Processing Capacity of Visual Perception and Memory Encoding

Jennifer E. McLean

A dissertation submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington

1999

Program Authorized to Offer Degree: Psychology

UMI Number: 9936444

UMI
300 North Zeeb Road
Ann Arbor, MI 48103

University of Washington
Graduate School

This is to certify that I have examined this copy of a doctoral dissertation by

Jennifer E. McLean

and have found that it is complete and satisfactory in all respects,
and that any and all revisions required by the final
examining committee have been made.

Chair of Supervisory Committee:

Geoffrey R. Loftus

Reading Committee:

Geoffrey R. Loftus

John M. Miyamoto

Davida Y. Teller

Date: May 14, 1999

University of Washington

Abstract

Processing Capacity of
Visual Perception and Memory Encoding

Jennifer E. McLean

Chairperson of the Supervisory Committee
Professor Geoffrey R. Loftus
Department of Psychology

A common conceptualization is that the visual system is a limited-capacity information-processing system. This notion of limited capacity explains why performance for any one task or item typically declines with increases in the number of other tasks or items to which attention must be paid simultaneously. The idea is that a limited amount of attention, or more specifically, processing capacity, must be divided between the tasks and items. This effect of divided attention has been attributed to capacity limits in perception, or memory encoding, or both, or neither. In this study, capacity of perception and memory encoding was measured and compared using closely matched, yes-no response, search and memory tasks in which stimuli set size was manipulated. The search and memory tasks differed only in that the target was presented before the stimuli set for search versus after the stimuli set for memory. Stimuli were small ellipses varying in contrast, orientation, and size. Set sizes were 1 and 4. Performance was always worse for the larger set size, and this set-size effect was larger for memory than for search. The magnitude of the set-size effect was interpreted in terms of a theoretical measure of capacity limits that varied on a scale in which zero corresponds to unlimited capacity and one to fixed capacity. By this measure, perception had only slightly limited capacity while memory encoding was more sharply limited. The limit of the joint processes of perception and memory encoding was consistent with fixed capacity.

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGMENTS

The author wishes to express sincere appreciation to Professor Geoffrey Loftus and Dr. John Palmer for their assistance in the preparation of this manuscript.

INTRODUCTION

It is a common experience that when we try to do many things at once, our performance on any one of those things is typically worse than when doing it alone. For example, if a person is talking on a cellular phone and fixing his hair in the rear-view mirror while driving, he may be more likely to get into a traffic accident than when his attention is focused solely on driving. An explanation of this phenomenon begins with the idea that the human mind has a limited amount of attention that must be divided among the various tasks or activities being executed simultaneously. Thus, if talking on the phone and fixing the hair each take up some amount of attention, there is less available to devote to driving than there would be otherwise, and driving performance suffers.

Other examples of divided attention abound in everyday life. When three children are simultaneously asking for their mother's attention, she cannot respond to all three in the same way she would be able to each child alone. When a policeman asks a potential witness on a park bench whether or not she had seen an average height male wearing a brown jacket go by, that witness' memory may be quite different depending upon whether the park was crowded (and her attention was divided between many people) or nearly empty (and her attention was divided between fewer people). Similarly, the ability to locate a friend's face in a crowd becomes more difficult as the size of the crowd increases.

Thus far, the word "attention" has been used rather loosely. To remedy the situation, it is necessary either to clearly define what is meant here by attention or to abandon the term in favor of a more precise term, or multiple terms, for the phenomena and processes under investigation. In this decision, I follow the lead set by Pashler in his recent review of the psychology of attention (1998) and choose the latter. Thus, while this is a study of divided attention, these words will only rarely come up such as when referring to the effect of doing multiple tasks at once or processing multiple items at once as a divided-attention effect. To retain the unrestricted use of the term "attention" could easily lead to misunderstanding because of the familiarity of the term in ordinary language. The concepts and experiences commonly associated with the term are likely to enter and weave their way into the reader's thoughts, thus potentially obfuscating the issues presented here. The more precise term that has already been introduced, and henceforth will be used instead of "attention," is "capacity". The term "capacity" also carries with it connotations from ordinary language, but hopefully fewer and less confusing ones, and indeed, it will be gradually distilled into a precise, mathematically-formulated meaning.

Capacity

Before abandoning the term "attention," it is useful to discuss the meaning of capacity and how it relates to the ordinary experience and understanding of "attention". Capacity and selectivity are considered to be the two primary aspects of attention. Capacity is the focus of this study. It has to do with limits in the ability to do multiple things at once as illustrated in the examples above. Selectivity, which will not be addressed further in this paper, concerns the ability to select and process only a portion of the stimuli available for processing. For example, in reading this paper, you are able to select one section of text at a time to process. From here on, the focus will be on the specific aspect of attention referred to as capacity and on the limits of capacity. It is a term that Broadbent (1958) used when he ushered in the modern era of attention research and proposed a limited-capacity view of human information processing to account for divided-attention effects.

A capacity limit may be defined in two ways: First, as *the maximum amount of information that can be stored*, when referring to a storage-capacity limit, or second, as *the maximum amount of information processed per unit time*, when referring to a processing-capacity limit. An over-simplified but useful analogy is the following. Think of information like water. Information moves through the human information-processing system like water flowing through a system of pipes and tanks. A storage-capacity limit of a tank is its volume—the maximum amount of water it can hold. Similarly information may be stored in short term memory or other stores in the information-processing system, and potentially, there is a limit to the capacity of those stores. A pipe also has a capacity limit. There is a maximum volume of water than can pass through the pipe per unit time, a maximum flow rate. Given a sufficiently high water pressure input, smaller diameter pipes have a smaller flow rate than larger pipes. Similarly, processes in the human information-processing system may have a maximum information processing rate. In this study the focus is on measuring and comparing the processing-capacity limits of visual perception and memory encoding, two of the "pipes" in the visual system.

Because human information processing is much more complex than plumbing, this simple analogy is not complete. Instead, an understanding of capacity limits in processing requires in-depth investigation and a more complex model. The architecture of the information-processing system itself is not yet well understood, let alone the nature and source of capacity limits. To begin the investigation of processing-capacity limits, consider the following definitions. *Unlimited capacity* means that there is enough processing capacity available for all the stimuli being considered to be processed independently. That is to say that the processing of one stimulus is unaffected by the processing of other stimuli. *Limited capacity* then means that there is *not* enough processing capacity available for all the stimuli to be processed independently. A particular kind of capacity limit is known as *fixed capacity*. In this case, a fixed amount of capacity is available, and it is divided up among the stimuli. There are other possible kinds of limits that will be described later, but the fixed capacity concept is a simple idea that will serve as an anchor in the investigation of capacity limits.

On the scale of the whole information-processing system, it is known that divided-attention effects exist, as illustrated in the several everyday-life examples presented above and confirmed through laboratory research to be discussed in detail below. Thus the entire system as a whole is, by the definitions presented, limited in capacity. However, it is not yet well understood what is the source, or sources, of this limit. Where in the system do the processing bottlenecks occur? What are the capacity limits at each stage or component of the system? These are questions addressed in this study.

In accomplishing a simple task such as searching a set of stimuli for a particular target stimulus (e.g. searching for a friend's face in a crowd), it is thought that the information comprising the stimuli must pass through a series of stages in the human information-processing system including sensory encoding, sensory storage, perception, memory encoding, and memory storage. To complete the task and provide a response such as saying, "yes, the target is present in the set" (or "there's my friend!"), the stages of memory retrieval, response selection, and response execution must also occur.

Capacity limits could potentially occur at any stage of processing. In order to measure the capacity limit of a particular stage, a divided-attention task must be devised that isolates that stage. In other words, the ideal is to control all stages but the one of interest with respect to a set-size manipulation, and then measure the effect of set size on task performance. The effect is typically a decline in accuracy and/or increase in reaction time with an increase in set size. In this way, the set-size effect can be attributed to the particular stage of interest and the capacity limit of that stage can be interpreted from the set-size effect. In this study, the capacity limits of interest are those present in visual perception and memory encoding.

Perception and Capacity

A prototypical task for studying the effects of divided attention in perception is the yes-no visual search task. The observer searches for a target stimulus among a set of stimuli and responds "yes" or "no" if the target is present or absent, respectively. The stimuli-set size is manipulated and the set-size effect on performance (accuracy and/or reaction time) is measured. This is the laboratory version of the everyday-life tasks such as that of searching for a friend's face in a crowd or for your favorite cereal on the supermarket shelves.

The ideal implementation of this task isolates the perceptual stage of processing by controlling the other stages with respect to the set-size manipulation. The sensory stage may be controlled by careful placement of stimuli in the display to avoid crowding, configural, and eccentricity effects that could be confounded with set size. The memory and response stages are controlled because the simple yes or no response is the same for each set size.

Researchers observed that search performance either remains approximately constant or decreases with increasing set size, depending upon the task and stimuli. One common view of search due to Treisman & Gelade (1980) emphasizes two findings. First, when the target differs from the distractors in only one feature, such as searching for a red dot among green dots or a horizontal line among vertical lines, reaction time is constant with respect to set size. This first type of search is referred to as *feature search*. Second, when the target differs from the distractors in more than one feature, such as searching for a vertical red line among horizontal red and vertical green lines, reaction time increases linearly with set size. This second type of search is referred to as *conjunction search* because a conjunction of features defines the target.

The picture is more complicated than two kinds of results for the two kinds of search. It was found that target-distractor discriminability variation could cause feature search to yield large set-size effects like those found in conjunction search (Bergen & Julesz, 1983; Duncan & Humphreys, 1989; Nagy & Sanchez, 1990; Palmer, 1994; Palmer, Ames, & Lindsey; 1993; Pashler, 1987) and conjunction search to yield little or no set-size effect like previously found only in feature search (McLeod, Driver, & Crisp, 1988; Nakayama & Silverman, 1986; Steinman, 1987). Specifically, when the target is difficult to discriminate from the distractor, such as searching for a vertical line among very slightly tilted lines, feature-search performance decreases dramatically with increasing set size. Also, when the target is easy to discriminate from the distractor through the appropriate choice of features and discriminability on each feature, conjunction-search set-size effects can be eliminated.

Some researchers account for the results found in visual search with theories that include the assumption of a very limited-capacity perceptual process. However, other researchers have shown that an assumption of limited-capacity is not necessary to account for the results. These two perspectives are discussed next.

Limited-Capacity Perception Models

Treisman and her colleagues designed what they called a feature-integration theory of attention (Treisman & Gelade, 1980; Treisman & Sato, 1990) to account for the two basic early findings and later made adjustments to it to account for the more recent findings. The guided-search model of Wolfe and his colleagues (Wolfe, 1994; Wolfe, Cave, & Franzel, 1989) is similar to the feature-integration theory. An assumption in both feature-integration and guided search is that perception is limited-capacity.

In the feature-integration theory, the features (color, orientation, spatial frequency, etc.) of the set of stimuli in a visual display are processed separately and in parallel in the first stage of perception. Although parallel processing does not imply unlimited capacity, it is commonly assumed to be such in this model and others similar to it. In feature search this first stage dominates perceptual processing, yielding the result of little if any effect of set size on performance. The capacity limit comes in at the second stage in which feature representations are integrated into object representations in a serial fashion based upon location in the visual field as represented in a "master map," a concept similar to the attentional spotlight metaphor described by Posner (Posner, Snyder, & Davidson, 1980). This second stage accounts for the findings of large set-size effects in conjunction search: Large set-size effects are produced because features must be integrated before the target can be identified and the integration occurs one stimulus at a time. Serial processing is necessarily limited-capacity: stimuli are not processed independently. Thus this two-stage perceptual process that consists of a potentially unlimited-capacity first stage followed by a limited-capacity second stage is overall a limited-capacity process. The feature-integration theory was later adjusted to account for the other findings in visual search. It was claimed that when target-distractor discriminability is low, the first stage does not work as efficiently and the serial "attentional" mechanism is necessary to perform the fine discriminations yielding larger set-size effects in feature search; when target-distractor discriminability is high, a feature-inhibition mechanism allows faster search on other features, yielding smaller set-size effects in conjunction search.

Wolfe proposed a similar two-stage model of perception called guided search based in part upon Treisman's feature-integration theory but with more specificity allowing quantification and simulation (Wolfe, 1994; Wolfe, Cave, & Franzel, 1989). In this model, the first-stage is the same kind of parallel-channel feature processor used by Triesman. Each feature has a set of broadly-tuned channels, a concept used widely in vision research and supported by physiological evidence (e.g. Maunsell & Newsome, 1987). What is new is that the output of this stage guides the second, serial, limited-capacity stage so that locations are not sampled randomly, as in the feature-integration model, but rather in order of likelihood for containing the target. The visual field is represented as an activation map. The activation at each location is the weighted average of bottom-up and top-down evidence for the target at that location. The bottom-up evidence comes from the first-stage feature processor and is based upon differences between stimuli in a local region. The top-down evidence selects the output of one of the broadly-tuned channels per feature thus activating one particular feature value over others (e.g. selecting the red channel when searching for a red dot among green dots or a vertical red line among horizontal red and vertical green lines). Wolfe accounts for discriminability effects as follows. In the first stage, discriminability affects utilization efficiency of the parallel channels. In the second stage, discriminability affects signal-to-noise ratio of the activation map that determines order of location sampling.

Unlimited-Capacity Perception Models

Whereas in the models described above, a limited-capacity perceptual process was used to account for visual-search results, other researchers have shown that no such assumption is necessary (e.g. Eckstein, 1998; Palmer, 1994; Palmer & McLean, 1995; Shaw, 1984; Tanner, 1961). The key difference between the two perspectives is in the explanation of the source of errors.

In the feature-integration theory, which is focused on reaction-time predictions, errors are largely ignored. In the guided-search model, errors can occur only as a result of limited stimuli exposure duration shortening the serial search so that there is not enough time to process every stimulus. A guess is made about the identity of the remaining stimuli, and errors can occur as a result of guessing. Errors are not a result of stimulus representation properties. These models implicitly assume a "high threshold," meaning that a distractor cannot be mistaken for the target.

The alternative is to assume that distractors *can be* mistaken for the target: the "low threshold" assumption. Two classical theories that are based upon this assumption are Signal Detection Theory (Green & Swets, 1966) and the Biased-Choice Theory (Luce, 1959). The models to be discussed next that are designed to account for visual search results are based upon the Signal Detection Theory (SDT) framework. In this framework, stimulus representations are noisy, not perfect. Such internal noise can be attributed to the variability of neuron firing, a fact supported by physiological studies of neural activity, and to noise in the physical stimulus. Because of noise, a distractor can engender a representation that falls above the "low threshold" criterion for classifying it as a target – an error termed a false alarm. Errors can also occur when, due to noise, a target engenders a representation that is misclassified as a distractor — an error termed a miss. Set-size effects can arise when multiple noisy representations are mapped onto a single response in the decision process. With each additional stimulus, there is an additional noisy representation and hence another chance for an error. Thus, by this perspective, when performance decreases with increasing set size, it is not necessarily due to limited-capacity perception.

