, 64, 80, and 96.

Sessions began with 40 practice trials (8 trials per distance-condition) with display-mask SOA going from 700 ms (first 15 practice trials) to 500 ms (next 10 practice trials) to 300 ms (last 15 practice trials) in the course of the practice block. PEST adjustment of SOAs did not start until the practice trials were finished. Subjects were given breaks after every 60 trials.

Results

Experiment 1 employed an adaptive staircase procedure, with SOA’s adjusted to produce 80% accuracy in identifying both targets. (Accuracy data is presented in Table 1). Mean SOAs for each density condition crossed with inter-target distance are presented in Figure 6. (Means for each subject by condition by distance are shown in Figures 7a and 7b). Higher SOA value suggests worse performance than a lower SOA (a longer exposure was required for equal accuracy). In Figure 6, both lines show a general tendency for performance to improve as the separation between targets increased: The 2-letter condition shows a monotonic increase in performance (that is, lower SOAs) as separation increases. In contrast, the 24-letter condition has a noticeable dip in performance at the second closest separation (2.2°) compared with the smallest separation (1.1°). All four subjects demonstrated this dip in performance (Figure 7a). A 2 (density) x 5 (inter-target distance) ANOVA confirmed these observations. There was a significant main effect of distance [$F(4,12) = 8.4, p < .005$] supporting the finding of general improvement of performance with increased separation. The main effect of density was
also significant \([F(1,3) = 13.5, p < .05]\), largely resulting from the difference between the density conditions in SOAs for the 2.2° and 4.3° separations. Finally, there was a significant interaction between density and distance \([F(4,12) = 3.7, p < .05]\), indicative of the different pattern of SOAs—monotonic for the 2-letter condition versus a performance dip for the 24-letter condition.

Post-hoc (Bonferroni) contrasts were performed comparing the differences between mean SOAs (at 80% accuracy) across the inter-target distances within each density condition. For the 24-letter condition, all differences were significant \((p < .05)\) except for those between the 1.1° and 4.3° separations, and between the 7.4° and 8.3° separations. Note that the SOA at 2.2° was significantly different from both of the next closest separations (1.1° and 4.3°), indicating a real dip in performance at this separation. For the 2-letter condition, all contrasts were significant \((p < .05)\) except for all those between consecutive separations (1.1° and 2.2°, 2.2° and 4.3°, 4.3° and 7.4°, 7.4° and 8.3° were not significant) and that between 4.3° and 8.3°. This pattern concurs with the monotonicity of the function: performance rose steadily, but gradually, as separation between the targets increased, eventually reaching asymptote by 7.4° of separation.

<table>
<thead>
<tr>
<th></th>
<th>1.1°</th>
<th>2.2°</th>
<th>4.3°</th>
<th>7.4°</th>
<th>8.3°</th>
</tr>
</thead>
<tbody>
<tr>
<td>2-letter condition</td>
<td>76.8</td>
<td>79.3</td>
<td>74.5</td>
<td>77.6</td>
<td>82.3</td>
</tr>
<tr>
<td>24-letter condition</td>
<td>78.1</td>
<td>79.6</td>
<td>80.6</td>
<td>80.1</td>
<td>82.3</td>
</tr>
</tbody>
</table>

Table 1. Experiment 1 accuracy by experiment conditions. Data is for density x distance in degrees of visual angle between targets, for trials 48 through 96 in each PEST run.
Discussion

The results of Experiment 1 replicate Bahcall and Kowler's (1999) finding of improved performance with increased separation between targets—a pattern that held up for both the sparse (2-letter) and dense (24-letter) display conditions. The finding that distance was a factor even for sparse displays contradicts the findings of Sagi and Julesz (1985) and, later, Kwak, Dagenbach, and Egeth (1991); using displays of only 2 items, those authors concluded that separation did not modulate performance of a difficult dual-target discrimination task. Note that in both those studies, there was, in fact, a decrement in performance at the smallest inter-target separations, which the authors attributed to lateral masking. However, as shown in the present study and by Bahcall and Kowler, effects of inter-target separation are not simply the result of a low-level lateral masking: In the circular arrays of 24-letters, target letters were flanked on two sides by distractors. Thus, even though masking by neighboring letters was held constant, performance improved as the separation between attended letters increased, indicating a form of attentional suppression at short inter-target separations.

A reexamination of the data of Sagi and Julesz (1985a) and, Kwak, Dagenbach, and Egeth (1991) suggests an explanation for the apparent discrepancy between those studies and the current Experiment 1. Both Sagi and Julesz (1985, their Figure 4) and Kwak, Dagenbach, and Egeth (1991, their Figure 2) show performance to be poorest at smallest separations and not yet at asymptote until inter-target separations reach at least 4° visual angle; a figure normally considered beyond the range of low-level masking. Thus, these studies may indeed have been showing evidence of an inhibitory surround.
(e.g.) operating at some distance from the attended locations. This possibility was not raised in either article probably because the aim of each study was to assess whether attentional shifts were analog or discrete: in a manner of speaking, they were basically conducting one-tailed tests of whether shifts of attention were unaffected by separation, or benefited from smaller separations between targets. Neither set of authors expected to find the attentional-type masking described by Bahcall and Kowler (1999), and thus, did not consider this possible interpretation of their results.

Although performance under both density conditions in Experiment 1 tended to improve with inter-target separation, the differently shaped functions for each density condition suggest that this factor indeed plays some role in separation effects. Performance with 2-letter displays tended to uniformly improve as the distance between the targets increased, whereas with 24-letter displays there was a pronounced dip in performance at the second smallest (2.2° visual angle) inter-target separation—i.e., the condition in which one distractor letter is between the targets. Bahcall and Kowler (1999) also frequently reported a similar dip in performance, depending on cueing conditions and the particular subject. Those authors did not propose that any one stimulus-related factor was responsible for this dip; they were, however, able to adjust parameters in their inhibitory surround model (i.e., by increasing the variance) to simulate this pattern of performance. By extension to the current study, perhaps selectively attending to a denser stimulus array entails widening the region of inhibition (contingent, of course, on the veridicality of that particular model of attentional processing, in the first place).
What characteristics of the densely populated display are responsible for the dip in performance at small inter-target separations? Experiment 2 addressed whether it is the mere physical presence of the distractors, rather than their confusability with targets, as members of the same object category that is paramount. Experiments 4 and 5 then examined whether the effect of distractor elements stems from the way they are arranged (i.e., being placed between, as opposed to around, targets), or, simply from their proximity to the target letters.
Actual targets were not bold, but were a different color from distractors.

Figure 5. Stimuli and procedure for Experiment 1.
Figure 6. Experiment 1 data for all subjects combined. Note that y-axis is flipped (higher values are toward bottom) so relatively superior performance is toward the top of the graph.
Figure 7a. (top) Experiment 1 data: 24-letter (High density) condition, by subject.
Figure 7b. (bottom) Experiment 1 data: 2-letter (Low density) condition, by subject.
Experiment 2

If distractors interfere with dual-target perception, at what level of processing do they interfere? Eriksen and colleagues (Eriksen & Hoffman, 1972, 1973; Eriksen & Yeh, 1985; Eriksen & St. James, 1986) have shown that responses to a briefly presented target are slowed when the target is located near distractors associated with a competing response, suggesting interference at a response or decision level of processing. In contrast, Taylor and Brown (1972) demonstrated that, even when subjects were permitted unlimited viewing time, identification of a letter was impaired when the letter was placed near other letters, suggesting interference occurring at a perceptual level of processing. Reexamining the results of Experiment 1, it is then uncertain whether the effects of density should be attributed to low- or high-level aspects of the accompanying distractors.

The purpose of Experiment 2 was to directly test two hypotheses concerning the level at which distractors interfered in the dense display condition of Experiment 1, to produce the dip in performance at 2.2° (visual angle):

The perceptual interference hypothesis states that an alteration in display density leads to an adjustment of the regions of facilitation and inhibition near each attended location. This would correspond to resetting the variances in Bahcall and Kowler's (1999, inhibitory surround) model as described above. The dip in performance could result if the depth and reach of the inhibitory-surround becomes greater as display density increases, making the area of maximum inhibition extend farther from an attended location. Thus, whereas in a two element display, the trough of suppression was greatest at 1.1° (visual angle) from an attended location, in a dense display, the trough shifts to 2.2° (visual
angle), and likewise, makes target identification most difficult at that inter-target separation. (Note that in modeling their own data, Bahcall and Kowler provide a number of instances in which model parameters produce just this pattern of performance). One could hypothesize that re-tuning of the surrounding activation profile occurs to reduce the interference or competitive influence (in the terms of the FeatureGate model) of multiple distractor elements near an attended location.

The response interference hypothesis states that confusion between targets and nearby distractors increases as inter-target separation decreases. One explanation why such confusion might increase when targets are close to each other is provided by the imprecise targeting model of Bahcall and Kowler: as the separation between targets becomes smaller, distractors become more and more likely to be selected for processing. The detrimental effects of such mis-selection increase if the distractors and targets are of the same category, since distractors cannot be ruled out of consideration as possible targets once they are mistakenly identified.