Palmer and his colleagues (Palmer, 1994; Palmer, 1995; Palmer, Ames, & Lindsey, 1993) have shown that this kind of SDT-based, unlimited-capacity perception model can account for set-size effects in feature search for a variety of different stimuli: searching for a longer line among short lines, a higher luminance disk among low luminance disks, and several others. In these experiments, stimuli were presented briefly and accuracy was measured. Presentation duration was kept well within the duration of a single eye fixation to control eye movements which are possible in the typical reaction-time experiments in which the stimuli are displayed until response. Thus, one advantage to the brief presentation is that any set-size effects that are due to eye movements, a necessarily serial process, are controlled. Another advantage is that a brief presentation can bring performance down from ceiling level and variations in performance due to set size and other factors can be observed.

Target-distractor discriminability effects were addressed in this research using a technique borrowed from psychophysics. For each kind of stimuli and each set size, target-distractor difference was varied, and a psychometric function relating performance to this difference was fit to the data. For example, when searching for a longer line among short lines, the line-length difference was varied. When the difference was very small, performance was near chance. As the difference was increased, performance increased, eventually up to perfect performance. Based upon the psychometric function, a threshold target-distractor difference that would produce a set criterion performance level was estimated. In this way, a threshold was obtained for each set size, and set-size effects were based upon changes in that threshold with set size. This analysis technique provided the advantage that set-size effects could be compared across the different kinds of stimuli.

The SDT-based unlimited-capacity perception model accounted for discriminability effects in terms of the separation between the target and distractor representation distributions (d'). The more discriminable the target and distractor, the greater the separation between the means of the two distributions, resulting in fewer errors. The less discriminable the target and distractor, the smaller the separation, resulting in more errors. The model produces a probability correct prediction that is a function of discriminability and set size.

Eckstein (1998) has used a similar SDT-based unlimited-capacity model to account for set-size effects in contrast and orientation conjunction search. By his account, the larger set-size effects in conjunction search over feature search occur because noisy information about each feature is combined to produce a stimulus representation. The representation of a conjunction of features has more noise than the representation of one of the features alone. The additional noise increases the set-size effect due to decision. No limited-capacity, serial-search mechanism is required.

Palmer and McLean (1995) have also demonstrated that a limited-capacity perceptual process is even unnecessary to account for reaction time results in visual search with stimulus displays that are presented until the response is made. An unlimited-capacity, parallel search model with error-free processing predicts no search time set-size effect. However, when error is allowed, the unlimited-capacity, parallel model predicts arbitrarily large search-time set-size effects, depending upon target-distractor discriminability. Palmer and McLean made use of the diffusion-process random walk to predict visual search reaction time.

Mixed Evidence in Simultaneous vs. Successive Paradigm

Another approach to capacity-limit investigation of visual perception is the simultaneous vs. successive visual-search paradigm. The simultaneous condition is the same as the standard visual-search task. The stimuli are presented all at once. The observer searches the set for a target stimulus. In the successive condition, the individual stimuli in the set are presented one at a time. If perceptual-processing capacity is limited, then performance should be better in the successive condition in which full capacity can be dedicated to each stimulus, than in the simultaneous condition in which stimuli compete for a limited amount of capacity. If perceptual-processing capacity is unlimited, then there should be no difference in performance in the two conditions.

Results from the simultaneous vs. successive paradigm are mixed. Duncan (1987) and Kleiss and Lane (1986) found an advantage of successive over simultaneous display for words among similar letter strings and for difficult-to-discriminate letters, respectively, suggesting limited-capacity perception. On the other hand, Shiffrin and Gardner (1972) found no performance difference in the two display conditions with letter stimuli as would be expected if perception is unlimited-capacity. Duncan (1980) also found no performance difference between the simultaneous and successive conditions in visual search for two independent letter targets except when separate responses for the two targets were required. He concluded that capacity limits occur after stimulus identification, not in perception.

Memory Encoding and Capacity

The process by which information is encoded into short-term memory is another possible locus of capacity limits in the visual system. The capacity limit of memory encoding has not been researched as much as perception capacity limits or memory storage-capacity limits. Whenever memory is employed in a task, memory-encoding must occur. A typical memory task involves the presentation of stimuli followed by a memory test in which the observer is asked to recognize or recall items in the stimulus set. Varying stimuli set size in this kind of task allows for the study of divided attention in memory encoding. Again, the ideal would be to control all aspects of processing with respect to the set-size manipulation with the exception of the process of interest, in this case, memory encoding.

For example, consider the following postcue memory task: A set of stimuli are presented briefly. A cue is then presented at the location of one of the stimuli in the previous display along with a memory probe stimulus presented at the center of the display. The observer decides whether or not the cued stimulus matched the memory probe and responds "yes" or "no", respectively. The stimuli-set size is manipulated and the set-size effect is measured. As in visual search, the response end of processing is controlled with respect to set size because of the simple binary response, and the front/sensory end is also controlled due to careful placement of stimuli. What is different is that in search, the information encoded into memory need not depend upon set size: only information about the single target need be remembered to perform the task (was it present or absent?). The distractors need not be remembered, and indeed there is evidence that they are not (Brand, 1971; Gleitman & Jonides, 1976). In the memory task, however, all of the stimuli must be remembered to perform correctly. Unfortunately, in this task, the memory encoding process is not totally isolated. Perception must occur prior to memory encoding, and so set-size effects must be attributed to the joint processes of perception and memory encoding. However, if the capacity limit of perception is known, then the memory-encoding capacity limit may be determinable from the joint-processes limit. Any potential capacity limit imposed by memory storage may be avoided by keeping the largest set size smaller than the maximum short-term memory storage capacity.

As with perception, there is disagreement among researchers regarding capacity limits of memory encoding, although the weight of the evidence seems to fall on the side of limited-capacity, and in particular, fixed-capacity, memory encoding. The evidence comes from a variety of paradigms, some of which are better than others at isolating the memory-encoding process.

Evidence for Unlimited-Capacity Memory Encoding

First, on the side of evidence for unlimited-capacity memory encoding, some researchers have found that the recall of the last several words of a list presented auditorily is unaffected by a concurrent card-sorting task (Murdock, 1965), a visual-manual choice reaction time task (Anderson & Craik, 1974), or an arithmetic task (Silverstein & Glanzer, 1971). Similarly, in a study by Pashler (1993), visual recognition performance was unaffected by a simultaneous auditory-manual choice reaction time task. In these experiments there was no evidence of an effect of dividing attention between dual tasks in different modalities.

There is very little evidence for unlimited-capacity memory encoding in single modality paradigms. However, McLean, Palmer, and Loftus (1998) found evidence for unlimited-capacity memory encoding in coarse localization of letters. In this task, either two or four letters were presented briefly, followed by a memory probe. The observer's task was to indicate whether the probe letter had appeared on the left or the right side of the previous display. Display contrast was varied, contrast thresholds were estimated, and set-size effects were based upon changes in threshold with set size. Using a SDT-based model much like the one used by Palmer and Eckstein for visual search but applied to the localization-memory task, it was determined that decision alone could account for the observed set-size effects, implying unlimited-capacity perception and memory encoding.

Evidence for Limited-Capacity Memory Encoding

There is a larger body of evidence for limited- than for unlimited-capacity memory encoding. As one example, Palmer (1990) conducted a memory experiment similar to the example post-cue task described above. In his experiment, the study display was a set of 2 or 4 short horizontal lines. A test line appeared after a delay at one of the study stimulus locations. Observers determined whether the test line was longer or shorter than the corresponding study line. In this type of task, he argues, decision processes are controlled with respect to set size because, once the cue/test stimulus is presented, only one study stimulus representation need be evaluated. There is no integration of multiple noisy representations, a many-to-one mapping from representations to response selection, such as was required in the SDT-based visual search model.

Using the threshold-based, set-size-effect measure, Palmer found that the set-size effects in this memory task were relatively large as compared to the visual search results. To account for these results he applied the SDT-based model to this task, adding on a limited-capacity mechanism based upon the sample-size model by Shaw (1980). In the sample-size model, the variability of a stimulus-representation distribution depends upon the number of stimuli being processed. A stimulus representation is built up from information gathered over a number of samples of that stimulus. If there is an unlimited number of samples available, then there is no capacity limit due to this mechanism. However, if there is a fixed number of samples available those must be divided over the stimuli. Then, as the number of stimuli increases, the number of samples per stimulus will decrease, resulting in greater variability of representation. This additional noise in the representations causes more errors to occur in decision with larger set sizes. Hence set-size effects are larger in the fixed-capacity version than the unlimited-capacity version of the model.

A way of explaining the fixed-capacity sample-size model in terms of the water flowing through pipes metaphor is that the pipe is a fixed size, as in fact, physical pipes are. That is, the capacity of information processed per unit of time is constant, regardless of set size. The more stimuli being processed, the lower the transmission rate of information about each individual stimulus.

The fixed-capacity model accurately predicts the set-size effects Palmer found. Hence, he concluded that perception plus memory encoding is limited in capacity. McLean et al. (1998) generalized these results to letter stimuli using a similar postcue task.

Another line of evidence for limited-capacity memory encoding comes from Shibuya and Bundesen (1988). Their fixed-capacity, independent-race model predicted performance in a partial report task. A circular array of digits and letters was presented at varying brief exposure durations followed by a pattern mask. The number of letters and the number of digits also varied. The task was to report as many of the digits as possible and to ignore the letters. The observer typed the digits remembered from the display into a keyboard in any order. There was a complex pattern of results, but the most basic and general findings were that the number of correctly reported items increased with increasing exposure duration and increased with decreasing letter set size for each digit set size. Also, the probability that a given digit was reported correctly was reduced by the presence of other stimuli, and more so by the presence of other digits than other letters.

Shibuya and Bundesen argued that their results imply limited-capacity processing. They suggested that an unlimited-capacity process predicts that the number of correctly reported digits should increase proportionally with the digit-set size provided that memory-storage capacity is not a limiting factor. They suggest that this condition would be met at low exposure durations because the number of stimuli sampled is close to zero. What they found was quite different than this prediction. The number of digits correctly reported showed little dependence upon the digit-set size when there were no letters in the display, a pattern they claim would be predicted by a fixed-capacity process.

There may be some problems with the description of the unlimited-capacity model predictions provided by Shibuya and Bundesen. Assuming memory-storage capacity is not a limiting factor, there are at least two reasons why an unlimited-capacity processing model would not predict the number of correctly reported digits to be proportional to the digit-set size. First, if representations are noisy, more digits means more chances for incorrect reporting. Also, the reporting response varies with digit-set size. Report takes more time with more digits, hence more forgetting may occur.

Also, the assumption is questionable that memory-storage capacity is not a factor at low exposure durations because the number of stimuli sampled is close to zero. The implicit assumption in this statement is that stimuli are processed serially. In a parallel model, all of the stimuli are sampled simultaneously, and presumably, each stimulus though perhaps represented rather imprecisely, will take one of a limited number of slots in memory storage. The size of storage capacity has been shown to be about 4 items (Luck & Vogel, 1997; Irwin & Andrews, 1996). In the experiment by Shibuya and Bundesen, the number of items in a display was between 2 and 10 and digit-set size was as large as 6. Hence memory-storage capacity-limits may very well be a factor in their results, even at short durations. Furthermore, if the digit-set size exceeds memory-storage capacity (which it may in their experiments), and if only stored digits can be reported, then the report response will limit performance more for larger set sizes than for smaller ones.

Similar problems exist with a second argument that they posed. This argument is based upon the result that when the digit-set size was constant, the rate of increase in performance with duration declined dramatically with increasing letter-set size. Shibuya and Bundesen explained that this result suggests that letter processing occurred, and the more letters that were present, the less processing capacity was available for the digits, indicating the existence of a processing-capacity limit. The problem is that an unlimited-capacity model may also predict a change in rate, albeit a small one, but there is no way of knowing without generating quantitative rate predictions for the two models where the data falls with respect to these predictions and which model fits the data better.

Nevertheless, Shibuya and Bundesen do offer a fixed-capacity memory-encoding model that provides quantitative predictions that account for the patterns of their results. In their model, information about the stimuli is encoded into memory via a sampling process from perceptual representations. They implicitly assume unlimited-capacity perception: The perceptual representations do not depend upon the number of stimuli. There is a fixed sampling capacity of C items per second, distributed over the stimuli set. Samples occur one at a time, randomly over the set of stimuli. One sample is sufficient to encode a stimulus perfectly, provided there is space in memory storage. The sampling process begins at a fixed amount of time following stimulus display onset and continues until either the number of items sampled equals K, which is the memory-storage capacity limit, or a certain fixed time from mask onset has elapsed. The number of correctly reported digits is basically the number of digits sampled. Like the feature-integration and guided-search models of perception, given sufficient storage capacity, errors only occur as a result of limited processing time, not as a result of noise in stimulus representation.

Comparison of Capacity in Perception and Memory Encoding

Beyond determining that a process is either limited- or unlimited-capacity, it is of interest to measure those potential limits and then compare limits in different processes. In particular, what are the capacity limits in perception and in memory encoding and how do they compare?

In their model, described above, Shibuya and Bundesen (1988) assume unlimited-capacity perception and fixed-capacity memory-encoding. Shibuya and Bundesen based their work in part upon the early research of Sperling who also addressed these issues though not within the context of divided attention.

Sperling (1960) originated the partial report task to investigate how much can be "seen" in a single brief exposure. He reported, "observers enigmatically insist that they have seen more than they can remember afterwards, that is, report afterwards." He devised a task that would eliminate the memory-storage capacity limit on performance. In this task, a matrix of letters (such as 3 rows of 4 letters) was presented briefly followed by a cue that indicated one row of the letter matrix. Observers reported as many letters as they could remember from that particular row. For short cue delays, the number of letters correctly reported in this partial report task was the same (about 4) as when no cue was given and observers reported as many letters as they could remember from the entire matrix (whole report task). As cue delay was increased, performance in the partial report task declined. Sperling accounted for these results with a model of the visual system that consists of an initial detailed representation of the stimuli, termed iconic memory, that fades quickly, and a fixed-capacity sampling process that encodes the stimulus information from the fading iconic representation into short-term memory, a process he termed "readout from visual storage" (Gegenfurtner & Sperling, 1993; Sperling, 1960; Sperling, 1969). Reinterpreting Sperling's model in terms of the present discussion, it could be said that his model assumed unlimited-capacity perception (to create the iconic representation) and fixed-capacity encoding into short-term memory.

Another example of a comparison between perception and memory-encoding can be drawn from the research of Palmer and his colleagues as described above. Set size effects in a visual search line-length discrimination experiment were predicted by an unlimited-capacity model (Palmer, et. al., 1993), but a similar experiment with an added memory component yielded larger set-size effects that were accounted for by a fixed-capacity model (Palmer, 1990). These two results taken together imply unlimited-capacity perception and fixed-capacity memory encoding plus perception. However, there is a problem with drawing this comparison: In one experiment partial report cues were used and not in the other.

Finally, McLean, et. al.(1998) measured and compared capacity limits in perception and memory encoding of letters. A portion of their results were described above. They designed matched search and memory tasks so that perception capacity and memory-encoding plus perception capacity could be directly compared. McLean et. al. used a SDT-based model much like the one used by Palmer (1990, 1994, Palmer et. al., 1993), but with an added capacity-limit parameter that reflects the degree to which representation variance is affected by set size. When this parameter is equal to 0, the model is unlimited capacity. When the parameter is equal to 1, it is fixed capacity. These two models are equivalent to Palmer's SDT-based models. However, the parameter can also take values other than 0 or 1 indicating varying magnitudes of capacity limits.

The best-fitting capacity-limit parameter was estimated for each of the tasks. It was found that for coarse localization search, set-size effects were consistent with unlimited-capacity perception, as described above. Surprisingly, set-size effects were similar for the memory version of the task, indicating unlimited-capacity memory encoding, also described above. Set-size effects for yes-no search and memory were similar with the effects slightly larger for memory, but neither were consistent with unlimited-capacity or fixed-capacity. Instead, the best-fitting parameter was in between 0 and 1, at approximately 0.3, indicating limited capacity but not to the extent of the fixed-capacity model.

Unfortunately, the results of McLean, et al. were not definitive. The data was noisy which may have been due to using letters as stimuli rather than simpler stimuli such as those used in the Palmer studies. Some of the letters were easier to recognize than others. Also the letters, relatively complex multi-dimensional objects, were modeled as single dimensional representations where that one dimension reflected discriminability from the target. Discriminability was manipulated indirectly, not by varying target-distractor difference, but rather by varying display contrast or duration. Furthermore, there were substantial individual differences in the data, perhaps caused by use of different strategies. The main goal of the present study is the same as that of McLean, et. al. (1998): to measure and compare capacity limits in perception and memory encoding. However, improvements were made over the earlier study to reduce noise and individual differences with various controls, stimulus design, and multi-dimensional modeling.

Purpose of this Study

The discrepant conclusions regarding the capacity limits of perception and memory-encoding processes may arise because of differences in question-framing and theoretical context as well as differences in task and stimuli. Indeed capacity limits may be stimulus and task specific. In this study, I seek to begin to resolve the discrepancies by using very simple tasks and stimuli for both perception and memory encoding and by providing a unified theoretical context to measure and compare capacity limits in these two processes.

By measuring capacity, the question of unlimited or limited capacity can be answered, but we can also go beyond this gross distinction and determine the degree of limited capacity. On this scale that will be defined in the Theory section below, unlimited capacity is one landmark at 0 and the particular form of limited capacity referred to as fixed capacity is another landmark at 1.

The following is an outline of the remainder of the article. In the next section entitled *Experiments — Part 1*, two experiments are reported that were designed to replicate prototypical perception and memory encoding tasks such as those discussed above. This is followed by a *Theory* section in which is presented a unified theoretical context for capacity-limit measurement in perception and memory encoding. Control experiments relevant to the construction of the theoretical context are reported in *Appendix A* and *Appendix B*, and the mathematical modeling details of the theory are provided in *Appendix C*. Then, in *Experiments — Part 2*, the capacity measurements of the first two experiments are reported using a metric derived in the theory section. A final experiment is also reported in which perception and memory-encoding capacity are measured together in matched divided-attention tasks in a single experiment and thus may be directly compared.

EXPERIMENTS - PART 1

The first two experiments are reported in this section, the purpose of which was to replicate prototypical experiments that address capacity limits in perception and memory encoding, respectively. The first experiment, which addresses perception, is a feature search experiment. The second experiment, which addresses memory encoding, is a postcue memory experiment. In both experiments, set-size effects are measured in yes-no tasks using small ellipses as stimuli.

Feature Search Experiment

In this experiment, observers search a briefly displayed stimuli set for a target stimulus and then indicate whether the target was present in or absent from the set. Performance is measured as percent correct. The distractors, which vary in number, are homogeneous on each trial. The target, if present, is a stimulus that differs from the distractors in only one feature: contrast, orientation, or size. This experiment is modeled after the feature search experiments conducted by Palmer (1994) that are described above.

Method

Much of the methodology was the same over all experiments reported in this paper. The Feature Search Experiment methodology is described here in detail. For each subsequent experiment, only the methodology unique to that experiment will be described.

Observers. Six young adult observers who were University of Washington students or research staff were paid $10 per hour. All had normal or corrected-to-normal acuity.

Apparatus. Stimuli were displayed on an Apple Two-Page monochrome 1152 x 870 pixel video monitor controlled by a Macintosh computer with 75 Hz temporal resolution. The resolution was 30 pixels/cm which, with the viewing distance of 56 cm, resulted in 2 arc min per pixel for the central 1 degree viewing area. The entire display was 38 x 29 cm which subtended 37 degrees x 29 degrees. Dark stimuli were presented against a gray background of 34 cd/m^2 measured with room lights turned off. A dim lamp, lit during the running of the experiment, illuminated the display face by less than 0.1 cd/m^2.

Stimuli. The stimuli were small ellipses that varied in contrast, orientation, and size. There was a target and three types of distractors. All stimuli had an aspect ratio of 3 (major axis length / minor axis length). These ellipses were dark on a gray background. Contrast was calculated as (e-b)/b where e is the ellipse luminance and b is the background luminance, the usual calculation for aperiodic stimuli. The target ellipse was defined by -25% contrast, vertical orientation (i.e. 0 degrees of rotation from the vertical), and a major axis length of 30 arc min of visual angle. The three types of distractors differed from the target in one feature apiece: contrast, orientation, or size. The contrast distractor had a 14% decrement from the -25% contrast target. It was the same as the target in orientation and size. The orientation distractor was rotated clockwise 5 degrees from the vertical orientation of the target, but was the same as the target in contrast and size. The size distractor was 36 arc min in major axis length, a 6 arc min increment from the target, but the same as the target in contrast and orientation. The increment values were chosen to match target-distractor discriminability for the three different distractors at a level in which performance was midrange between chance and perfect as will be described in the Theory section with additional details in Appendix B. See Figure 1 for a schematic illustration of the four types of stimuli.

Set size varied between 1 and 4 ellipses as shown in the sample displays presented as Figure 2. In this figure, the set size 1 sample display contains the target and the set size 4 sample display contains the target and 3 distractors of the contrast increment type. Distractors are homogeneous: only one type of distractor is used in any given trial, and in fact, in any given block of trials.

The sample displays in Figure 2 schematically demonstrate the spacing and positioning of the stimuli and their size relative to the size of the display. To prevent

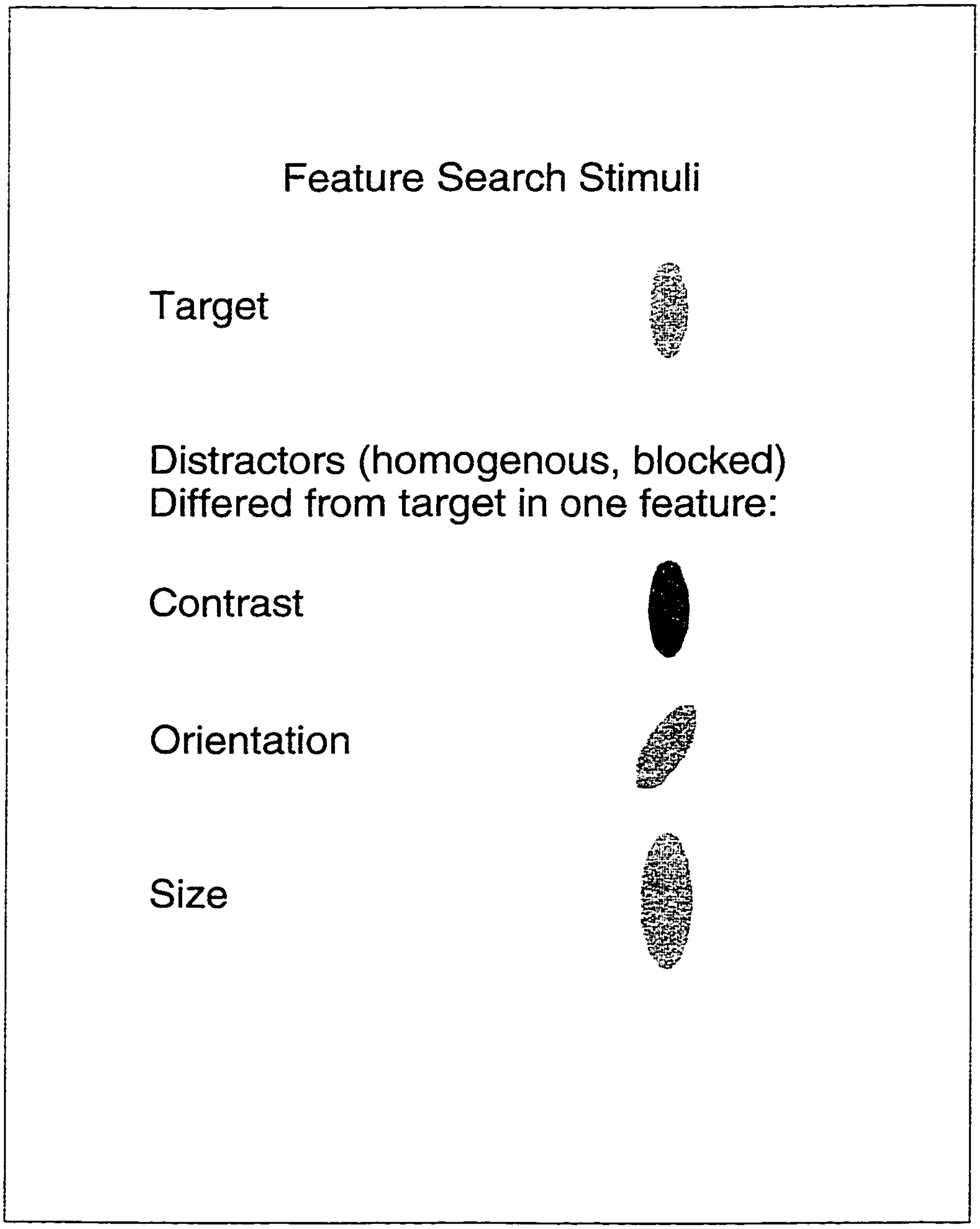

Figure 1: Feature Search Stimuli

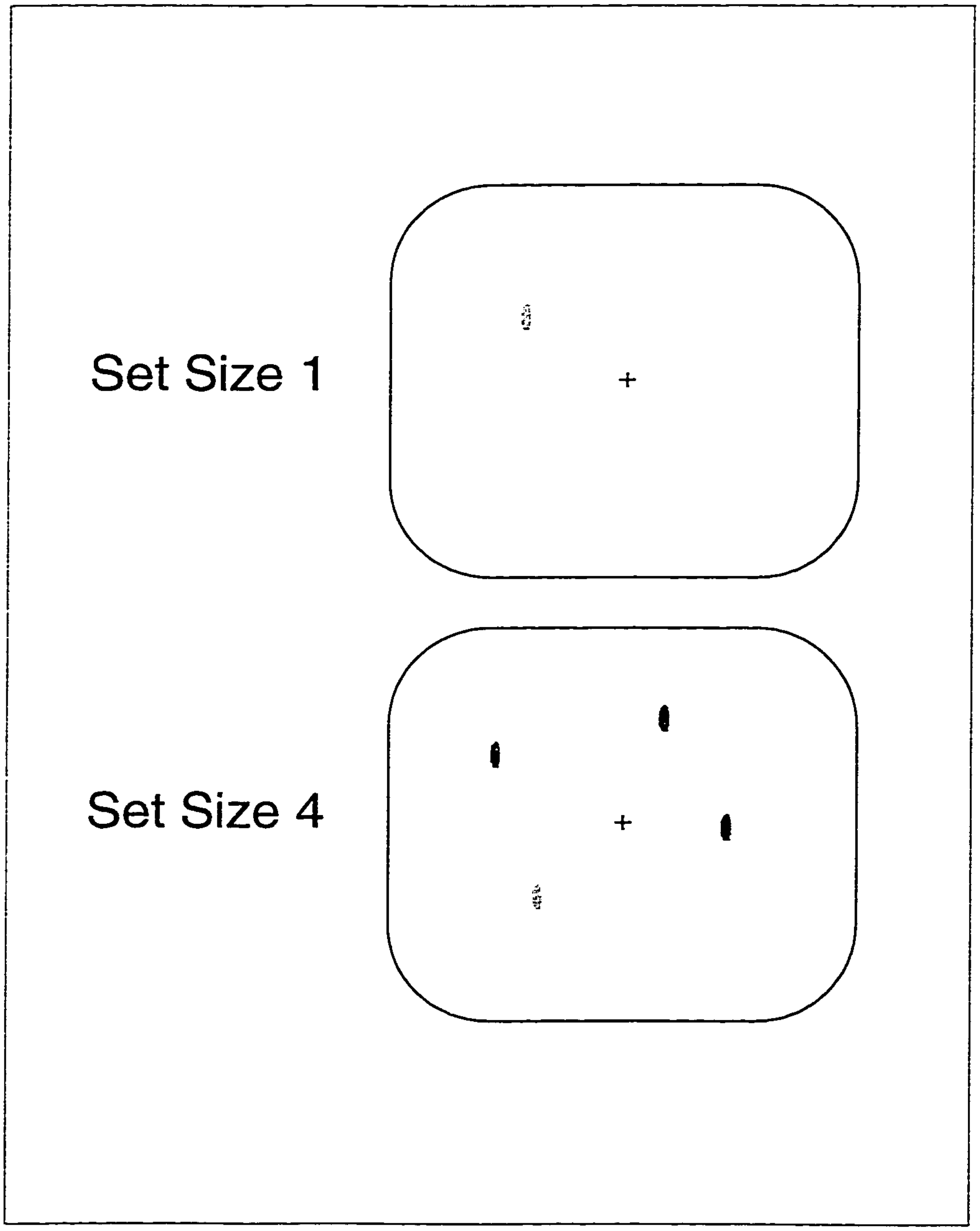

Figure 2: Feature Search Sample Displays

configural cues, the display locations were randomized within three constraints: (a) Eccentricity of the center of each stimulus was between 5 and 8 degrees of visual angle. (b) Center-to-center spacing between stimuli was at least 3 degrees of visual angle. (c) For set size 4, stimuli were chosen in pairs roughly opposite one another across the vertical midline of the display.

The position and movement of the eyes during the stimuli set display was controlled in two ways: First, the display was presented for a brief 80 ms duration to minimize the time available to make an eye movement. Second, observers were instructed to fix their gaze at a central fixation point prior to and during the stimuli-set display. This fixation point was marked by a plus-sign as shown in the center of the sample display in Figure 2.

Design and Procedure. As shown in Figure 3, the display sequence on a trial was the following: (a) the target ellipse was presented at the center of the display for 500 ms. (b) A blank screen was presented for 500 ms. (c) A plus-sign fixation point was presented at the center of the display for 500 ms. (d) The fixation point was joined by the study display — 1 or 4 ellipses — presented for 80 ms. (e) A blank screen was presented for 500 ms. (f) A question mark was presented as a response prompt.

Following this display sequence, the observer responded by pressing the "p" key to indicate "target present" or the "a" key to indicate "target absent." The correct response in the Figure 3 example is "p" because the target was present in the study display. There was no time pressure to respond. Two successive tones were used to provide feedback after an incorrect response. The next trial began 1500 ms after the response on a correct trial or the feedback tones on an incorrect trial.