In Experiment 2 there was only one display density used (always 24 letters). however, there were two distractor-type conditions: distractors were either all letters (same as Experiment 1) or distractors were all numerals. The letter distractor condition was employed to replicate the findings of Experiment 1 using a slightly different procedure (fixed SOA instead of an adjusted staircase). The numeral distractor condition pitted the two hypotheses directly against each other, because they make different predictions as to the outcome. The perceptual interference hypothesis leads to the prediction that the effect of numeral distractors will be similar to that of letter distractors.
(producing the characteristic dip in performance) because both similarly occupy the
display. The response interference hypothesis predicts that there will be no performance
dip in the numeral distractor condition: the numbers will not be confused with letter
targets, because numbers are not members of the possible response set, which consists
solely of letters.

Method

Participants. Seventeen members of the Johns Hopkins University community
participated in this experiment. Subjects were paid for taking part. Eight were originally
assigned to the letter-distractor condition, and eight were assigned to the digit-distractor
condition. One subject in the letter-distractor condition was later replaced because their
accuracy was less than 10%.

Stimuli. Displays were the same as Experiment 1, except as noted. All displays
consisted of 24 elements arranged in a circular array with a diameter of 8.5° visual angle
(see Figure 8). Targets were always letters. There were two distractor conditions: in one
condition, all stimulus elements were letters (as in Experiment 1), in the other condition,
all non-targets were digits, as were the masking elements. The digits were taken from the
same font set as the letters (MEL font fg-25) and were approximately the same size as the
letters (.56° x .38°). Inter-target separation could be any of five possible values (same as
Experiment 1). Luminances were equated by flicker photometry.

Procedure. The target display remained on screen for a fixed period of time, 225
ms, before being masked. (Unlike the staircase procedure of Experiment 1, accuracy was
the dependent measure in this experiment). Otherwise, the procedure used was identical to that of Experiment 1.

*Design.* Each subject was assigned to either the letter-distractor or digit-distractor condition. Two sessions were performed by each participant over the course of 5 days (or fewer). Sessions were composed of four experimental blocks of 90 trials each. Blocks consisted of 18 trials at each of the five target separations, presented in a quasi-randomized sequence. Target and distractor color (red or green) were fixed within a session, but were varied across sessions. The order of color conditions was counterbalanced across subjects. Sessions began with 40 practice trials (8 trials per distance-condition) with display-mask SOA going from 700 ms (first 15 practice trials) to 500 ms (next 10 practice trials) to 250 ms (last 15 practice trials) in the course of the practice block. The first two blocks of the first session were also considered practice. Subjects were given the opportunity to take breaks midway through, and at the completion of each block.

*Results*

In Experiment 2 the dependent measure was accuracy in identifying the two target letters (i.e., a trial was counted as “correct” only if both targets were correctly identified). Mean subject accuracies for each distractor condition crossed with inter-target distance are presented in Figure 9. (Note that the y-axis in this figure refers to the proportion of trials on which both targets were correctly identified.) Both lines show a positive slope, indicating that performance generally improves (increased proportion correct) as inter-target distance increases. An ANOVA conducted separately for each experiment
condition found a significant effect of inter-target distance [letter distractors: F(4.28) = 37; P < .001; numeral distractors: F(4.28) = 12.6; P < .001].

In Figure 9, the line for numeral-distractors depicts a monotonically increasing function, whereas the letter-distractor condition shows a dip in performance at the second smallest inter-target separation (2.2°). The accuracy data for the eight subjects in each experimental condition are shown in Figure 10a and 10b. Post-hoc (Bonferroni) contrasts) were performed across inter-target distances within distractor-type condition. For the number-distractor condition, no contrasts were significant (p < .05) between any consecutive separations (1.1° and 2.2°, 2.2° and 4.3°, 4.3° and 7.4°, 7.4° and 8.3°). For the letter-distractor condition, none of the contrasts between consecutive separations was significant, except for that between 2.2° and 4.3°, indicative of a dip in performance for the 2.2° separation, not present in the number-distractor condition.

Discussion

The findings of Experiment 2 support the response interference hypothesis as put forth in the experiment introduction. As in Experiment 1, a dip in performance at the 2.2° (visual angle) target separation occurred when distractor items and targets were of the same category (letters). However, a monotonic function of improving performance with increased target separation prevailed when numerals served as distractors: a pattern resembling that found in the sparse 2-letter displays of Experiment 1. The added interference associated with letter distractors when targets are separated by short distances (in Experiments 1 and 2) must, therefore, not have resulted from interference at the perceptual level, but rather, from response competition between targets and distractors.
The results of this experiment can be explained through a Guided Search (Wolfe, 1994) reading of the imprecise targeting model (Bahcall & Kowler, 1999): Based on a Guided Search-type explanation, in the present paradigm, bottom-up (color salience) and top-down (target color) factors combine to produce peak activations at the two target locations. As the targets get closer together, however, their salience is diminished, and the activation at each target peak begins to flatten out. (This description of the change in activations with decreased separation, taken directly from Guided Search, is almost identical to that described in the imprecise targeting model.) The resulting smaller difference in activation between the targets and surrounding locations (as compared with the targets being spread wide) increases the likelihood that distractors near the targets will inadvertently be selected for further processing. Mis-selection of distractors (as a result of their increased activation relative to the targets), may occur equally for numeral and letter distractors, but is less problematic for numeral distractors, because numbers are not members of the set of possible targets and thus are not confusable with a target. Whenever a number is seen, it is disregarded. However, if a letter distractor is mis-selected, because it is a member of the target set, it could be encoded as a target, and yield an identification error. This explanation is supported by Bahcall and Kowler’s finding of increased neighbor errors (that is, adjacent distractor letters identified in place of targets) with smaller inter-target separations.⁵

⁵ Neighbor errors could not be assessed in this study because subjects were free to enter target identity in any order. As a result, it could not be determined which neighbors were associated with which response.
But why is performance better when the targets are adjacent than when one
distractor letter separates them? Again, in the language of Guided Search, when the
targets are adjacent, a single region of high activation results, which can easily be selected
for processing. Attention will be restricted to this highly salient region, making the
process of identification less susceptible to neighbor intrusions than when the targets are
separated by one letter. (However, as in the 2-letter condition of Experiment 1, accuracy
in identifying these adjacent targets will be degraded as compared to the wider inter-target
separation conditions as a result of local attentional suppression.)

An alternate reading of the findings of Experiment 2 may be derived by
combining Bahcall and Kowler’s inhibitory-surround model with a conceptual inhibition
hypothesis. Perhaps the area of inhibited processing around an attended item extends
farther from the attended location when distractors are confusable with the targets than
when distractors are categorically different from the targets or not present at all. In this
view, the activation distribution around an attended item is not modulated by the physical
density of visual stuff, but by the density of relevant abstract information in the display.
Thus, when targets are letters, a field of competing distractor letters yields a wider region
of suppressed processing around an attended target letter than when numerals serve as
distractor items. Some precedent for inhibition occurring at a conceptual level comes
from a study by Egeth and Santee (1981; see also Walley and Weiden, 1973), who found
that identification of a briefly presented capital letter (“A”) was worse when the letter was
presented alongside a letter sharing only the same name (“a”) than when presented near
other letters (“K”, “E” or “e”).
Finally, one could make the case that the effects of separation in attention tasks result from both an inhibitory surround, and imprecise targeting operating in tandem. The first premise of this revised scheme is that the attentional locus is accompanied by a surrounding region of suppressed activation, which is independent of display density. Evidence for this comes from the first two experiments and from Mounts (2000a). When only two letters were present (as in the present Experiment 1), we may assume that performance decrements at short target separations were the result of the inhibitory surround, as the scarcity of items in these displays would argue against an imprecise targeting explanation for distance effects. However, the similar performance functions obtained when there were no distractors (in the present Experiment 1) or when distractors were numerals (in Experiment 2) suggests that the distribution of the inhibitory surround was not effected by the distractors. This concurs with Mounts (2000a) who found that the effects of inter-target separation were equivalent whether displays consisted of 28 or 56 items. The second premise of this scheme is that imprecise targeting comes into play only when there are noise elements present in the display. Even so, the consequences of imprecise targeting vary depending on whether or not the distractor items are confusable with the targets (i.e., from the same response set.). This premise is also supported by the data from Experiments 1 and 2.

This final rendering of the first two experiments may in fact be the most satisfying. The inhibitory surround (as opposed to imprecise targeting) explanation for most separation effects is in agreement with recent studies that have shown performance to be impaired near an attended location when no other elements are present in the
display, such as Experiment 1 (2-letter condition) or Cave and Zimmerman (1997), in which a post stimulus probe was presented in an otherwise empty display. However, the apparent (additional) effects of response competition are better explained by imprecise targeting. The findings of the first two experiments for conditions in which letters served as distractors are reminiscent of those of Eriksen and colleagues (Eriksen & Hoffman, 1972, 1973; Eriksen & Yeh, 1985; Eriksen & St. James, 1986) as both describe situations in which the process of object identification receives interference, in the form of response competition, from unattended neighbors. This "mixed model" concedes that in the case of separation effects, there may be no "unified theory", though we understand how different factors contribute to these effects.