Trials were presented in blocks of 36, not including 4 practice trials at the beginning of each block. A session consisted of 9 blocks and took up to one hour to complete. Observers typically completed one session per day or at most two, and the second only if they did not feel fatigued. Each observer participated in 4 sessions, for a total of 1296

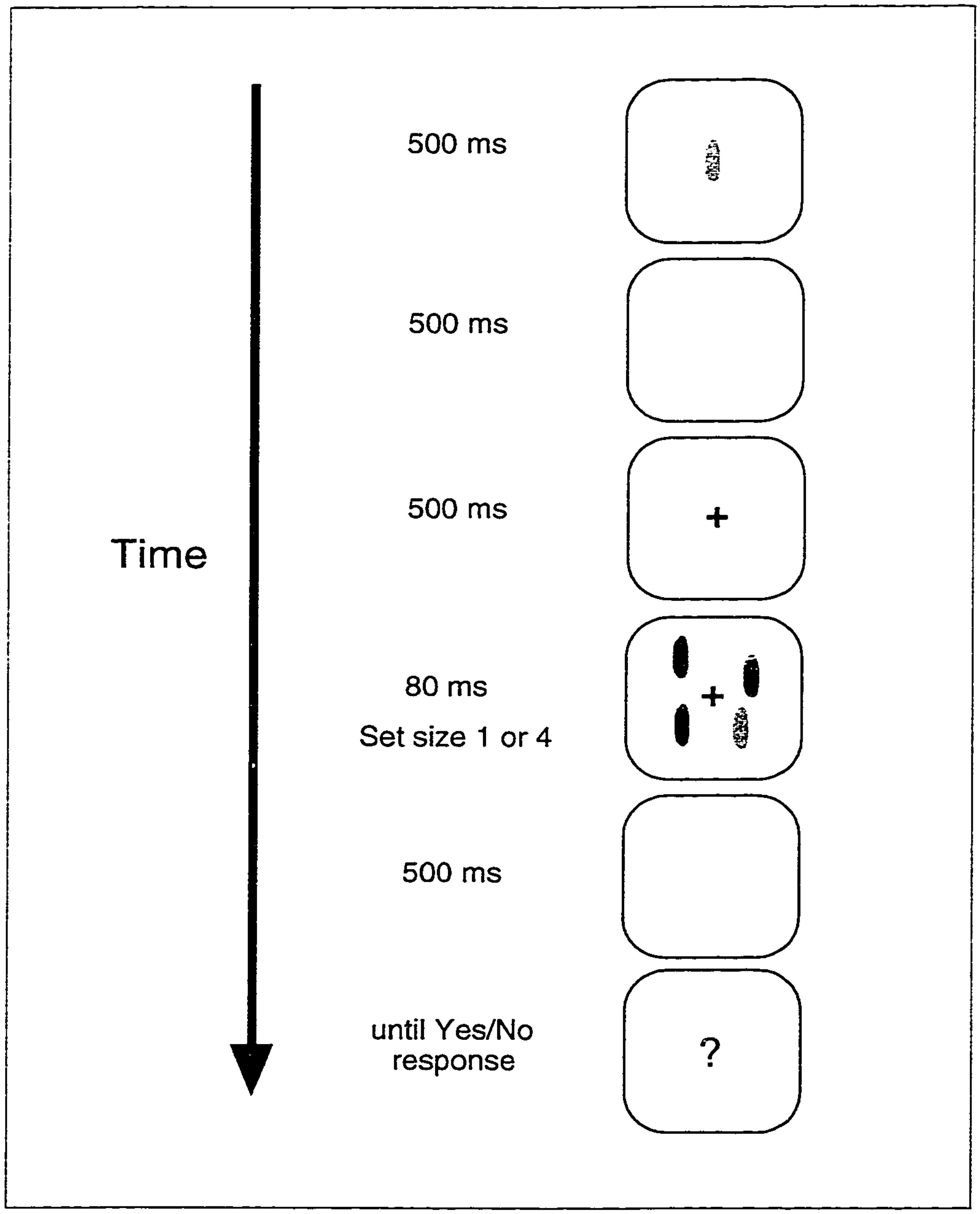

Figure 3: Feature Search Task

trials. Distractor type varied between blocks. Each of the 3 consecutive triplets of blocks within a 9-block session consisted of a block for each of the 3 types of distractors in random order within that triplet. Set size (1 or 4) and target status (present or absent) varied within a block: The 4 possible factorial combinations of these two variables each occurred 9 times within a block yielding 36 trials, presented in random order. Set size and target status were selected randomly for each of the 4 practice trials.

Observers were instructed that when guessing to respond "p" about as often as "a". They were given bias feedback at the end of each session: The final display of the session reported the percentage of trials in which the response was "p" and reminded them to continue to aim for 50% in the next session.

Each observer had participated in other similar experiments prior to this one and had at least one session of practice (360 trials) in this experiment before providing any of the reported data.

Results

Mean percent correct over observers for set size 1 was 85.0 ± 0.8 and for set size 4 was 71.9 ± 0.6. The results are shown graphically in Figure 4 in a form similar to that used by Mulligan and Shaw (1981) and Shaw (1980). The two dependent variables are plotted against each other: percent correct for set size 4 is plotted against percent correct for set size 1, yielding a scatterplot of 6 data points, one for each observer. As will be demonstrated below, this format is useful for displaying the full range of theoretical predictions produced by varying target-distractor discriminability. This type of graph will be used repeatedly throughout this manuscript for presenting both experimental results and theoretical predictions.

In addition to the data points for the individual observers, mean percent correct for each set size is represented as a point on this plot and is marked by an open circle with standard error bars. There was an effect of set size: percent correct for set size 4 was

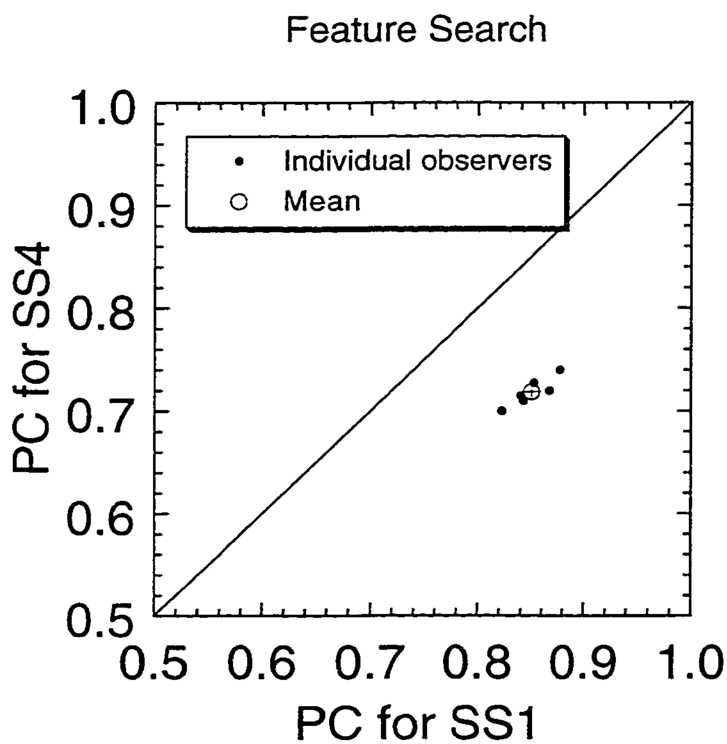

Figure 4: Feature Search Results

always lower than for set size 1 as indicated by the individual data points all falling below the identity line (included in Figure 4 for reference). The mean difference in percent correct for the two set sizes was 13.2 ± 0.4.

Discussion

Because performance declines with increased set size, it can be said that the information-processing system as a whole is limited-capacity — an individual stimulus is processed more effectively alone than when it is part of a display of 4 stimuli. However, at this point, a capacity limit in perception can only be inferred to the extent that all other processes are controlled or accounted for with respect to the set-size manipulation. Also, at this point, the magnitude of the set-size effect is difficult to compare with results found by other researchers in similar experiments because the magnitude depends upon the discriminability between the target and distractors. A theoretical context is needed to address these issues and provide a context for interpretation, but first consider a prototypical memory-encoding experiment.

Postcue Memory Experiment

In this experiment, observers view a briefly-displayed stimuli set, much like that of the previous Feature Search Experiment, but this time it is followed, rather than preceded, by a target stimulus. This target, acting now as a memory probe, is accompanied by a location cue. The observer indicates whether or not the target appeared at the cued location in the stimuli set. Unlike the Feature Search Experiment, the target varies from trial to trial and distractors are heterogeneous so that the target identity cannot be inferred from the stimuli set display as the only unique stimulus in the set. Set sizes used were 1 and 4. In a control experiment reported in Appendix A, it was verified that 4 items was within the memory storage capacity of the observers participating in this study. Thus limits on performance would not be attributable to memory storage capacity limits. This experiment was modeled after aspects of the partial report type experiments of Sperling (1960), Shubuya and Bundesen (1988), and Palmer (1990).

Method

The methodology of the Postcue Memory Experiment was the same as the Feature Search Experiment with the following exceptions.

Stimuli. The stimuli were 8 ellipses, illustrated schematically (not to scale) in Figure 5. Two values along each of 3 dimensions (contrast, orientation, and size) were chosen to define the 8 stimuli. Stimuli were -25% or -44% contrast, vertically oriented or rotated 7 degrees to the right, and 30 or 38 arc min in major axis length. These values were chosen to approximately match target-distractor discriminability and keep performance midrange as described in the Theory section with additional details in Appendix B. As in the Feature Search Experiment, the aspect ratio of the stimuli was 3.

The target stimulus varied from trial to trial and the distractors were heterogeneous so that the identity of the target could not be inferred from the stimuli set display and would instead be known when it was presented after the stimuli set. Distractors were sampled randomly with replacement from the full set of 8 stimuli on each trial, hence occasionally a stimulus, potentially even the target, would be repeated within a display. Sample displays for the two set sizes (1 and 4) are shown in Figure 6.

Design and Procedure. As shown in Figure 7, the display sequence on a trial was the following: (a) a question mark was presented at the center of the display for 500 ms. (b) A blank screen was presented for 500 ms. (c) A plus-sign fixation point was presented at the center of the display for 500 ms. (d) The fixation point was joined by the study display — 1 or 4 ellipses — presented for 80 ms. (e) A blank screen was presented for 500 ms. (f) The target stimulus was presented at the center of the display along with a cue as a prompt for response. This cue was a small but highly visible black dot that marked the location on which was centered one of the stimuli presented in the study display.

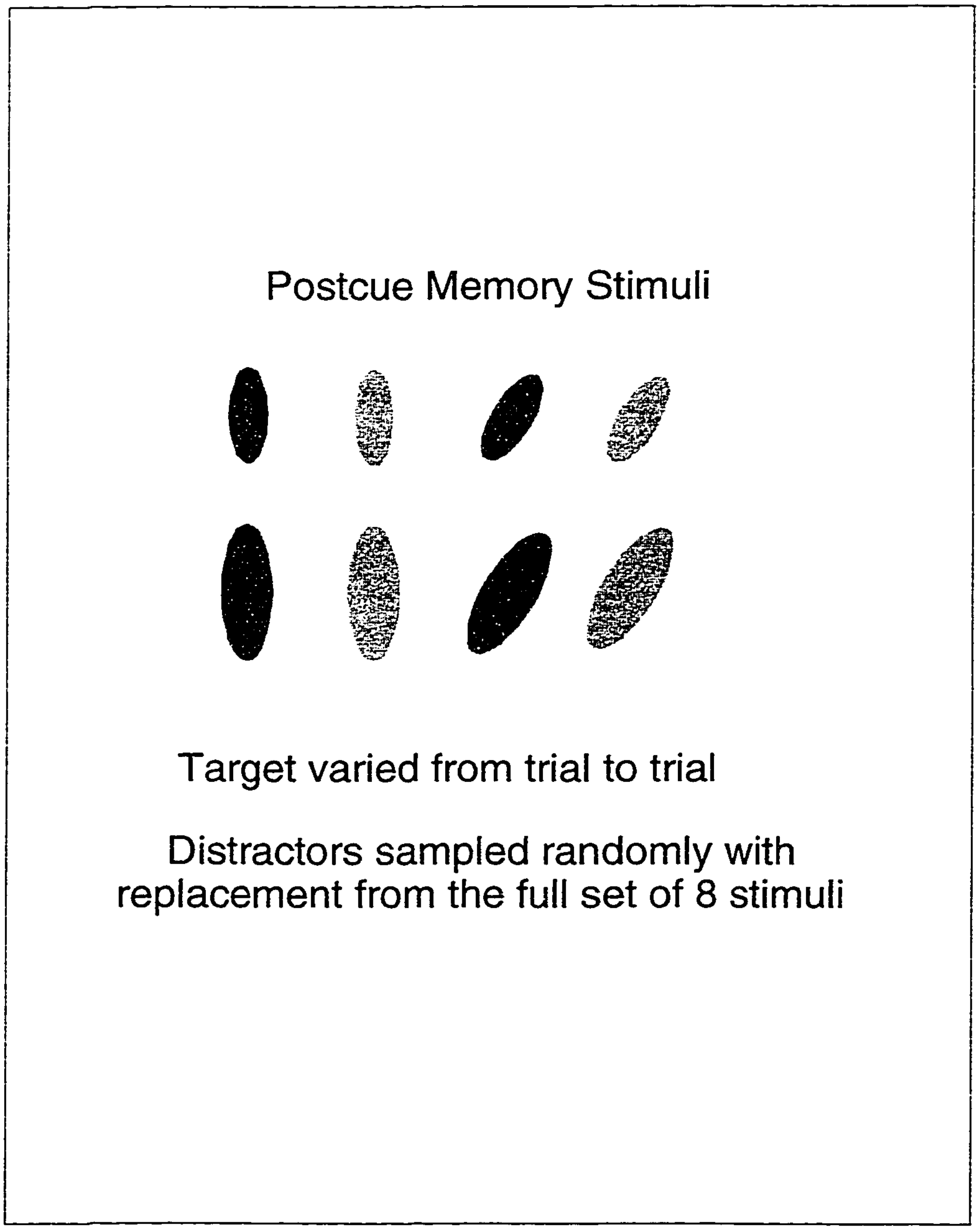

Figure 5: Postcue Memory Stimuli

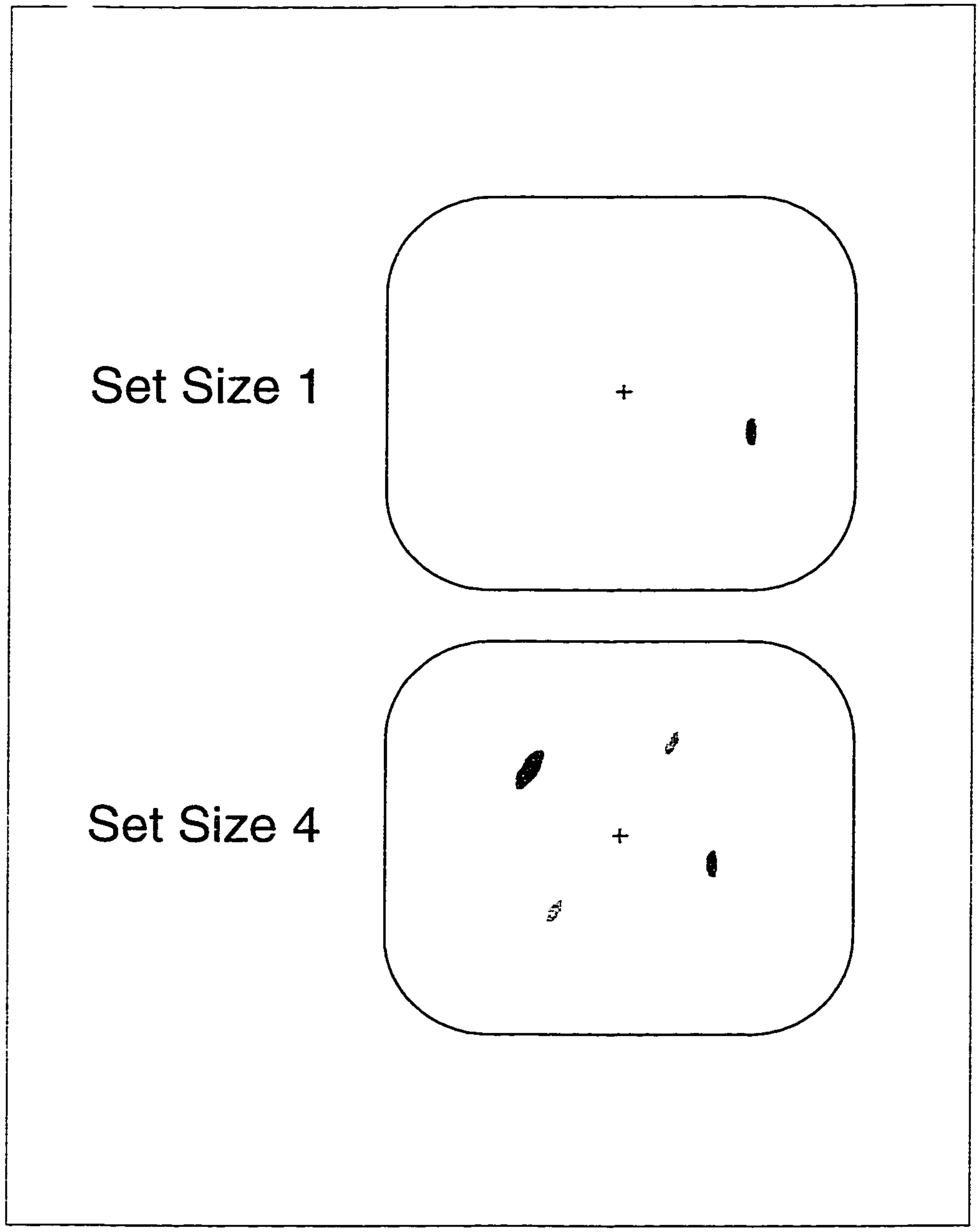

Figure 6: Postcue Memory Sample Displays

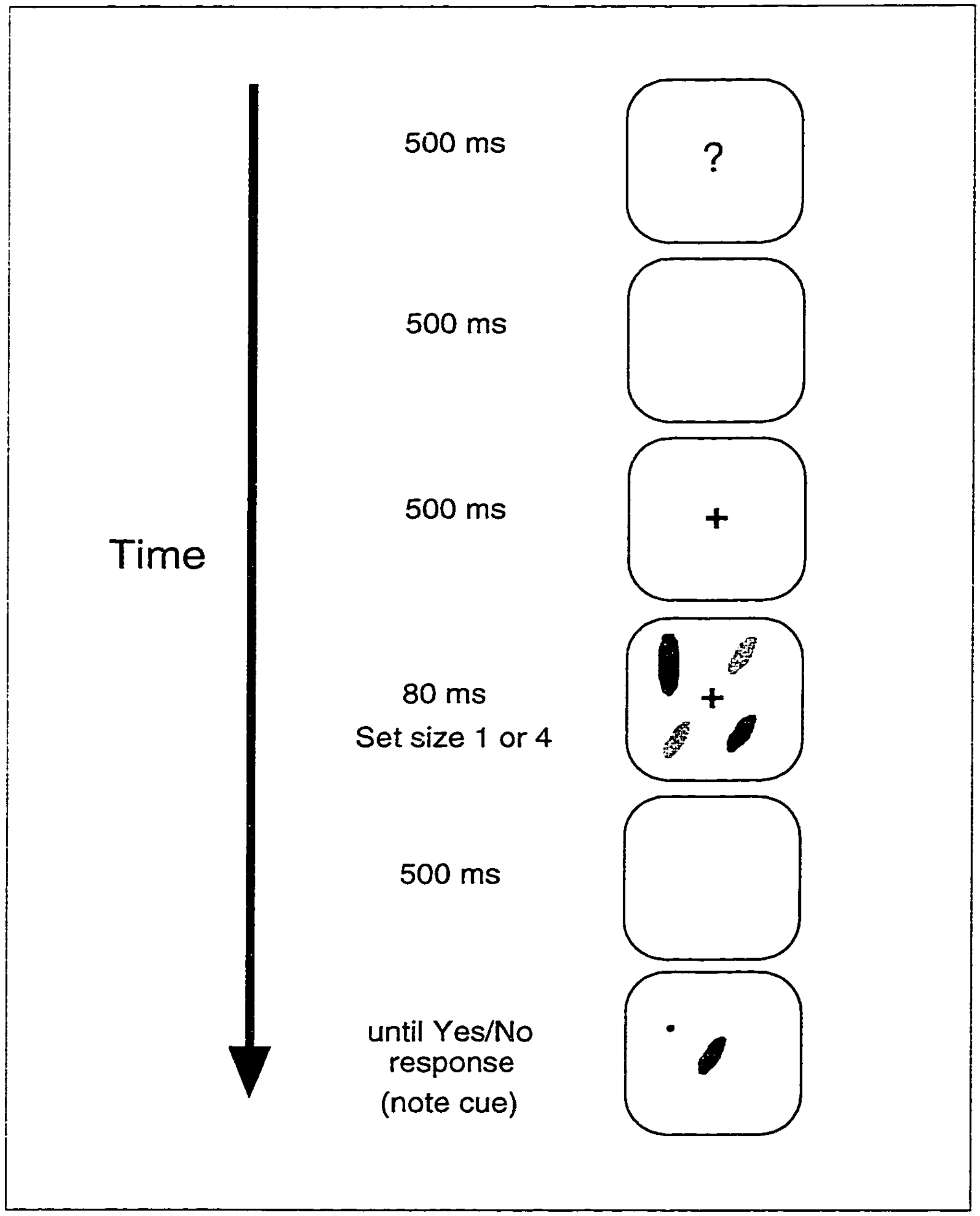

Figure 7: Postcue Memory Task

Following this display sequence, the observer responded by pressing the "p" key to indicate "target was *present* at cued location" or the "a" key to indicate "target was *absent* from cued location" The correct response in the Figure 7 example is "a". Although the target was present in the study display, it was absent from the cued location. Information about non-cued stimuli was not useful in accomplishing this task.

Trials were presented in blocks of 32, not including 4 practice trials at the beginning of each block. A session consisted of 10 blocks. Each observer participated in 4 sessions, for a total of 1280 trials. The target stimulus was selected randomly on each trial with the constraint that each stimulus serve as target equally often. Hence each of the 8 stimuli served as the target 4 times per block. Set size (1 or 4) and target status (present or absent) varied within a block: The 4 possible factorial combinations of these two variables each occurred 8 times within a block yielding 32 trials, presented in random order. Target stimulus, set size, and target status were selected randomly for each of the 4 practice trials.

Results

Mean percent correct for set size 1 was 88.6 ± 1.1 and for set size 4 was 73.0 ± 2.3. The results are shown in Figure 8 in the same form as the Feature-Search-Experiment results in Figure 4. Percent correct for set size 4 is plotted against percent correct for set size 1, yielding a scatterplot of 6 data points, one for each observer. Mean percent correct for each set size is represented as a point on this plot as well, marked by an open circle with standard error bars. There was an effect of set size: percent correct for set size 4 was always lower than for set size 1 as indicated by the individual data points all falling below the identity line (included in Figure 8 for reference). The mean difference in percent correct between the two set sizes was 15.6 ± 2.4.

Discussion

The same issues apply here as above in the Feature Search Experiment. Because performance declined with increased set size, it can be concluded that the information-

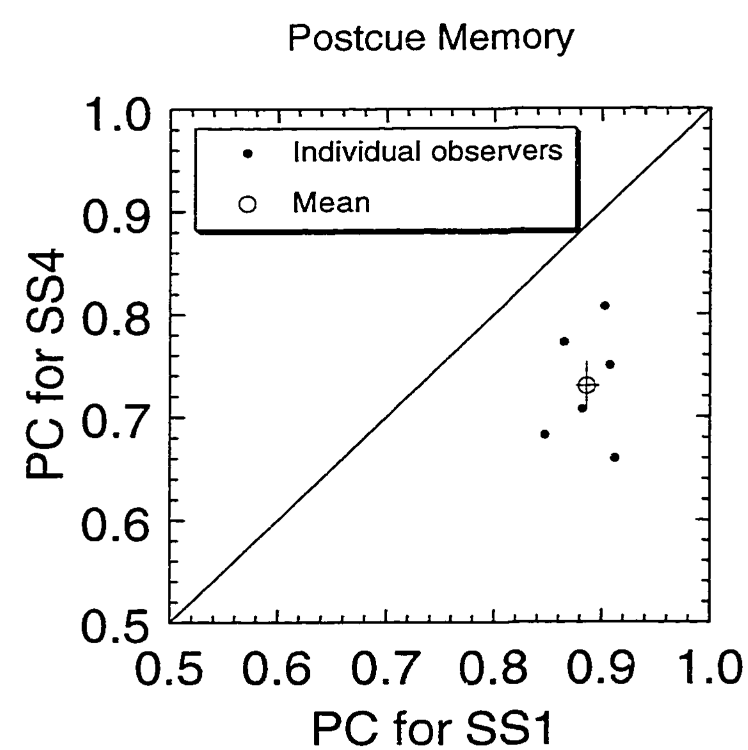

Figure 8: Postcue Memory Results

processing system as a whole has a capacity limit, but the contribution from memory encoding to that limit cannot be determined at this point. Also the magnitude of the effect, while larger than the Feature Search set-size effect when measured in terms of the difference between percent correct for the two set sizes, cannot validly be compared because of potential differences in target-distractor discriminability, not to mention other differences in the task, stimuli, and so forth. The degree of capacity limit cannot be inferred from the set-size effect magnitude at this point either. What is needed to answer these questions is a theoretical context which is provided next.

THEORY

The purpose of the theory is to provide a context for interpreting observed set-size effects in terms of the capacity limits of perception and memory encoding. In the presentation of this theory, there are three parts. First, the stimuli are chosen. Three control experiments were conducted to calibrate the stimuli for the observers by measuring discriminability thresholds on each of the three dimensions and then using those thresholds to choose the stimuli. Second, the visual system representations of those stimuli are defined. Potential capacity limits come in to the theory at this point in the way that the number of stimuli being processed affects the representations of the stimuli. Third, representations are linked to task performance for purposes of making quantitative predictions that can be compared to data.

Stimuli

The stimuli used in the experiments in this study are small ellipses. They were chosen because they are simple, having three well-defined, controllable dimensions: contrast, orientation, and size. Some other studies of divided attention have used more complex stimuli such as letters and words. Such complex stimuli have many dimensions that are difficult to define, vary from one another in ways that are difficult to control and model, and lead to the use of strategies. Some strategies arise because of associations due to familiarity such as when the letters in the stimuli set can be put together to form a word, acronym, or set of initials. Other inconsistencies arise due to varying levels of discriminability: A and D are more discriminable from one another than O and D, so the task of searching for a D is easier when the set consists of A and D than when it consists of O and D. Also, some letters are easier to identify than others. One subject reported that when she couldn't identify a letter, she would guess it was the most difficult-to-identify letter in the set of possible stimuli. Counterbalancing and randomization can help neutralize the effects of these inconsistencies, but at the same time, the averaging can create a result that is not representative of any of the trials. Additionally, if strategies are possible, it is more difficult to ensure that observers are all doing the same thing. Using the simple ellipses minimizes these problems.

Specific values along each of the three dimensions were chosen to define the ellipses used in the experiments in such a way that each of the stimuli were approximately equally discriminable from one another. This strategy avoids the confound of set size and discriminability and many guessing strategies such as those alluded to above. To illustrate the goal here, consider the 3-dimensional psychological space of perceived contrast, orientation, and size (assumed to be orthogonal within the local region of consideration), in which distance is proportional to discriminability. Each stimulus is a point in this space. A base stimulus, at the origin of this space, is defined in physical dimensions: -25% contrast, vertical orientation, 30 arc min major axis length, aspect ratio of 3. Now the other stimuli must be placed in this space. Ideally, they would be equidistant from one another to have equal discriminability, but in 3 dimensions, the maximum number of equidistant points is 4, and at least 8 are needed to accommodate the set size of 4 in the experiments. The solution was to form a cube in this psychological space with a stimulus at each corner of the cube, as shown in Figure 9. Thus, if k is the length of a side of the cube, out of the 28 pairs of stimuli, 12 pairs are a distance k apart, 12 pairs are a distance $\sqrt{2}k$ or approximately 1.4k apart, and 4 pairs are a distance $\sqrt{3}k$ or approximately 1.7k apart.

Next, the question must be answered: what is k? It must be some unit of discriminability, and for each dimension it will correspond to some physical increment from the base stimulus. Here, a physical increment that yielded 75% correct in a discriminability task is adopted as the unit. These 75% correct discriminability increment thresholds were measured for each observer along each of the three dimensions in three control experiments, one for each of the feature dimensions. These experiments are described in detail in Appendix B. On each trial, observers were presented with the