Stepping back from the hypothesizing, the key finding of the first two experiments needs to be restated: as the distance between two attended objects increased, subjects’ ability to identify the objects improved. And, this performance function obtained whether or not distractor items were present, and whether or not distractors and targets were of the same conceptual category. The data, however, don’t appear to favor one model over another. Overall, though, these findings are quite consistent and conclusive in 1) corroborating Bahcall and Kowler’s (1999) finding in regard to the positive effect of increased separation in dual-attention tasks, and 2) generalizing this finding to both dense and sparse stimulus displays.
Figure 8. Stimuli and procedure for Experiment 2. (Bold letters represent target locations, which were a different color from distractors.)
Figure 9. Experiment 2 data for all subjects combined.
Figure 10a. (top) Experiment 2 data; letter-distractor condition by subject.
Figure 10b. (bottom) Experiment 2 data; number-distractor condition, by subject.
Experiment 3

An unexpected result of the first experiment was the comparable level of performance for 24- and 2-letter displays under a number of conditions. This was unexpected because previous studies have often shown that as distracting items are added to a display, performance of attentional tasks worsens. The most basic example of interference from distractor items is in the form of a filtering cost as demonstrated by Kahneman, Treisman and Burkell (1983). In their Experiment 1, subjects simply had to identify a word written above or below central fixation. Compared to when the word was presented alone, the mere presence of a small rectangular patch of dots at the opposite side of the display was sufficient to slow responding to the word.

In addition, as has been discussed throughout this dissertation, adding distractor items creates the possibility for crowding and masking. Recall that Bouma (1970) found increasing interference in a letter identification task as one and then two distracting letters were placed around a target letter. As also noted earlier, Strasburger, et al. (1991) found that four flanking letters—2 on each side of a letter—had larger masking effects than just two flanking letters—one on each side of the target letter. It is unlikely that the increase in interference resulted from masking per se because the extra distractors were placed on the peripheral side of the first two (relative to the central target). Again, this would be an example of crowding, which is another kind of filtering cost.

So, what led to the surprisingly good performance in the 24-letter condition in Experiment 1? The simplest explanation perhaps, is that the difference in color between the targets and distractors was sufficient to cause the targets to pop-out from the display.
as predicted by Feature Integration Theory (Treisman & Gelade, 1980). A hypothesis based on the processing theory of Duncan and Humphreys (1989) is that interference was reduced by grouping among same-color distractors, and the subsequent suppression of their processing (see also Banks, Bodinger, & Illige, 1974; Banks & Prinzmetal, 1976; Banks, Larson & Prinzmetal, 1979; Bundesen & Pedersen, 1983). According to this theory, a large number of similar contiguous distractors would only minimally interfere with identification of highly salient targets (i.e., dissimilar to the distractors).

Another (although somewhat related) explanation for the comparable performance for displays with and without distractors implicates crowding among the distractors themselves. Similar adjacent distractor letters may serve to mask each other. As shown by Taylor and Brown (1972), letters embedded in a string of letters were less perceptible than equally eccentric letters at the end of such a string. A string of similar items are somehow perceptibly “smushed” or blended together, making any one of them difficult to identify. However, when freed of these blending effects on one side (i.e., when the letter is on the end of a string), the letter becomes easier to identify.

The main goal of Experiment 3 was to investigate the role of distractor grouping in dual-target identification. In the experiment there were two density conditions. One was the 24-letter condition of the previous experiments; in the other condition, each target was displayed alongside two distractors, one on either side of it (making two sets of three letters, or the 6-letter condition). Targets were presented at two separations: 4.3° (60° polar coordinates) and 7.4° visual angle (120° polar coordinates).
It is predicted that grouping among distractors will reduce their interference with target identification (as per Duncan & Humphreys, 1989, e.g.), making performance better in the 24-letter condition than the 6-letter condition. (In the 6-letter condition the distractors are separated, which reduces the opportunity to group with other adjacent distractors.)

**Method**

*Participants.* Four members of the Johns Hopkins University community participated in this experiment. Subjects were paid for taking part.

*Stimuli.* Displays and procedures were the same as Experiment 1, except as noted. Displays had either 24 letters or 6 letters present. The 24-letter arrangements were identical to those of the first experiment (2 targets, of one color + 22 distractors of a different color). In the 6-letter arrangement, letters were clustered into two sets of three, with each cluster consisting of a target letter flanked by a distractor letter on either side (see Figure 11). In both arrangements, all letters were placed along the perimeter of an imaginary circle having a radius of 4.25° visual angle. The spacing between adjacent letters was the same as in Experiment 1.

Two different target separation distances were tested: 4.3° visual angle (60° polar coordinates), and 7.4° visual angle (120° polar coordinates). These translate into there being either three or seven intervening letter slots between targets along a circle divided into 24 parts. Thus, for the 6-letter arrangement, when the target separation was 4.3° there was one empty slot between the two three-letter clusters. As in the previous experiments, targets were all one color (red or green), and distractors were the other
color. The placement of the two three-letter groups around the display was (quasi-) randomized, although their placement relative to each other was set according to experimental conditions (i.e., 4.3° or 7.4° visual angle).

Procedure. The sequence of events in a trial was the same as in Experiment 2: i.e., the critical display remained on screen for a fixed period of time: 225 ms. (The PEST procedure was not used here; the dependent measure was accuracy, as in Experiment 2.)

Design. Each participant performed two sessions of experiments over seven days (or fewer). Sessions were composed of four experimental blocks of 90 trials. A single display arrangement (6- vs. 24-letters) was maintained throughout a block, and there were two blocks of each display arrangement per session. The order of display-arrangement conditions was counterbalanced across sessions and between participants. Target separation was mixed within blocks, and there were 45 trials at each target separation per block. Target and distractor color was fixed within a session, but was varied across sessions. This factor was also counterbalanced across subjects. Each session began with 40 practice trials (20 trials per display-arrangement condition) with display-mask SOA going from 700 ms (first 15 practice trials) to 500 ms (next 10 practice trials) to 250 ms (last 15 practice trials) in the course of the practice block. The first two blocks of the first session were also considered practice. Subjects were given the opportunity to take breaks midway through, and at the completion of each block.

Results

Mean accuracy measures for each condition in Experiment 3 are presented in Figure 12a; data for the four individual subjects are presented in Figure 12b. Visual
inspection of the figures indicates a clear benefit for the 24-letter condition over the 6-letter condition, and better performance for more widely separated targets. A 2 (display-arrangement) x 2 (inter-target distance) ANOVA confirmed the presence of a significant main effect of arrangement (24- versus 6-letters) \[ F(1,3) = 36.7, (p < .01) \]. The effect of inter-target distance showed a significant trend \[ F(1,3) = 19.87, (p = .094) \], while the interaction between arrangement and inter-target distance was not significant \[ F < 1 \]. Looking at Figure 12b, the data for all four subjects show a consistent pattern, echoing these findings.

**Discussion**

Despite the large difference in color between the distractors and targets, the flat search slope associated with target pop-out did not occur in experiment 3. In fact, the finding of improved performance with increasing distractor number is opposite what has been found for visual searches with randomly placed distractors (but see Bacon & Egeth, 1991, who also found negative search slopes due to grouping). The results suggest that some form of grouping among same-color distractor letters played a role in facilitating target identification in the 24-letter condition (i.e., in both Experiment 1 and 3). As a result, in the 6-letter condition, in which the general sparseness of distractors, and the separations between them, significantly reduced any chance for grouping, performance was worse than for the 24-letter condition.

The performance advantage for the 24-letter condition over the 6-letter condition was approximately the same for both the shorter and longer inter-target separations. This was somewhat surprising because at 4.3° separation, the number of distractors between
the targets only changed from three to two across the density conditions. Apparently, even though there was only a minor change in the letter arrangement between density conditions for the closely spaced targets, other factors, indirectly related to the reduction in grouping operated to impaired performance in the 6-letter condition: In the 6-letter condition, distractors only had neighbors on one side, and wide space on the other. Thus, they were more visible than when they were crowded between two distractors (Bouma, 1970; Taylor & Brown, 1972). The distractors were also grouped with the target letters (in each three-letter cluster, 6-letter condition), increasing the likelihood of a distractor being selected for processing with a target (Kahneman & Henik, 1981). In addition, targets were less salient, as they were not displayed against a homogeneous ring of distractors.

Experiment 3 addressed an issue which, while not central to the current thesis, still needed resolution in order to better explain the data of Experiment 1: How can target identification performance be so similar for displays having zero or 22 distractors? Regardless of the precise explanation for the present results, the importance of Experiment 3 was in confirming that adding distractors to a display does not imply that poorer performance will necessarily follow in attentional tasks. On the contrary, a crowded collection of homogeneous distractors might be grouped, suppressed and perhaps even blended together perceptually to, in effect, form a background from which salient targets easily emerge. It is when a distractor is left on its own, unmasked by its peers, that the opportunity for trouble arises.
Figure 11. Experiment 3 stimuli. (Bold letters represent target letters, which were a different color from distractors.)
Figure 12a. (top) Experiment 3 data; all subjects combined. Figure 12b. (bottom) Experiment 3 data, by subject. Each subject has two lines, one for each condition, and one line-type (i.e., broken, dotted, etc.) is assigned to each subject.
Chapter 4: The Experiments, Part II

Experiment 4

In the first part of the dissertation, it was shown that when attention is allocated to two locations, the effects of inter-target separation on letter identification are mediated by distractors in the display. However, the effects of distractor elements may have been dependent on the specific arrangement of stimuli in these experiments. Displays in Experiments 1 and 2 always consisted of a circular arrangement of elements. As a result, at the shorter inter-target separations, distractors occupied locations both between the targets (though, not when the targets were adjacent, of course) and surrounding the targets. (i.e., distractors were near an imaginary straight line connecting the two targets and extending beyond them). In contrast, at the longer inter-target separations, distractors were neither directly between nor did they surround a pair of targets (i.e., no distractors were on the imaginary line connecting the two targets). Thus, even though the letters were always arranged in a circle, the spatial relationship between the target pair and neighboring distractors varied. It may be that effects of density are primarily the product of distractors intervening between two targets, as opposed to sheer numerosity of display elements. This issue was investigated in Experiments 4 and 5 by assessing the effects on performance of varying the arrangement of distractors between and around a pair of attended locations.