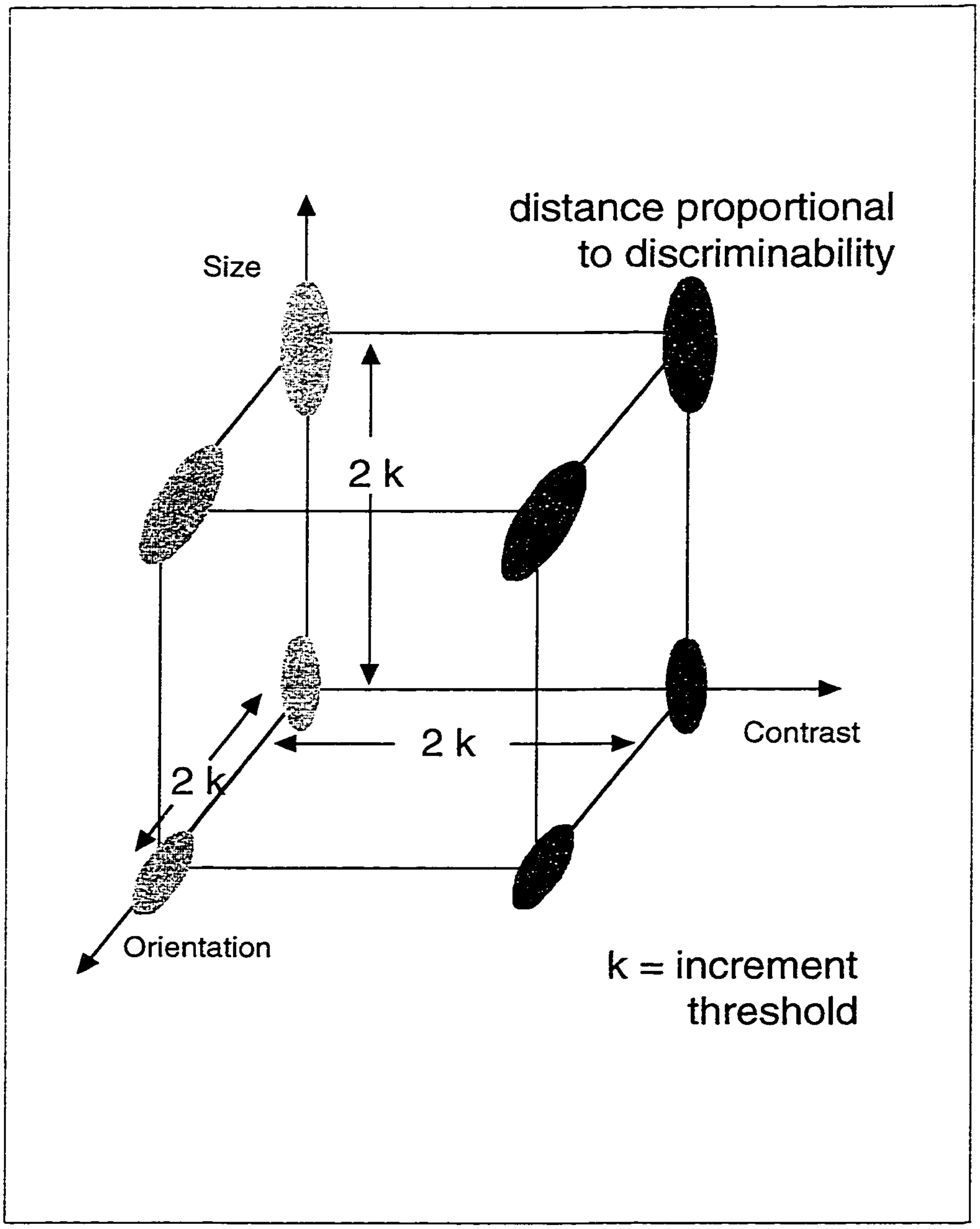

Figure 9: Stimuli Representation Distributions in Psychological Space

target/base stimulus followed by a study display that consisted of a single stimulus. They were to determine whether the target was present in or absent from the study display. On half of the trials the target was present in the study display. On the other half, the target was absent and instead a distractor that differed from the target in one feature was present.. The experiments were very similar to the Feature Search Experiment, but with set size 1 only. The size of the increment between the target and the distractor varied. The increment that yielded 75% correct based upon psychometric function fits to the data were estimated for each observer. The results were the following: From the base/target stimulus of -25% contrast, vertical orientation, and 30 arc min in major axis length, the mean estimated 75% correct discriminability increment thresholds were $-9.6 \pm 0.3\%$ contrast, 3.7 ± 0.1 degrees of rotation from the vertical orientation, and 4.0 ± 0.2 arc min of major axis length. These three threshold values are the physical increments that yield the same discriminability increment, k, from the base stimulus along each dimension and that define the cube in psychological space. The cube can then be expanded or contracted proportionally by using multiples of k to define the length of the side, thus increasing or decreasing the discriminability between stimuli for the purpose of increasing or decreasing performance levels to span the range from chance to perfect.

Representations

The issue of stimulus representation in the visual system has already been introduced above in the context of explaining how stimuli were chosen: In this theory, the stimulus is represented as a point in the 3-dimensional psychological space described where distance is proportional to discriminability. Presumably, as information is encoded and recoded through the information-processing system, there are a series of representations of the stimuli. What are examined in this study are the perceptual and memory representations and, in particular, how those representations are affected by the capacity limits, if present, of the perception and memory-encoding processes, respectively, that yield those representations.

The assumptions of standard multidimensional signal detection theory (Green & Swets, 1966) are used as a basis for this theory, an extension of the one-dimensional theory of Palmer et. al. (1993). Each stimulus engenders a representation. These representations are not perfect, that is, the representation has noise. In this case, the representation is a three-dimensional random variable, a point in the three-dimensional psychological space of Figure 9. The probability density functions of these random variables are centered at each of the vertices of the cube and for convenience, assumed to be normal.

At this point the capacity-limit concept is built into the theory. The variance of the representation distribution, in units of internal-noise variance which is assumed to be larger than zero, is a function of stimulus set size, n, and a capacity-limit parameter, a. Specifically,

$$var_r = n^a \bullet var_i$$

where var_r is representation variance and var_i is internal-noise variance. One of the reasons for selecting this function, other than that the power function is simple and familiar, is because it creates a capacity-limit measurement scale along which there are two particular and well-known landmarks: a=0 is the unlimited-capacity model and a=1 is the fixed-capacity model.

When a=0, n^a=1 and $var_r = var_i$. In this case, the representation is independent of set size: the mean and variance of the representation distribution are constants now, not functions of set size. Stimuli are processed independently, and hence the number of stimuli has no effect on the precision of each representation. Set-size effects may arise, however, not due to perception or memory encoding, but due to later processing, in particular, due to the integration of these noisy representations in the decision phase of processing. It is a model originally proposed by Tanner (1961) and developed and used more recently by Shaw (1980), Palmer, Ames, and Lindsey (1993), and others.

When a=1, n^a=n and $var_r = n \bullet var_i$. This is the value for variance that issues from the assumption that there is a fixed amount of processing capacity that must be divided up among n stimuli, an Information Theory concept used by Broadbent (1958; 1971), Lindsay, Taylor, and Forbes (1968), Shaw (sample-size model, 1980), and Palmer et. al. (1993). In this model, set-size effects arise from perception or memory encoding, as well as decision.

In general, when $a > 0$, representation variance is dependent upon set size in that as set size increases, variance also increases. This is the limited-capacity situation, fixed capacity being one instance of limited capacity. This increase in variance with set size is mediated by the value of a. As a gets larger, so does the magnitude of the increase in variance with set size. This is evident in Figure 10 where this variance gain factor function is plotted against set size for various values of a. As a approaches zero, the variance is affected by set size less and less. However, as a gets larger, the variance is affected by set size more and more. It is reasonable, therefore, to characterize a as a measure of the degree to which set size influences representation variance. Thus a is a measure of capacity limits.

It has been said that a=0 represents unlimited capacity and a>0 represents limited capacity, a=1 being a particular instance of limited capacity called fixed capacity. The other theoretical possibility is that $a < 0$. In this case, variance decreases with increasing set size, meaning that the precision of the representation actually improves the more stimuli are being processed simultaneously. This has been termed super-capacity (Townsend, 1983). The more stimuli that are being processed, the more capacity is available to process additional stimuli.

A second reason for selecting the power function of set size as the variance gain factor is that capacity-limit contributions from multiple processes combine in a simple way.

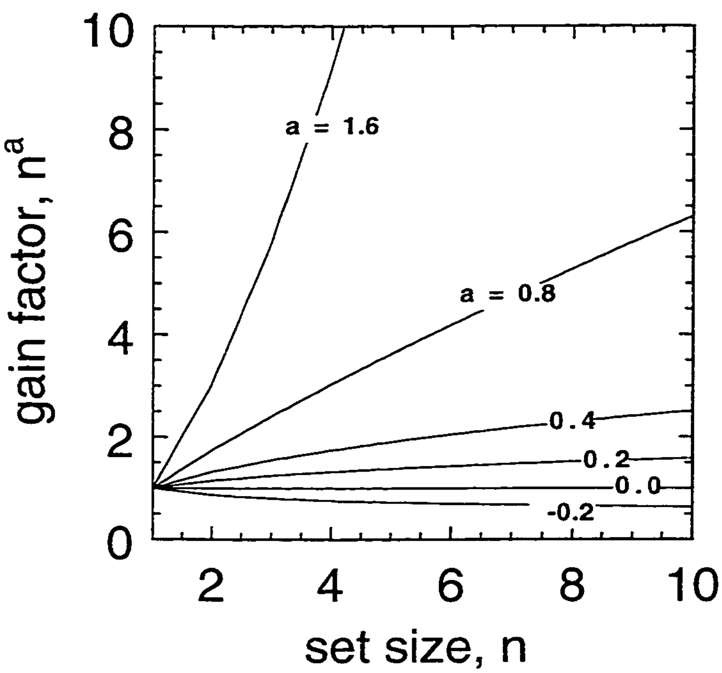

Figure 10: Variance Gain Factor

As stated above, the memory-encoding process is difficult. if not impossible, to isolate in an experiment. Perception, in particular, must occur prior to memory encoding and may be affected by set-size manipulations. Let *a1* be the capacity limit parameter corresponding to perception and let *a2* be the capacity limit parameter corresponding to memory encoding. Then

$$var_p = n^{a1} \bullet var_i$$

where var_p is the variance of the perceptual representation, and

$$var_m = n^{a2} \bullet var_p$$

where var_m is the variance of the memory representation. Thus

$$var_m = n^{a2} \bullet n^{a1} \bullet var_i$$

or

$$var_m = n^{a1+a2} \bullet var_i$$

Therefore, if the capacity-limit parameter corresponding to perception (*a1*) is measured in a search task, and the capacity-limit parameter corresponding to memory encoding plus perception (*a1+a2*) is measured in a memory task, then the capacity-limit parameter corresponding to memory encoding alone (*a2*) can be determined by simple subtraction.

Decision

Once the representations of the stimuli are formed, the decision must be made based upon the representations: Was the target present or absent? If there was no noise in the visual system, that is if $var_i = 0$, the decision would be straightforward. The representation would always be at one of the vertices of the cube in psychological space, because representation variance would be zero, and there would be no uncertainty about whether a stimulus was the target or not. The decision would be a simple mapping between perfect evidence and response selection: If there is a representation at the target vertex, then the target was present, select target present response. Note that in this case of no noise, the stimuli set size has no impact on representation variance or decision.

However, if there is noise in the visual system, that is, $var_i > 0$, which is what is assumed here as in all SDT-based theories and is justified by the facts that neural firing is variable and the physical stimulus is noisy, the decision is not so straightforward. The representation is imperfect, it can be anywhere in psychological space, not necessarily at a cube vertex. Each representation must be classified as a target or non-target, given imperfect evidence. The target could potentially engender a representation that is quite distant from the target vertex at the origin; similarly, a distractor could engender a representation that is quite near or even at the target vertex. In this theory, as in all SDT-based theories, it is assumed that a rule is followed for deciding whether a representation is of the target or not so that a response can then be selected.

The decision rule used in this theory is similar to one that has been used by Eckstein, Thomas, Palmer, and Shimozaki (1999) referred to in their paper as the max-min decision rule. It is implemented as follows. The stimulus representation has 3 psychological dimensions corresponding to the 3 physical dimensions of the stimulus. For convenience, in both the psychological and physical space, the target is centered at the origin. Thus the 3 physical dimensions are increment values in contrast, orientation, and size, and the 3 psychological dimensions correspond to these physical increments. A decision variable, V, is defined as the maximum value of the three psychological dimensions, that is, the greatest distance (in discriminability units) from the target of each of the three dimensions. Other alternatives are possible for this variable, for example, Eckstein et. al. (1999) compared this rule to another in which the decision variable was a linear combination of the values along the different dimensions, but the predictions were similar and the data fits did not favor one over the other. The rule chosen here is was selected because of its simplicity given the details of the task and stimuli.

A criterion is defined next. This criterion is a particular distance (i.e. degree of discriminability) from the target vertex in psychological space. If V exceeds the criterion, the representation is relatively distant from the origin and is classified as not having arisen from the target distribution centered at the origin in this space. If V is within the criterion, the representation is relatively close to the origin and is classified as having arisen from the target distribution. An illustration is provided in Figure 11 where (x,y,z) which corresponds to (orientation, contrast, size), is the representation of a stimulus and c is the criterion. In this example, $V = \max(x,y,z) = z > c$, hence the decision would be that this stimulus was not the target. In general, the decision must be made for each representation corresponding to each of the n stimuli. If any one of the representations is classified as a target, then the response selected is target present. If none of the representations is classified as a target, then the response selected is target absent. Note that there is a possibility of incorrect classification for each stimulus, hence as the set size increases, so will the error rate. Thus, the decision process will contribute to the set-size effect. In the case of the Postcue Memory experiment, only the representation engendered from the postcued stimulus need be classified and hence there is no decision process contribution to the set size effect.

Clearly, given this decision rule and the assumption of internal noise, errors are possible. Indeed, as k (discriminability) decreases, the representation distributions will overlap more and more, and the error rate will increase until chance performance is reached. On the other hand, as k increases, errors will decrease to zero which is perfect performance. Theoretical prediction curves that relate percent correct for set size 4 versus percent correct for set size 1 over the range of discriminability are mapped out in this way as will be described below. Because errors do occur in the experiments reported here and because errors have been shown to increase with decreasing discriminability, the validity of the noise assumption is supported.

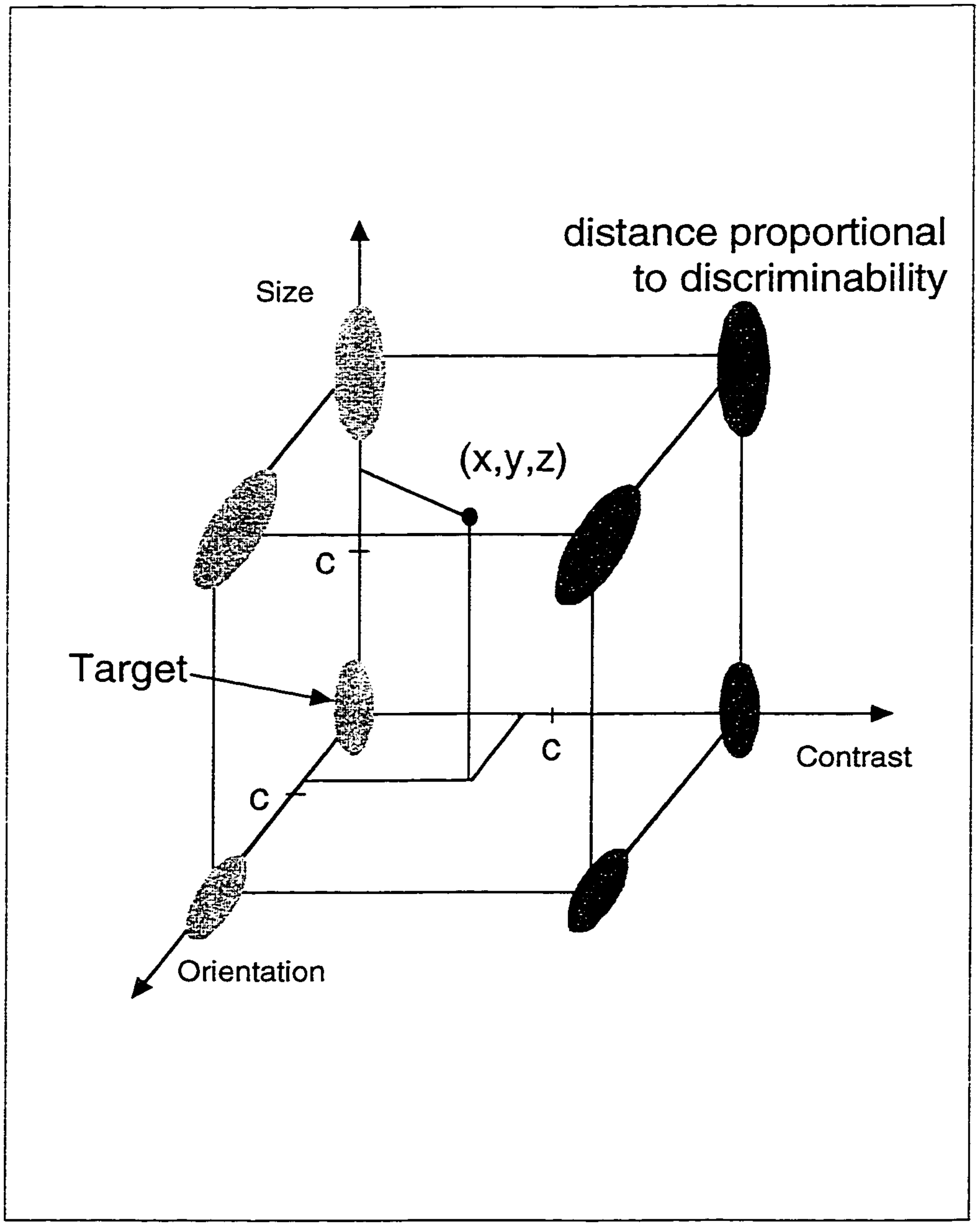

Figure 11: Decision Rule

Link to Observables

Next, the theory must predict performance. In the discussion that follows the general concepts are provided. Additional mathematical details are given in Appendix C.