A pertinent issue with regard to distractor effects in a dual identification task concerns whether attention can be split into noncontiguous regions. A theory positing a
single spotlight of attention would assert that the space between two attended items is also processed to some extent (Eriksen & Yeh, 1985; Eriksen & St. James, 1986). (This would hold whether the spotlight expanded to encompass both attended locations simultaneously, or if the spotlight traversed the locations sequentially...provided it remained turned on, in between.) Conversely, distractors outside of this attended range of locations would be less likely to call on attentional resources. Hence, a single spotlight theory predicts that (in the present task) distractors intervening between attended locations produce more interference than more peripherally located distractors. A number of studies, however, have shown that the attentional beam can in fact be split (Bichot, Cave, & Pashler, 1999; Egly & Homa, 1994, but see Pan & Eriksen for a failure to find divided attentional allocation). Awh and Pashler (2000) cued two locations, separated by 2.2° visual angle. Identification of target numbers at either cued location was significantly better than for a target at the location between them. Intriligator (1997) used a task in which subjects tracked a number of spots in a field of moving spots. Subjects pressed a button when one of the spots changed color. Not surprisingly, response times were fastest when one of the attended spots changed color. The main finding though, was that there was no performance benefit for a spot located between a pair of attended spots in comparison with a spot outside this region. The studies by Awh and Pashler, and Intriligator, both provide evidence that attention need not be allocated within the outline of a single spotlight, but may be manifest as a multimodal distribution across a display.

Even if spatial attention can be divided, this does not imply that intervening items are necessarily ignored, and don't impair performance in the primary attention task.
Kramer and Hahn (1995) addressed this issue using displays consisting of 4 letter stimuli. Two target letters were presented, e.g., at 10:30 and 1:30 on an imaginary clockface (12° visual angle separation), and two distractors were placed between them, e.g., at 11:30 and 12:30 along the clock face. In this response-competition experiment, subjects made a “same” or “different” judgment about the targets. Target locations were first cued (150 ms), and critical displays (targets and distractors) were then presented for 60 ms. When all elements had an abrupt onset, distractor compatibility significantly affected performance. This finding is similar to that of Pan and Eriksen (1993), who also found effects of response competition from a distractor placed between two attended locations. Thus, even though attention can be split between noncontiguous locations, the intervening locations are not always ignored.

Apropos the current study, the findings of Kramer and Hahn (1995) lead to the prediction that subjects would not be able to ignore intervening distractors at small target separations. But, could they ignore the distractors surrounding the target pair? Intrilligator (1997) demonstrated that there is no processing advantage for distractors intervening between attended locations, as opposed to distractors around two targets. Does this imply that interference produced by intervening distractors is also no greater than that produced by distractors in other nearby locations? Kramer and Hahn (1995) might have answered this question had they included a condition in which the distractor

---

* When Kramer and Hahn (1995) used stimuli constructed by removing lines from figure eight premasks (i.e., stimuli which were already present in the display), distractor effects were eliminated. These researchers attribute this change to the abrupt onset disrupting the allocation of attention to the target locations.
letters were in the positions *surrounding* the targets, i.e., at the 9:30 and 2:30 locations…(and, in effect, this is what was done in Experiment 4 of the present study).

The purpose of the Experiment 4 was to assess whether there is a difference in the level of interference produced by intervening distractors versus surrounding distractors in dual-target identification. Because the effects of density were detected primarily at small inter-target separations in Experiments 1 and 2, a short separation was used in the experimental conditions. Stimuli were similar to those used in Experiment 1—letters arranged in a circle, etc. (see methods). In each of the two experimental conditions, targets were separated by the same distance (two slots around an imaginary circle divided into 30 slots), but distractors were arranged differently around them. In one of these conditions, the intervening space contained no letters while the targets were each flanked by letters on their other side. In the other experimental condition, two letters were placed between the targets, and blank spaces were on their outside. In both conditions, each target was flanked by a distractor and a blank space. What differed was whether the distractors or spaces were between or surrounding the target pair.

*Method*

*Participants.* Eight members of the Johns Hopkins University community participated in this experiment. Subjects were paid for taking part.

*S.imuli.* Displays and procedures were the same as Experiment 1, except as noted. Placement of letters was based on an imaginary circle divided into 30 equally spaced “slots” (as opposed to the 24 in Experiments 1 and 2). On all trials, letters were arranged in five groups of four, with two empty slots between each group (see Figure 13). This
made for 20 letters and 10 open slots. The letters at the ends of a single grouping were thus, two slots apart, as were the letters at the near ends of two adjacent groupings. As in Experiment 1, targets were two letters that were colored differently from the other letters. Luminances (red and green) were equated by flicker photometry. There were three conditions: 1) intervening-distractor condition: targets were the two letters at the end of a single four-letter grouping, 2) surrounding-distractor condition: targets were the two letters at the near ends of two adjacent groups, or 3) control condition: targets were separated by twelve slots (letters + empty spaces). Measured from center to center, adjacent letter slots were separated by 0.9° visual angle (12° polar coordinates). In the experimental conditions, the targets were separated by 2.6° visual angle (36° polar coordinates), and in the control condition the targets were separated by 8.1° visual angle (144° polar coordinates).

Although the arrangement of letter groups relative to each other never varied, the absolute position of the groups around the imaginary circle was randomized (e.g., a specific slot might be occupied by a letter on one trial, and be empty on the next). This made for six possible orientations of the basic 20-letter arrangement.

Procedure. The critical target display remained on screen for a fixed period of time: 225 ms. (This was an accuracy study, not a staircase procedure.) Otherwise, the procedure used was identical to that of Experiment 1.

Design. Each participant performed two sessions of experiments over seven days (or fewer). Only one session was done on any single day. Sessions were composed of four experimental blocks of 90 trials. Blocks were mixed, consisting of 30 trials at each
of the three conditions, presented in a quasi-randomized sequence. Target and distractor colors were fixed within a session, but were varied across sessions (e.g., a subject who viewed red targets in the first session viewed green targets in the second session). The order in which target color was assigned to a session was counterbalanced across subjects. Each session began with 42 practice trials (14 trials per condition) with display-mask SOA going from 700 ms (first 15 practice trials) to 500 ms (next 10 practice trials) to 250 ms (last 17 practice trials) in the course of the practice block. The first two blocks of the first session were also considered practice. Subjects were given the opportunity to take a break at the midway point and at the completion of each block.

Results and Discussion

Mean accuracy measures for Experiment 4 (presented in figure 14) are as follows (the independent variable was proportion correct): Surrounding-distractor condition: 0.55, [SD = .121]; intervening-distractor condition: 0.57, [SD = .122]; wide target-separation condition: 0.67, (SD = .117). A one-way ANOVA found a significant effect of condition [F(2,14) = 17.31; p < .0002].

A matched samples t-test was performed comparing the two conditions of interest (surrounding- versus intervening-distractor conditions.). The difference between these conditions was not statistically significant [t(7) = 1.17; p = .280. n.s.].

The main finding of Experiment 4 was that for distractors isoecentric with a pair of targets, distractors in-between the targets produced no more interference than distractors surrounding the targets (in fact, the effect was in the opposite direction, although this difference did not reach statistical significance). The improvement in
accuracy for the control (long inter-target separation) condition over the experimental conditions, indicates that the procedure was sensitive enough to detect differences in performance, and lends support to the claim that there is no difference between the experimental conditions.

The results of Experiment 4 suggest that subjects were no more compelled to process items in the space between targets than items around the targets. This finding is consistent with Bahecall and Kowler's (1999) data, which showed that distractors between target letters were not misreported (i.e., in place of targets) any more than were distractors outside the target pair. Put another way, the space between targets was not illuminated by the spotlight of attention any more than the space outside the inter-target area, in agreement with Intriligator's (1997) study of attentional tracking reported in the introduction to the experiment.

Under each experimental condition, targets were subject to the same degree of crowding—a blank space on one side and a letter on the other—yet performance was not impaired any more by objects occupying the space between two targets than by empty space between them. As such, one could reach a more general conclusion from the present experiment: that processing of two attended objects is affected by the extent of crowding at each location, but not by the arrangement of one target-distractor cluster in relation to the other. This hypothesis is further discussed and tested in Experiment 5.

A reexamination of the stimuli used in Experiment 4 indicates a possible alternate explanation for the present findings. In the intervening distractor condition, the targets were at the ends of same perceptual group (composed of the two targets plus two
distractors). In the surrounding distractor condition, the targets were at the ends of two separate groups (each composed of one target plus three distractors). Based on theories of attention that emphasize grouping or object-based processing (Duncan, 1984; Egly, Driver, & Rafal, 1994; Kahneman & Henik, 1981) it might be argued that it is easier to identify two targets when they organized together than when they are spread across two groups. If this is true, then interference from distractors intervening between targets could be offset by the advantage gained from grouping between the targets. The null result that was found could thereby reflect two different effects balancing each other out in the intervening distractor condition: perceptual grouping (which aids performance) and distractor interference (which hurts performance), and not the absence of a difference between the experimental conditions.

In response to this analysis, it could likewise be claimed that perceptual grouping is unavoidable when distractors are introduced between two narrowly separated targets. Put another way, grouping between targets and distractors under these display conditions is not an artifact, but an essential part of the relationship: If we are going to discuss the effect of intervening distractors on identifying closely spaced targets, then grouping is part of the package.

One way to lessen this grouping factor might be to use displays adapted from those of Banks, Larson and Prinzmetal (1979, see Introduction to dissertation). Take the case in the current experiment when both targets are part of the same group: additional distracting letters could be added adjacent to the intervening distractors, but at right angles to the target-distractor group (i.e., along a radial drawn through the distractor).
The intervening distractors might then group more strongly with the added distractors then the target letters, and their effect could be gauged apart from the influence of target-distractor grouping.

Another solution is to place the target letters farther apart. With the targets adequately spread out, a gap can be created between the intervening distractors. This creates two separate groups in both experimental conditions—which addresses the issue described above. In addition, as demonstrated in Experiment 3, the interference by distracting letters is significantly increased when distractors are freed from grouping with similar stimuli. Such were the conditions of Experiment 5.
Figure 13. Experiment 4 stimuli. (Bold letters represent target letters, which were a different color from distractors.)

Figure 14. Experiment 4 data. (Error bars denote standard errors.)
Experiment 5

In Experiment 4 it was found that distractors between two targets produced no more interference than distractors extending beyond the target pair. In Experiment 5 targets were more widely separated (6.0° visual angle) to also test whether local crowding effects vary with relative placement of targets. In Experiment 5, the two experimental conditions were as follows. In one condition, distracting letters were on the imaginary line connecting, and extending beyond the targets: in this, the in-line condition, distractors were both between and exterior to the target letter pair (though always on the imaginary connecting line). In the other condition, the parallel condition, distractors were aligned perpendicularly to the line connecting the targets (as the distractors were adjacent to the targets, this formed two parallel target-distractor groups; see Figure 15). In both conditions, however, the arrangement of each target and its nearby distractors in relation to fixation was the same. Any difference between the conditions could then be attributed to attentional and not retinal factors.\(^7\)

The two experimental conditions employed in Experiment 5 were meant to be analogous to the conditions of greatest and least masking as found by Chambers and Wolford (1983) and Toet and Levi (1992). Recall that subjects in these studies viewed displays consisting of a target accompanied by one (Chambers and Wolford) or two (Toet

---

\(^7\) A control condition in which no distractors were presented with the targets was run concurrently with this procedure. This was done to assure the sensitivity of the present procedure should no difference be found between the other two experimental conditions: the assumption was that performance should be better for the no distractor condition than for either of the other conditions (e.g., see Experiment 3), regardless of distractor arrangement.
and Levi) distracting letters (see Chapter 1 of dissertation). Both studies found that interference was strongest when the masking letter was on the (imaginary) radial line running from central fixation through the target (a letter on the peripheral side of the target produced more interference than one on the interior side).

In both the Chambers and Wolford (1983) and Toet and Levi (1992) studies, positional effects were assessed in relation to central fixation, i.e., positions were described either as “isoeccentric”, or, as being on a radial line passing through fixation. In contrast, in the present experiment, positions may be described in relation to a reference axis connecting the targets. For example, in the in-line condition, distractors are along the imaginary line connecting the targets (analogous to the radial line of the cited studies), and in the parallel condition, each set of distractors and their accompanying target are equidistant (or isoeccentric) from the other target.

As in Experiment 4, the factors being investigated concerned the spatial allocation of attention across two locations, and how this allocation is effected by nearby distractors. However, in this experiment, a comparison was made between the influence of distractors along the imaginary line connecting the targets and those distractors off of this line. If two separate spotlights of attention illuminate the objects, or, if attention shifts in discrete/quantum jumps to the targets (Sperling & Weichselgartner, 1995), distractor placement should not affect target identification performance. However, if attentional processing is oriented along an axis connecting the targets, then distractors along this line should interfere with performance more than distractors off of this imaginary line.
The display configuration of Experiment 5 alleviates one particular problem from Experiment 4. In the previous experiment, the distractor conditions were conflated with the state of perceptual grouping between the targets: in the intervening-distractor condition, the targets were members of the same 4-letter grouping, and in the surrounding-distractor condition, the targets were members of different 4-letter groupings. In Experiment 5, whether there were distractors or empty space between them, each target was always part of a different 3-letter cluster. (Note, however, that stimuli in the in-line condition did form a collinear set of units. This aspect of grouping between the stimuli groups is addressed in the discussion.)

Method

Participants. Four Johns Hopkins University students participated in this experiment. They received class credit for taking part.

Stimuli. In the experimental conditions, displays consisted of two groups of three letters (see Figure 15). Each three-letter group (or “triplet”) was arranged in a single row. The central letter of each triplet (the target letter) was always a different color from its two neighboring flankers. Both targets were same color (red or green) and the distractors were the other color.

One target letter was at the center of each triplet. Each triplet could appear with its target letter at any of the twelve clock face positions around the display; the target was always 4.25° visual angle from fixation. The target letters in the two displayed triplets were always separated by 45° (polar coordinates; 6.02° visual angle) across the display. The flanking letters were placed such that a line drawn through the centers of each of the
three letters in a triplet would intersect a radius drawn through the central letter at either 45° or -45° (There were two possible triplet arrangements at each target location). The center to center distance between adjacent letters was 1.1°. The inner flanking letter in each group was 3.60° visual angle from fixation and the outer flanking letter was 5.12° visual angle from fixation.

The two experimental conditions were defined by the relationship between the two triplets on the display: in the in-line condition, all the letters in the two groups were arranged in single file across the display (see Figure 15, right panel). In the parallel condition, the groups formed two parallel, non-intersecting alignments: no letters were present between the two targets, and a line drawn connecting the central target letters intersected a line drawn across either triplet at 90° (see Figure 15, left panel).

It should be emphasized that the only factor that was varied across conditions was the relationship between the alignment of the two clusters of letters; the exact same (individual) triplets were used in both conditions, however, the combination of triplets varied. That is, for any single triplet configuration, there was one complementary triplet with which it could form the in-line condition (e.g., at 45° polar coordinates around the display), and another with which it could form the parallel condition (e.g., at -45° polar coordinates around the display). An example of this is seen in Figure 15, in which the identical N-A-H triplet appears in each condition.

Procedure. The sequence of events within a trial was identical to that of Experiment 1: the PEST algorithm was used to adjust the display-mask SOA.
Design. Participants performed two sessions per day on five different days (ten sessions total, including control condition, over the course of 14 days (or fewer). Each day, one session consisted entirely of experimental condition trials (in-line and parallel condition) and the other session consisted entirely of control condition trials (2 targets/no distractors). The assignment of target color was counterbalanced across sessions (within subjects) and was counterbalanced between subjects.

Each session consisted of two PEST runs, a run being a sequence of staircase-adjusted trials for a particular condition. During a session, one run was performed for each of the two experimental conditions (in-line condition and parallel condition). Trials in each condition were interleaved with each other, in a semi-random order. There were 96 trials in each condition, or 192 trials in a session. Each session began with 40 practice trials (20 per condition) with display-mask SOA going from 700 ms (first 15 practice trials) to 500 ms (next 10 practice trials) to 300 ms (last 15 practice trials) in the course of the practice block. PEST adjustment of SOAs did not start until the practice trials were finished. Subjects were given breaks after every 48 trials.

Results

Experiment 5 employed an adaptive staircase procedure, with SOA’s adjusted to produce 80% accuracy in identifying both targets. (Accuracy data for relevant trials is presented in Table 2). Mean SOAs for each subject in both target-distractor alignments are presented in Figure 16. Higher SOA suggests worse performance than a lower SOA (a longer exposure was required for equal accuracy). For all four subjects mean SOAs were greater for the in-line condition than for the parallel condition. The grand mean for
the in-line condition was 307.91 ms (SD = 72.98) and grand mean for the parallel condition was 280.50 ms (SD = 55.76). A paired-sample t-test showed that the mean SOA for the in-line condition was significantly greater than that for the parallel condition \[t(3) = 6.79, p < .01\]^8.

In a post-hoc analysis, the difference between the mean SOAs (in-line versus parallel conditions) was examined across the experimental sessions. The differences (in ms), by session, were as follows (a positive figure indicates that the mean for the in-line condition was greater than that for the parallel condition): +85, -14, +33, +22, +14. A 2 (alignment condition) by 5 (session) ANOVA showed a significant effect of session \[F(4,12) = 18.8, p < .001\] and alignment condition \[F(1,3) = .011, p = .011\]. A significant trend was also found in the interaction between these two variables \[F(4, 12) = 2.9, p = .067\], reflecting the pattern of differences shown above.

<table>
<thead>
<tr>
<th>Parallel Condition</th>
<th>In-line Condition</th>
</tr>
</thead>
<tbody>
<tr>
<td>80.1</td>
<td>78.6</td>
</tr>
</tbody>
</table>

Table 2. Experiment 5 accuracy, by experiment condition. Combined accuracy (per cent correct) for trials 48-96 of PEST runs.

**Discussion**

The main finding of Experiment 5 was that distracting letters along the imaginary line connecting two target letters produced poorer target identification performance than distractors arranged perpendicularly to this line. This result occurred in spite of the individual target-distractor clusters having the same retinal orientation in each

^8 In light of this statistically significant finding, results for the control condition are not reported.
experimental condition. Mean data from four subjects indicate that the effect was notably strongest in the first experimental session, was not detected in the second session, and was present, but decreased steadily across the final three experimental sessions. This interaction of distractor effect with session might well reflect a ceiling effect of subjects getting better at the task in general, and learning to ignore distracting letters.