First, consider the case in which only one stimulus representation is being evaluated. This occurs when set size is 1 or when a stimulus is postcued. Per Signal Detection Theory, there are four possible results of decision: a hit, a miss, a false alarm, or a correct rejection. A hit is when the target is correctly classified as the target and a miss is when the target is incorrectly classified as a distractor. A false alarm is when a distractor is incorrectly classified as a target and a correct rejection is when a distractor is correctly classified as a distractor. Using the decision rule described above, where V = max(x,y,z) is the decision variable and c is the criterion, the probabilities of a hit and a false alarm are the following

$$H = P\ (V_T < c)$$

$$FA = P\ (V_D < c\ )$$

where the subscript on V indicates whether the representation was engendered by a distractor (D) or the target (T). The probabilities of a miss and a correction rejection are just the complements of the hit and false alarm probabilities, respectively.

Next, consider the case in which n stimuli are being evaluated. The hit probability is the probability that the target is present and at least one of the n representations (it doesn't have to be the target representation) yields a V that falls within the criterion (i.e.,the complement of the probability that they all fall outside the criterion), resulting in a correct "target present" response. The false alarm probability is the probability that the target is not present but at least one of the n distractor representations yields a V that falls within the criterion, resulting in an incorrect "target present" response.

$$H = 1 - P\ (V_T > c,\ V_{D1} > c,\ V_{D2} > c,\ \ldots\ V_{Dn-1} > c)$$

$$FA = 1 - P\ (V_{D1} > c,\ V_{D2} > c,\ \ldots\ V_{Dn} > c)$$

Because the observers are trained to have equal bias with bias feedback, equal bias is assumed. Along with the assumption of normal distribution this implies that probability of a correct response is equal to the hit probability which is equal to the complement of the false alarm probability

$$PC = H = 1 - FA$$

This allows for the criterion value, c, to be determined.

At this point, prediction curves that relate probability correct for set size 4 and probability correct for set size 1 can be generated for each value of a as follows: a, n, and k values define the variance and means of the representation distributions. Then c is determined by solving H = 1 - FA. Finally PC = H. This is done for each set size, n. This yields a point on the probability correct set size 4 vs. set size 1 plot. That point traces out a curve as k is varied. This is repeated for various values of a, yielding a set of curves.

There are subtle differences in the implementation of this theory for each of the different experiments due to differences in stimuli that are used and how those stimuli are sampled on a given trial. For the Feature Search Experiment, the distractors are homogeneous and the target and distractor vary only along one dimension at a time (the three types of distractors are blocked), thus the representation of a stimulus can be reduced to a one-dimensional random variable. The decision variable, V, is then just equal to the value of the random variable: the maximum of a single value is just that value. In the Postcue Memory Experiment, the distractors are heterogeneous, differ from the target in three dimensions, are sampled randomly with replacement from the full set of 8 stimuli, and there is a postcue which simplifies the decision process. In the Combined Search and Memory Experiment to be reported below, the distractors are heterogeneous, differ from the target in three dimensions, are sampled randomly without replacement from the non-target stimuli, and there is no postcue. These differences result in different sets of prediction curves for the different experiments. Modeling details are given in Appendix C.

Prediction curves for the Feature Search Experiment and Postcue Memory Experiment are shown in Figures 12 and 13, respectively. They were generated using the Mathematica computer software package. In these figures, curves corresponding to *a*-values of 0, 0.25, 0.5, 0.75, and 1 are shown with probability correct for set size 4 on the ordinate and probability correct for set size 1 on the abscissa, just as in the data Figures 4 and 8. Each curve is mapped out by varying the target-distractor discriminability (k) and hence allows for the full range of difficulty levels. The identity line, where probability correct for set size 4 is equal to probability correct for set size 1, is also shown for reference. In the Feature Search predictions, all of the curves shown fall below the identity line, indicating set-size effects: probability correct for set size 4 is always less than probability correct for set size 1 with these *a* values. The set-size effects increase as *a* increases, as evidenced by the curves falling farther below the identity line as *a* increases. A similar pattern emerges with the Postcue Memory predictions, except that the curve corresponding to a=0 coincides with the identity line. The a=0, or unlimited-capacity, prediction is that there is no set-size effect. This is because only the postcued stimulus representation is evaluated, and if that stimulus representation is unaffected by the presence of other stimuli (the unlimited-capacity situation) then there should be no set-size effect.

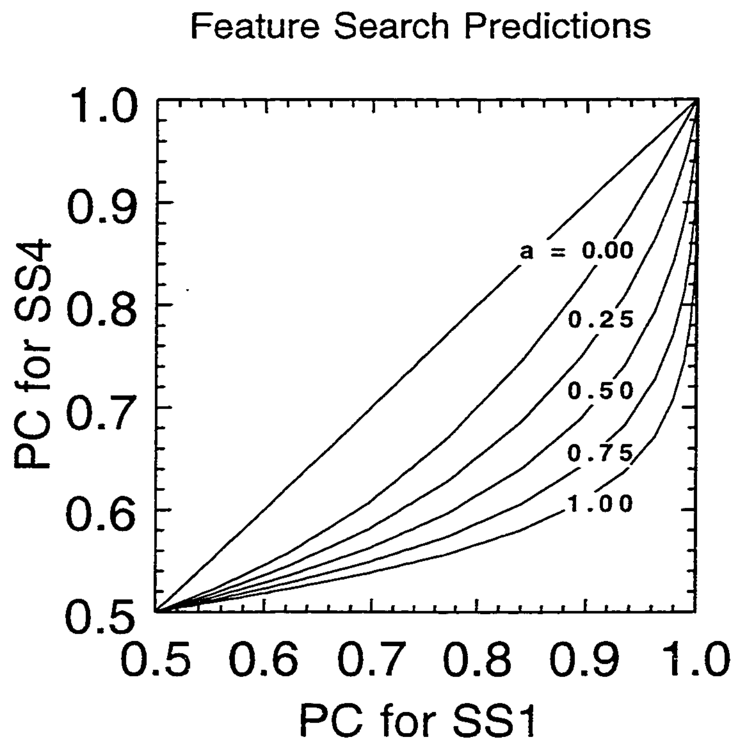

Figure 12: Feature Search Predictions

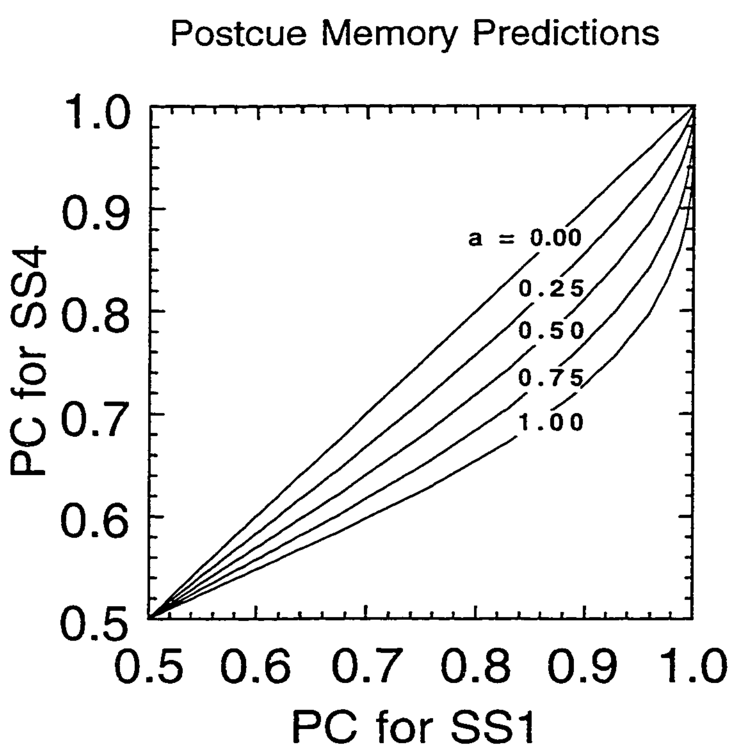

Figure 13: Postcue Memory Predictions

EXPERIMENTS — PART 2

In this section, the results from experiments reported in Part 1 are revisited in light of the theoretical context described above, and then the final experiment is reported. The purpose of this final experiment was to combine matched search and memory tasks into a single experiment in order to directly measure and compare capacity limits in perception and memory encoding.

Feature Search Experiment, Revisited

In this experiment, as reported in Part 1, observers search a briefly displayed stimuli set for a target stimulus and then indicate whether the target was present in or absent from the set. The Method and Results were described in Part 1. At this point more elucidation can be given regarding the choice of stimuli. Also, theoretical predictions and fits to the data will be described.

Stimuli

The stimuli, described in Part 1, were defined based upon the 75% correct discriminability increment thresholds measured in the control experiments described in the Theory section above and reported in detail in Appendix B. Because the threshold measurements were obtained from essentially a set size 1 feature search experiment, using the 75% threshold to define the stimuli in this experiment would yield approximately 75% correct performance for the set size 1 condition, but lower performance for the set size 4 condition. It was necessary to increase the discriminability of the stimuli to keep performance in the middle of the range. The threshold was multiplied by 1.5 to define the stimuli to ensure that set size 4 performance was off the floor (well above 50% correct) while the set size 1 performance was below the ceiling (well below 100% correct).

Theoretical predictions

Theoretical predictions, shown in Figure 12, were generated for this experiment; The mathematical details are given in Appendix C. If perception is unlimited-capacity, then the data should fall somewhere on the a=0 curve. If perception is limited-capacity, it should fall somewhere in the plot below the a=0 curve. In the extreme, if perception is fixed-capacity, then the data should fall somewhere on the a=1 curve. The value of a corresponding to the data will be some measure of degree of limited capacity.

Analysis

For each observer, the data point, represented as percent correct for set size 4 versus percent correct for set size 1, was transformed to an a value and a delta value. The a value represents which prediction curve the point is on. The delta value is proportional to k and represents where on the curve the point is located.

Results

The mean a over observers was 0.22 ± 0.03. The mean delta was 2.08 ± 0.08. The data, originally presented in Figure 4, is reproduced in Figure 14, along with the a=0 and a=1 prediction curves and the curve corresponding to the mean a value.

Discussion

At this point, having provided the theoretical context, the data can be interpreted in terms of capacity limits. The a value, 0.22 ± 0.03, is a measure of capacity limits in perception. The value is not consistent with a=0, unlimited-capacity perception. Nor is the value consistent with a=1, fixed-capacity perception. The value is greater than zero, hence it can be concluded, given this model, that perception is limited in capacity. However, to the extent that distance on this scale is proportional to degree of limited capacity, the degree of limited capacity is small. The value is much closer to 0 than to 1. Apparently, perceptual representation variance does not vary much with set size.

Postcue Memory Experiment, Revisited

In this experiment, as reported in Part 1, observers view a briefly displayed stimulus set followed by a target stimulus, acting as a memory probe, and a location cue. Observers indicate whether or not the target appeared at the cued location in the stimulus set

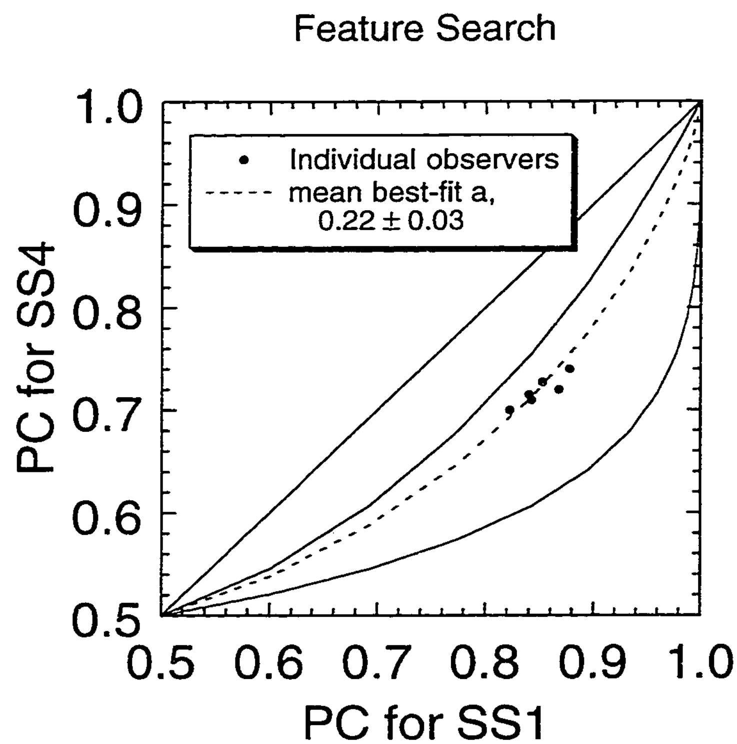

Figure 14: Feature Search Results Revisited

display. The Method and Results were described in Part 1. At this point more elucidation can be given regarding the choice of stimuli. Also, theoretical predictions and fits to the data will be described in this section.

Stimuli

The stimuli, described in Part 1, were defined based upon the 75% correct discriminability increment thresholds measured in the control experiments described in the Theory section above and reported in detail in Appendix B. It was necessary to increase the discriminability of the stimuli beyond the 75% correct increment threshold of the control experiment to keep performance in the middle of the range. The estimated increment threshold was multiplied by 2.0 to define the stimuli separation distance. This ensured that set size 4 performance was off the floor (well above 50% correct) while the set size 1 performance was below the ceiling (well below 100% correct).

Theoretical predictions

Theoretical predictions, shown in Figure 13, were generated for this experiment; The mathematical details are given in Appendix C. If memory encoding is unlimited-capacity, then the data should fall somewhere on the a=0 curve. If memory encoding is fixed-capacity, then the data should fall somewhere on the a=1 curve. If memory encoding is limited-capacity, it should fall somewhere in the plot below the a=0 curve. The value of a corresponding to the data will be some measure of degree of limited capacity.

Analysis

For each observer, the data point, represented as percent correct for set size 4 versus percent correct for set size 1, was transformed to an a value and a delta value. The a value represents which prediction curve the point is on. The delta value is proportional to k and represents where on the curve the point is located.

Results

The mean a over observers was 0.93 ± 0.16. The mean delta was 2.52 ± 0.10. The data, originally presented in Figure 8, is reproduced in Figure 15, along with the a=0 and a=1 prediction curves and the curve corresponding to the mean a value. The mean a value is not significantly different from 1 ($t(5) = 0.42$, $p > 0.4$).