As in Experiment 4, grouping factors in the present displays were confounded with the distractor conditions, however, this confound can not explain the present results. In the in-line condition, the Gestalt property of good continuation leads to the two clusters forming a collinear, albeit, disjointed entity. Grouping between these two groups would seem to be stronger than between the two clusters in the parallel condition. As noted in the Experiment 4 discussion, theorists who assert that attention is based on objects or perceptual grouping (Duncan, 1984; Egly, Driver, & Rafal, 1994; Kahneman & Henik, 1981) would predict an advantage for the in-line condition, as it should be easier to extract two letters from one grouped entity, than one letter from each of two. This pattern did not hold under the present conditions. An explanation is offered below.

So, how might the Experiment 5 stimuli have produced the results found here? I would hypothesize that the spatial relationship between the targets serves to guide attentional allocation to the display. Specifically, when the display appeared, attention became oriented along the axis (that is, the imaginary line) connecting the two targets. Processing favored the target locations, but because the displays were present for a very
short time, it was difficult to fully exclude or inhibit the distractor items along this axis. And, as these distractors were not fully grouped or masked (due to the space between clusters, as in Experiment 3) they would tend to produce interference in target identification. Thus, when distracting letters are off this target-target axis, they are less deeply processed, and cause less interference than distractors that are along this axis.

This description of attentional allocation is similar to one posed by Pan and Eriksen (1993) based their study which employed a response competition paradigm. In their Experiment 1, subjects viewed displays composed of three letters. One target letter was presented directly above and another was presented directly below fixation; a third letter, the distractor, was presented along the horizontal meridian. The distance between the target letters varied between .25 and 1.0° (the targets moved in symmetry, in that the distance of each from fixation was always the same). The distance of the distractor from fixation also varied from .25 to 1.0 degrees (the distractor letter was moved horizontally in line with fixation). To summarize the findings, 1) performance was best when the targets were farthest apart, and 2) the nearer the distractor was to fixation the worse was performance. Note that an attempt was made to control for effects of retinal eccentricity (in their Experiment 2) by presenting distractors in advance of the targets, so distractors could be clearly seen. The authors concluded that attention conforms to an area elliptical in shape, with its major axis defined by the target stimuli, and its minor axis varying in proportion to the major axis length.
Experiment 5 extends the findings of Pan and Eriksen (1993) to a situation in which the attentional axis is well off of fixation. (A line drawn through both targets would be about $3^\circ$ visual angle from fixation at its closest, based on targets $4.26^\circ$ visual angle from fixation). In addition, in the present study, distractor interference was not confounded with a) distractor distance from the targets and b) distractor distance from fixation, as was the case in the Pan and Eriksen study. What varies in the present experiment is simply the distance of the distractors from the major attentional axis: retinal eccentricity of distractors is otherwise balanced across conditions.

Finally, it may appear that the results of Experiment 4 were not consistent with Experiment 5 (in Experiment 4 no effect of distractor placement was found). However, the two results can be reconciled if we assume that the allocation of attention extends just beyond each target letter along the attentional axis. (This is not an unreasonable assumption, especially for a brief display). Thus, applying this to Experiment 4, both intervening and surrounding distractors were located along the attentional axis: that is, they were along the line connecting, and extending beyond the targets. As such, both sets of distractors would be predicted to cause similar levels of interference.
Figure 15. Experiment 5 stimuli. (Bold letters represent target letters which were a different color from distractors.) Note that the identical N-A-H triplet appears in each condition.

Figure 16. Experiment 5 data, by subject. Note that y-axis is flipped so relatively superior performance is toward the top of the graph. (Error bars denote standard errors.)
General Discussion

The experiments comprising this dissertation explored the role of spatial separation and irrelevant objects in attentional tasks. In all experiments, attention was allocated to a pair of locations occupied by letter targets, while the distance between the locations and the arrangement of distractors around them were varied. The purpose of Experiment 1 was to ascertain the basic effect of distance in dual-target identification. The main finding of this experiment was that as distance between the targets increased, performance also improved (as measured by target-display SOA required for 80% accuracy). This finding concurred with that of other recent studies that had found a pattern of suppressed processing near an attended location (Bahcall & Kowler, 1999; Caputo & Guerra, 1998; Cave & Zimmerman, 1997; Mounts, 2000a, 2000b; Turatto & Gandalfi, 1999). Conversely, other studies had reported no effect of distance in a similar task (Kwak, Dagenbach, & Egeth, 1991; Sagi & Julesz, 1985). A reexamination of the data from those studies suggests that attentional interference of the kind demonstrated in Experiment 1 may also have been present, but was attributed to lateral masking. Note that the presence of the positive effect of separation in the 24-letter displays in Experiment 1 indicates that this is not simply lateral masking, but a higher order, or attentional, effect.

Studies in the area of neurophysiology have similarly demonstrated suppression of neural responding during attentional tasks. Such work may provide clues as to the mechanism that underlies the behavioral findings thus far described. In a now classic
study, Moran and Desimone (1985) investigated how attention modulated single cell responses in Macaque visual cortex (these findings apply to areas V4 and IT). A preferred and a nonpreferred stimulus were placed in a cell’s receptive field (a preferred stimulus is a stimulus that evokes the highest response by the cell, in comparison to other tested stimuli). When the monkey attended to the location of the preferred stimulus, responding was high. However, when attention was directed to the nonpreferred stimulus, responding was significantly reduced, even while the preferred stimulus was still present in the receptive field. Moran and Desimone describe the receptive field as having “shrunk” around the attended stimulus, to the detriment of the response to the unattended stimulus. More recently, Reynolds, Chelazzi and Desimone (1999) performed a similar study of Macaques, finding, first of all, that when two objects were placed in the receptive field of a cell from either area V2 or V4, the resulting response was intermediate between its response to either object alone. However, when one of the objects was attended, the cell’s response moved in the direction of its response to that object alone: i.e., faster for a preferred stimulus, and slower to an nonpreferred stimulus. Both of these studies provide examples in which a cell’s response to its preferred stimulus can be extinguished when that stimulus is in the vicinity of a(n) (nonpreferred) attended stimulus. In the cited studies, it was also found that a cell’s response to the preferred stimulus was often unaffected when attention was directed to a location outside of the receptive field (see also Luck Chelazzi, Hilyard & Desimone, 1997).
Desimone and Duncan (1995) propose that suppressive functions such as those described arise to enable the organism to allocate limited resources in favor of a relevant (i.e., attended) stimulus. As receptive fields of cells downstream in visual cortex (especially V4, TE and TEO) can subtend many of degrees of visual angle, they are likely to receive inputs from lower level visual areas regarding multiple stimuli. To manage this mass of input, some stimuli are given preferential processing in lieu of others, whose inputs are suppressed. As such, attention becomes an emergent property, describing the competitive process in which a stimulus emerges (i.e., to full recognition) as a result of the biasing of lower level inputs in its favor. A similar process is described by Luck, Girelli, McDermott and Ford (1997). However, they describe the function in terms of resolving ambiguity in perceptual coding. When a large receptive field is occupied by multiple objects, attention allows the properties of a relevant item to be bound together into a coherent whole, while suppressing the inputs associated with properties of other nearby objects.

In experiment 1, two-letter displays yielded a monotonic function of improved performance with increasing separation, whereas for the 24-letter displays there was a significant dip in performance at the second closest separation. The cause of this disparity was investigated in Experiment 2. Two hypotheses were pitted against each other: perceptual interference versus response interference. In experiment 2, numeral
distractors did not induce the dip in the performance function that was produced by letter
distractors, supporting the response interference hypothesis: the dip resulted from the
conceptual similarity between the targets and distractors, and was not caused by the mere
physical presence of distractors.

The findings of the first two experiments were discussed in the framework of
Bahacall and Kowler's (1999) proposed models of attentional interference. The results
for the 2-letter condition (Experiment 1) support the presence of an inhibitory region (i.e.,
a difference-of-Gaussians) surrounding an attended location, regardless of display
density. In addition, the results of Experiment 2 point to "imprecise targeting" as
contributing interference when targets are separated by small distances.

The results of Experiments 1 and 2 might also be explained by the presence of an
inhibitory mechanism operating at a conceptual level of processing. Precedence for this
type of mechanism is found in cited work by Santee and Egeth (1981) and Walley and
Weiden (1973). In addition, studies by Egeth, Jonides, and Wall (1972), and Jonides and
Gleitman (1972,) have shown that an efficient visual search can result when there is a
categorical difference between a target and accompanying distractors (but also see
Krueger, 1984, who claims these findings are the result of differences between letter and
numeral categories at the physical level). In the Egeth, Jonides and Wall (1972,
Experiment 3) study, flat search slopes were obtained (that is, response time was
independent of the number of distractors) in searches for a letter among numerals, or a
numeral among letters. Thus, there may be visual processing channels capable of
detecting and suppressing stimuli based on abstract visual information. This possibility
certainly warrants further investigation as it relates to inhibitory-surround-type models of
attention.

In Experiment 3 a comparison was made between a display crowded with
distractors (24-letter displays) versus one in which there were only a few distractors, all in
the vicinity of the targets (6-letter displays). At both tested inter-target distances,
performance was better in the crowded (24-letter) displays. This experiment helped
explain the comparable performance in Experiment 1 between dense and sparse displays:
Target identification in the 24-letter displays was apparently aided by grouping (Duncan
& Humphreys, 1989) or crowding (Taylor & Brown, 1972) among distractors, combined
with the high salience of the targets against the uniform field of distractors. An
alternative explanation is that distractor effects were limited by the high perceptual load
(Lavie & Tsal, 1994) imposed by the filled display. This explanation would assert that
perceptual capacity was consumed by processing the large number of distractors, limiting
interference from any one of them. In the Experiment 3 displays, load would seem to
have been conflated with distractor numerosity, so it is difficult to tease apart the effects
of perceptual load and grouping. However, load is a difficult construct to define: the
many distractors may have comprised a large perceptual load. Or, the distractors may
have been easily encoded due to their similarity with each other and dissimilarity with
target properties—easy encoding being a hallmark of a low load. The ambiguity inherent in the perceptual load explanation puts it beyond the scope of the present study to explore more fully, although it is something to consider.

The final two experiments looked at how interference associated with distractors is influenced by the spatial relationship between the attended objects; that is, are crowding effects local to a target-distractor group, or are these effects global, and thereby affected by the arrangement of the groups in relation to each other?

Experiment 4 first examined the effect of distractors when targets are at small separations; specifically, the question under consideration was whether items interior to a pair of attended objects produced greater interference than items exterior to the pair? Using displays adapted from the first three experiments (all stimuli were isoecentric to fixation), no difference was found between the distractor conditions (i.e., intervening vs. surrounding distractors).

In experiment 5, targets were separated by a greater distance (6° visual angle) than in Experiment 4, and the question was whether attention is allocated along the axis connecting and extending beyond a pair of attended locations. Using display conditions designed to be analogous to those of Chambers and Wolford (1983), performance was found to be worse when distractors were located along the attentional axis than when distractors were orthogonal to this axis. In each condition, the relationship between each grouping of target and distractors to fixation was the same, so the finding can be
attributed to attentional and not retinal factors.

It is hypothesized that attention is oriented along the axis connecting targeted (attended) locations and extends beyond these locations: an explanation that can account for the findings of both Experiments 4 and 5. This finding is in agreement with theories positing the spotlight metaphor of attention, coming closest to the description put forth by Eriksen and Pan (1993) who claimed that attention assumed an elliptical form around a pair of separated targets (see also Kramer & Hahn, 1995). Note, however, that Pan and Eriksen did not test distractors placed external to the target axis locations (as was done in the present study). In future work I plan on using a variant of the Experiment 5 displays to separately test the effect of distractors interior, or exterior, to the target pair (this would entail using two variations of the in-line experimental condition). Such an experiment would provide evidence as to whether the elliptical allocation of attention extends beyond the targeted locations when the targets are more widely separated than in Experiment 1.

Another relevant follow-up experiment would be to fill the space between the targets with a complete uninterrupted array of distractors. It may be that under these circumstances, the distractors form a uniform backdrop for attentional allocation, and interference is reduced due to distractor grouping and increased target salience (in agreement with the findings of Experiment 3).

Support for the finding of Experiment 5 comes from a neurophysiological study by Heinze, Luck, Münße, Göss, Mangun, and Hillyard (1994). Event related potentials
were recorded in this study of attentional allocation. Subjects attended to two out of four possible stimulus locations, all arranged in a horizontal row just above fixation (locations were set about 6° and 12° to either side of fixation). The task was to make a same or different judgment regarding two symbols (the symbols were made up for the experiment) at the attended locations. Sets of four symbols, one at each location, were flashed on screen for brief periods. Intermittently, a probe—a white vertical bar—appeared at one of the stimulus locations. The event related response (specifically, the P1 response), to the probe was the dependent measure. When subjects attended to adjacent locations, the response was highest when the probe appeared at an attended location, and was attenuated when the probe appeared in an adjacent location. However, when subjects attended to alternate locations (for example, the first and third location, going across) recorded responses were highest when the probe was at the attended locations, and were just as high when the probe appeared at the intervening location. Behavioral measures also showed that performance was worse for the split attention condition than for adjacent target symbols. The behavioral and physiological evidence indicate that attention cannot be split between two locations without attending to the area in between.

Finally, it should also be noted that in the current experiments, displays were brief and subjects were not aware of target locations until the targets appeared. In contrast, Sperling and Weichselgartner (1995, see Chapter 2 of dissertation) used displays in which the two target “locations” (RSVP displays) were on screen for seconds before attention
was shifted from one stream to the other. In the Sperling and Weichselgartner study there was no effect of an intervening distractor stream. And, Kramer and Hahn (1995) found that distractors between two targets interfered with target identification when all stimuli had abrupt onsets, whereas, this response competition was eliminated when the stimuli were produced by removing lines from figure eights in a pre-stimulus display. The results of these two studies, in combination with the current Experiment 5, suggest that when attention can be allocated to relevant locations in advance, distractor interference can be overcome. However, at first quick glance, attention is ineluctibly guided by the underlying terrain.

Based on the results of the present study, no claim is being made as to the analog or discrete movement of attention. However, what is clear is that as the distance between attended objects increases, one’s ability to identify the attended objects improves. In addition, the interference caused by irrelevant items near a pair of attended locations is influenced both by the items’ category, and by the arrangement and resultant grouping among the stimuli in a scene. Finally, a pair of attended locations forms the axis along which attention is allocated.
References


_Bulletin of the Psychonomic Society, 21_(6), 459-461.


_Annual Review of Neuroscience, 18_, 193-222.


_Journal of Experimental Psychology: General, 4_, 501-517.


_Perception & Psychophysics, 38_(3), 296-298.


Appendix: The PEST Procedure

The goal of psychophysics is to relate a stimulus in the environment to the sensory experience of that stimulus. This is often done by calculating an absolute threshold, which is a measure of the minimum amount of stimulus energy necessary for the observer to detect the stimulus. Fechner (1860, cited by Goldstein, 1996) introduced three methods for measuring the absolute threshold in his Elements of Psychophysics, and these methods have remained staples of the field ever since. In all three methods, observers are presented with a sequence of sample stimuli, and indicate for each whether or not they can detect it. Two of these methods require that the sample stimuli be presented at fixed magnitudes and in a fixed order; the order may be ascending or descending for the method of limits, and the order is random for the method of constant stimuli. As its name implies, however, the third method, the method of adjustment, involves the observer or experimenter adjusting the stimulus energy until the stimulus is barely detectable by the observer.

Modern adaptive methods of psychophysics are designed to gradually home in on the magnitude of stimulus energy necessary to achieve a specified performance criteria, e.g., 80% accuracy on some measure. These methods work by adjusting the test stimuli based on some portion of the history of responses. By raising and lowering the stimulus energy in a prescribed way (hence, these are also called staircase procedures), the experiment eventually converges on a target value. As with the method of adjustment, stimulus presentation order is not preset, and thus these methods are more efficient than the fixed methods which expend trials retesting stimulus values which may otherwise
have been rejected earlier by adaptive methods. Note however, that the fixed methods allow for the derivation of psychometric functions showing the relative sensitivity of an observer to a wide range of stimuli.

Experiments 1 and 5 of the dissertation use the adaptive method of Parameter Estimation by Sequential Testing (or PEST; Taylor & Creelman, 1967), as the procedure for psychophysical testing. PEST was initially designed for use in psychoacoustical studies, and has been shown both to be highly efficient and to produce unbiased results (i.e., not to over- or underestimate the target variable; Taylor, Forbes & Creelman, 1983).

The PEST Procedure

Taylor and Creelman (1967) assert that any adaptive procedure must describe the following: 1) when to change levels of the stimulus, 2) what level of the stimulus to try next, 3) when to end the run of trials, and 4) how to determine the value of the target variable. The PEST rules for these four tasks will now be described in sequence.

As a matter of definition, a PEST “run” is a string of trials for a particular experimental condition, which follow the PEST procedure for adjustment of the stimulus of interest. (For example, in Experiment 1, a separate PEST run was performed for each distance condition within each display numerosity.) A PEST run begins with the presentation of the test stimulus at some specified level. This level may be determined in pilot studies, and is usually near the expected target level. After each trial, the decision whether or not to change the stimulus level is made based on results of the sequential likelihood ratio test (Wald, 1947). This procedure establishes the limits within which
performance must be sustained in order to maintain the current level of the stimulus for the next trial. If performance exceeds the upper limit, or is below the lower limit, the level is adjusted according to rules described later.

The sequential likelihood ratio test (Wald, 1947) works as follows: At each stimulus level, performance is assessed by tracking the total number of trials that have been run at that level, and the number of trials for which a correct response was recorded at that level. For any desired level of performance (e.g., 80% accuracy), there is also a corresponding expected value for the number of correct trials given the total number of trials run to that point (again, at that stimulus level). This may be expressed as follows:

\[
(1) \quad E[N(C)] = T \times P_t
\]

where \(N(C)\) equals the number of correct responses at the stimulus level, \(T\) equals the number of trials at that stimulus level, and \(P_t\) equals the target level of performance (Taylor and Creelman refer to this as “target probability”).

A constant, known as the deviation limit, \(W\), is then both added to, and subtracted from, the expected value to establish the test boundaries. That is,

\[
(2) \quad N_b(C) = E[N(C)] \pm W
\]

in which \(N_b(C)\) equals the upper and lower performance boundaries. If actual performance is outside the calculated boundary, then the stimulus level is adjusted. (The equation... above are both from Taylor and Creelman, 1967).

This procedure is better understood by examining a hypothetical PEST run. Table X depicts a run in an experiment in which an observer is required to make some judgment on each trial (such as, “Is a stimulus present?”). For this string of trials, the target
performance is 0.80 (i.e., 80% accuracy), and W is set to 1.0, the number suggested by Taylor and Creelman. Substituting these values into equations 1 and 2, the upper and lower bounds for performance after one trial are -0.2 and 1.8: Expected value for number correct is $1 \times 0.8 = 0.8$, so the boundaries are $0.8 \pm 1.0$. All possible levels of subject performance are also shown in the figure; in a single trial, the observer may either have been correct (1 correct) or not (0 correct). Since both 0 and 1 are within the boundaries established, no change is made in the stimulus level.

For the second trial, the new boundaries are 0.6 and 2.6 (for each trial, 0.8 is added to each boundary). The possible levels of performance after two trials are 0, 1, or 2 correct. (These and subsequent such totals for a trial number N, are derived by adding 0 or 1 to each possible number correct for trial N-1.) Of these numbers, 0 is outside (below) the performance limit, and results in a change (increase) in stimulus level. When the stimulus level is changed, the tally of total trials and number correct begins anew. Otherwise, if the observer had gotten 1 or 2 correct in the first two trials, a third trial would be run at the old stimulus level. (Note that the only possible number of correct responses after three trials are 1, 2, or 3, because an observer with 0 correct would have started onto a new level after just two trials.)

If the decision is made to change the stimulus level, the direction of the change is determined by which boundary is breached: If the lower boundary is crossed, the stimulus magnitude is increased in order to improve performance. If the upper boundary is exceeded, the stimulus magnitude is decreased in order to degrade performance (this can not occur until the sixth trial in the example shown in the table).
Each change in the stimulus magnitude is referred to as a step, and the size of each step as well as the direction (positive or negative) of the step can vary. The size of each step is adjusted according to the following rules devised by Taylor and Creelman (1967, pp. 783-4). (These rules are based on efficiency and bias calculations from computer simulations performed by those authors).

1. On every reversal of step direction, halve the step size. (A reversal occurs if the previous adjustment was in the direction opposite that of the current step).

2. The second step in a given direction, if called for, is the same size as the first. (The step numbers refer to consecutive steps in a given direction, i.e., positive or negative. When the direction of step movement changes, the counter is reset to 1.)

3. The fourth and subsequent steps in a given direction are each double their predecessor (Taylor & Creelman (1967) note that an upper limit on permissible step size may be needed).

4. Whether a third successive step in a given direction is the same as or double the second depends on the sequence of steps leading to the most recent reversal. If the steps immediately preceding that reversal resulted from a doubling, then the third step is not doubled, while if the step leading to the most recent reversal was not the result of doubling, then this third step is double the second.

Ending the PEST run: The PEST run continues, following the rules described for decisionmaking (whether to adjust the stimulus) and step size adjustment, until the step
size reaches some predetermined minimum size. When this minimum value is reached, “convergence” is said to have occurred. This final, or convergence, value is the target (i.e., threshold) for the run.

Multiple PEST procedures can also be performed concurrently to arrive at independent thresholds for a number of experimental conditions within a single experimental session. This can be done by randomly interleaving trials from each condition, and tracking each condition separately. Whereas a single run ends with its convergence to the minimum step size, with multiple simultaneous runs, one run may converge before the others. At this point, if the trials were no longer run in the converged condition, the order of trials by condition would no longer be randomized: that is, it would no longer be equally likely for any condition to occur on a given trial (given that the converged condition was no longer being run). In their report on “Multiple Pest”, Creelman and Kaplan (1973) advise continuing to run trials in any of the converged conditions, and using the minimum value step size when anything smaller is called for by the PEST procedure. Thus convergence is no longer the criteria for completion of the run, but a preset fixed number of trials. This was the procedure followed in Experiments 1 and 5.

By continuing to run trials after a convergence has occurred, a PEST run for a particular condition may converge multiple times. The target measure can then be calculated by taking the mean of the multiple convergences. As an alternate method of

---

9 In the present study, the minimum step size was determined by the raster display refresh rate. This is discussed later in Appendix.
finding the target measure, Creelman and Kaplan (1973) suggest sampling the stimulus value at fixed intervals along the run (i.e., every nth trial). In computer simulations, this fixed-sample procedure was demonstrated to be even more precise in finding the threshold value than sampling only the conversion measures (their Figure 3). As to the frequency at which values should be sampled, the authors found that no additional information was gained by sampling more than once every 16 trials. This figure was used in the current experiments: in PEST runs of 96 trials/condition, stimulus values (adjusted SOAs) were sampled at trials 48, 64, 80, and 96.

PEST in the Current Experiments

Except as noted, the PEST procedure was implemented exactly as described above.

As noted earlier, PEST was originally formulated for use in psychoacoustical experiments. The procedure was therefore adjusted for the present experiments to make PEST more compatible with the characteristics of computer-presented displays. Specifically, an issue arises from the raster scan of the video monitor (see Matin & Boff, 1990, for a related discussion on PEST and raster scan monitors). The computers on which the present experiments were performed had refresh rates of 60hz, which translates into each video frame being displayed for approximately 16.66 ms. A frame can not be displayed for less time than this. Thus, step changes were manipulated in multiples of 17 ms. so that only whole numbers of frames were ever called for.
The rules for halving and doubling of step size were also adjusted. In pilot work it was found that the procedure was more efficient if the large step sizes resulting from doubling could be avoided. In addition, dividing multiples of 17 often would have led to steps being measured in the fractions of milliseconds—which was not permitted by the computer software. The resulting rules of doubling and halving are as follows:

- **Adjusted Doubling Rule**: When the PEST rule called for doubling the step size, the step size was increased by 34 ms if the previous step size was 68 or greater; step size was increased by 17 ms if the previous step size was less than 68.

The maximum step size was set at 85 ms, to prevent extreme fluctuations in timing.

- **Adjusted Halving Rule**: When the PEST rule called for halving the step size, the step size was decreased by 34 ms if the previous step size was 68 or greater; step size was decreased by 17 ms if the previous step size was less than 68.

**Specifications:**

The following settings were used in Experiment 1 and 5:

- targeted performance level: 80% accuracy
- deviation limit (or Wald constant, W): 1.0.
- minimum step size: 17 ms
- maximum step size: 85 ms
- initial stimulus presentation time (SOA): 255 ms
- stimulus values were sampled at fixed-intervals of 16 trials: 48, 64, 80, 96 (out of 96 trials total)
<table>
<thead>
<tr>
<th>Trial no.</th>
<th>Lower Bound</th>
<th>Upper Bound</th>
<th>Possible Total Number Correct For the Given Trial No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-0.2</td>
<td>1.8</td>
<td>0 1</td>
</tr>
<tr>
<td>2</td>
<td>0.6</td>
<td>2.6</td>
<td>0 1 2</td>
</tr>
<tr>
<td>3</td>
<td>1.4</td>
<td>3.4</td>
<td>1 2 3</td>
</tr>
<tr>
<td>4</td>
<td>2.2</td>
<td>4.2</td>
<td>2 3 4</td>
</tr>
<tr>
<td>5</td>
<td>3.0</td>
<td>5.0</td>
<td>3 4 5</td>
</tr>
<tr>
<td>6</td>
<td>3.8</td>
<td>5.8</td>
<td>3 4 5 6</td>
</tr>
<tr>
<td>7</td>
<td>4.6</td>
<td>6.6</td>
<td>4 5 6</td>
</tr>
</tbody>
</table>

Figure 17. Example of a PEST run. All possible events in a seven trial run at any one stimulus level. The bold+underlined numbers indicate cases in which the stimulus level would be adjusted for the following trial. These are the numbers outside the bounds set in columns 2 and 3.
VITA

Larence Becker was born August 22, 1959 in Queens N.Y. He grew up in Brooklyn, New York, and attended Lafayette High School (as did Sandy Koufax) where he won awards for mathematics and chemistry upon graduation. He next attended the University of Pennsylvania. After a successful first year (GPA: 3.78, Chemical Engineering major). existential angst or malaise or plain youthful folly set in (choose one)...and he did not graduate. Desperate for anything, Larence began what was supposed to be a temporary job with the Morgan Guaranty Trust Company in 1980, but, perspicacious management there recognized his talents, and he was brought on full time one year later. Larence moved to the Computer Security area at Morgan and eventually reached the level of Associate.

But, something was missing. That darn college degree. And, in 1992, Larence returned to college, entering the New York University School of Continuing Education to study psychology. For the next three and a half years, Larence worked full time at Morgan while taking eight credits per semester. (It makes me exhausted just writing it.) In 1995 Larence graduated summa cum laude from NYU, and, received the Walter Goebels award for the highest GPA in his class (4.0).

Larence happily left the corporate world to pursue the academic life upon graduation. He has been at Johns Hopkins since 1995, where he has studied attention and vision with Howard Egeth. He has no regrets whatsoever.
VITA, cont’d

At the 2000 Psychology Department Holiday Party, Dr. Egeth was roasted by the department for years of service as Chair. At this gala, Larence performed an original song (the lyrics were original, at least) honoring Dr. Egeth. It was the largest audience before which he had ever played, and it was one of the highlights of his life.