Discussion

At this point, having provided the theoretical context, the data can be interpreted in terms of capacity limits. The a value, 0.93 ± 0.16, is a measure of capacity limits in memory encoding plus perception (recall that perceptual processes could not be controlled with respect to set-size variation). The value is not consistent with a=0, unlimited-capacity. The value is consistent with a=1, fixed-capacity. Apparently, memory representation variance does vary quite a bit with set size. Indeed, in units of internal noise variance, it is nearly equal to set size.

Feature Search and Postcue Memory Results Summary

The purpose of the Feature Search and Postcue Memory Experiments was to use prototypical tasks, much like those used by other researchers, to investigate capacity limits in perception and memory-encoding. The idea was to replicate the kinds of results found by other researchers in these two areas of investigation.

The data patterns found in feature search are not in dispute, and the pattern found here was consistent with what has been found previously. Set-size effects are present when target-distractor discriminability is sufficiently low. What is in dispute is the interpretation of the data. Some, such as Treisman and Wolfe, have used a limited-capacity perception model to account for the results whereas others, such as Palmer and Eckstein, have accounted for them with an unlimited-capacity perception model. The interpretation of the present results using a model very similar to the SDT-based models used by Palmer and by Eckstein is that perception is limited in capacity, but not very much. The results are closer

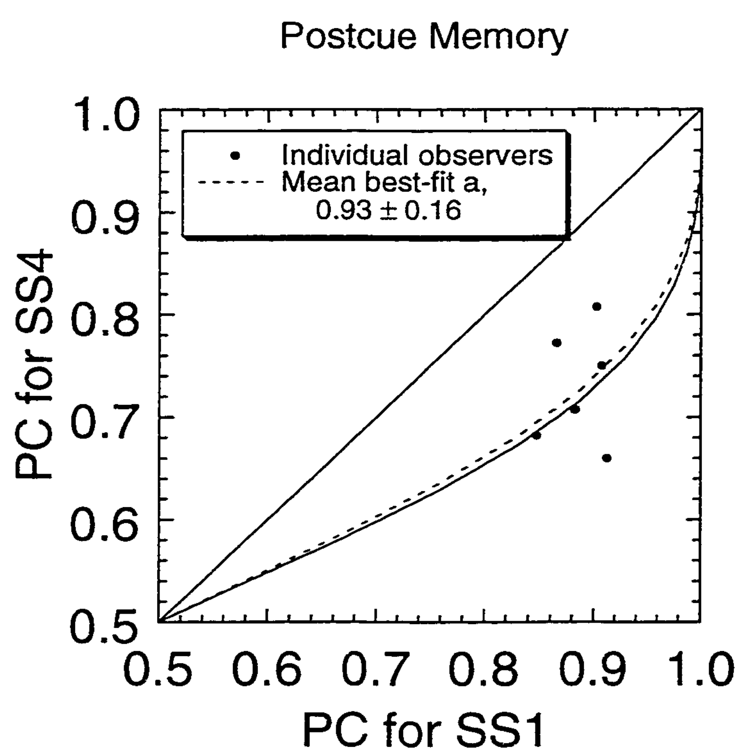

Figure 15: Postcue Memory Results Revisited

to being consistent with an unlimited-capacity model than a fixed-capacity model. The Postcue Memory Experiment results are consistent with previous research which has found memory-encoding plus perception to be fixed-capacity. The mean *a* values from these first two experiments are shown in Figure 16. The black-filled circle represents the search result. The open circle represents the memory result. Standard error bars are shown, but in the search result are smaller than the radius of the circle.

A major goal of this study is to compare capacity limits in perception and memory encoding. However, there are several factors that preclude making this comparison based upon the Feature Search and Postcue Memory experimental results. First, capacity limits in perception and memory encoding plus perception were measured in separate experiments. Also, the two tasks had some important differences such as different decision demands, different stimuli, and different stimulus sampling procedures. These differences were hopefully accounted for theoretically as shown in the different set of predictions and described in Appendix C, however, it would be better to have tasks that were more similar and to have the two tasks in one experiment so that capacity limits in perception and memory encoding plus perception could be compared directly and the perception contribution to the capacity limit would be the same in both. Thus a final experiment is reported below in which these controls are implemented.

<u>Combined Search and Memory Experiment</u>

This final experiment was designed to measure and directly compare capacity limits in perception and memory encoding. A search and a memory task were designed to be as similar as possible and combined into a single experiment with appropriate counterbalancing. Indeed the tasks were identical except that the target preceded the stimuli set display in the search task but it followed the stimuli set display in the memory task. This experiment is modeled after a similar experiment using letters as stimuli by McLean et. al. (1998) that was itself designed as a combination of prototypical search and memory tasks.

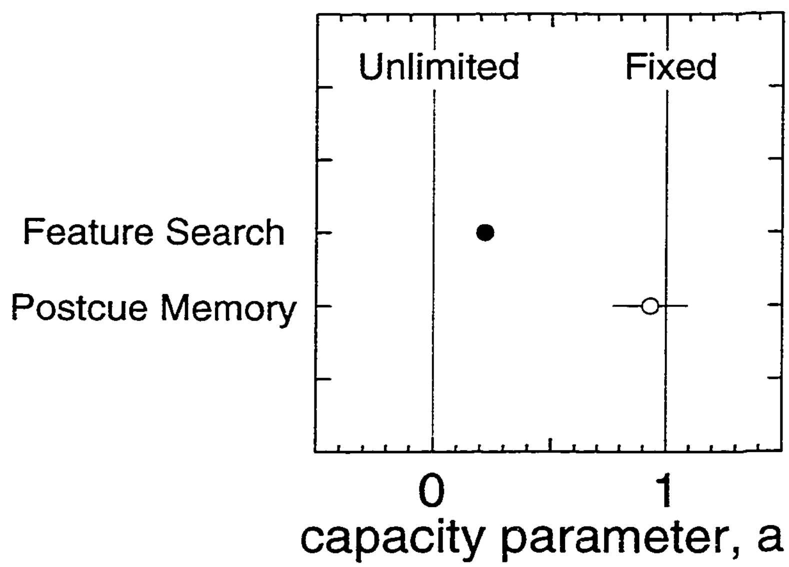

Figure 16: Summary of Feature Search and Postcue Memory Results

Method

The methodology of the Combined Search and Memory Experiment was the same as the Feature Search and Postcue Memory Experiments with the following exceptions. The main difference was that search and memory tasks were combined into one experiment. Other differences between the search task in the Combined Experiment and the Feature Search task as well as between the memory task in the Combined Experiment and the Postcue Memory task were necessary to make the search and memory tasks as similar as possible in the Combined Experiment: The distractors were heterogeneous non-targets, selected randomly without replacement in both the search and memory tasks of the Combined Experiment, and there was no postcue. These and other details of the exceptions are described below.

Observers. The same 6 observers participated in this experiment as in the Feature Search and Postcue Memory Experiments and the author also participated.

Stimuli. The stimuli were the same 8 ellipses used in the Postcue Memory Experiment. They are illustrated schematically in Figure 17. As in the Postcue Memory Experiment, the target stimulus varied from trial to trial and the distractors were heterogeneous. However, unlike the previous experiments, distractors were sampled randomly without replacement from the 7 non-target stimuli. Hence there were no repeated stimuli in the display set. Sample displays for the two set sizes (1 and 4) are shown in Figure 18.

Design and Procedure. As shown on the left side of Figure 19, the display sequence on a search trial was the following: (a) the target ellipse was presented at the center of the display for 500 ms. (b) A blank screen was presented for 500 ms. (c) A plus-sign fixation point was presented at the center of the display for 500 ms. (d) The fixation point was joined by the study display — 1 or 4 ellipses — presented for 80 ms. (e) A blank screen was presented for 500 ms. (f) A question mark was presented as a response prompt.

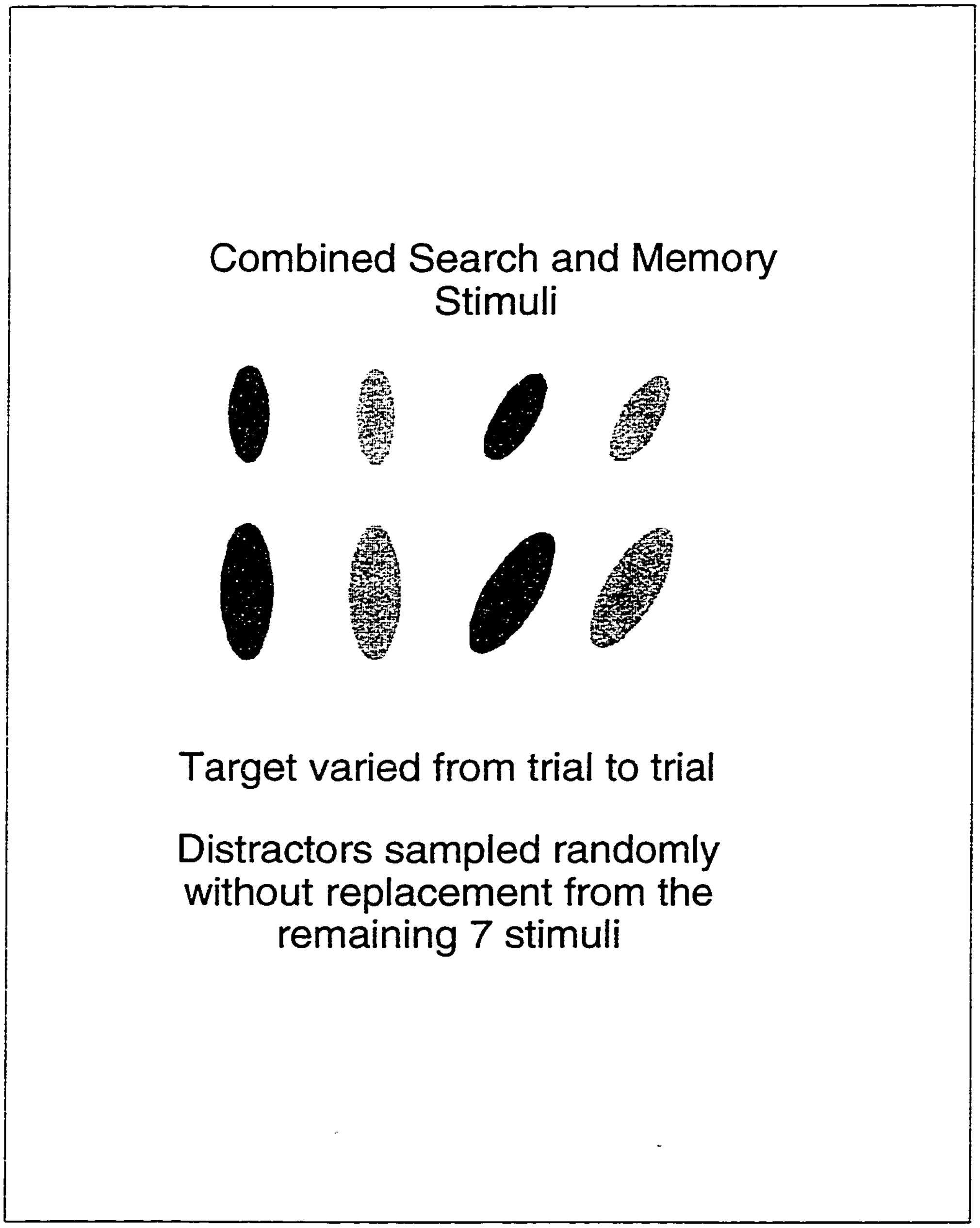

Figure 17: Combined Search and Memory Stimuli

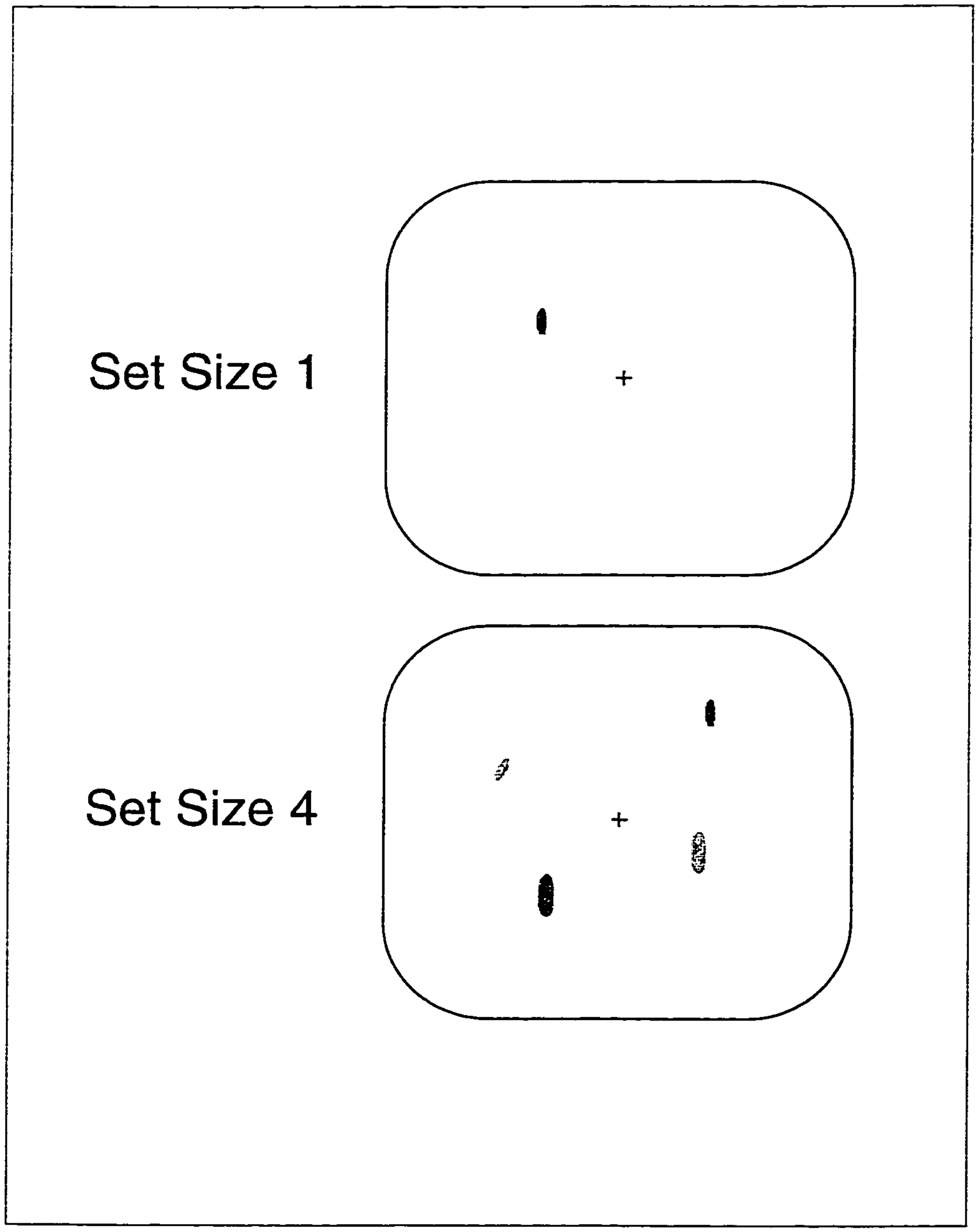

Figure 18: Combined Search and Memory Sample Displays

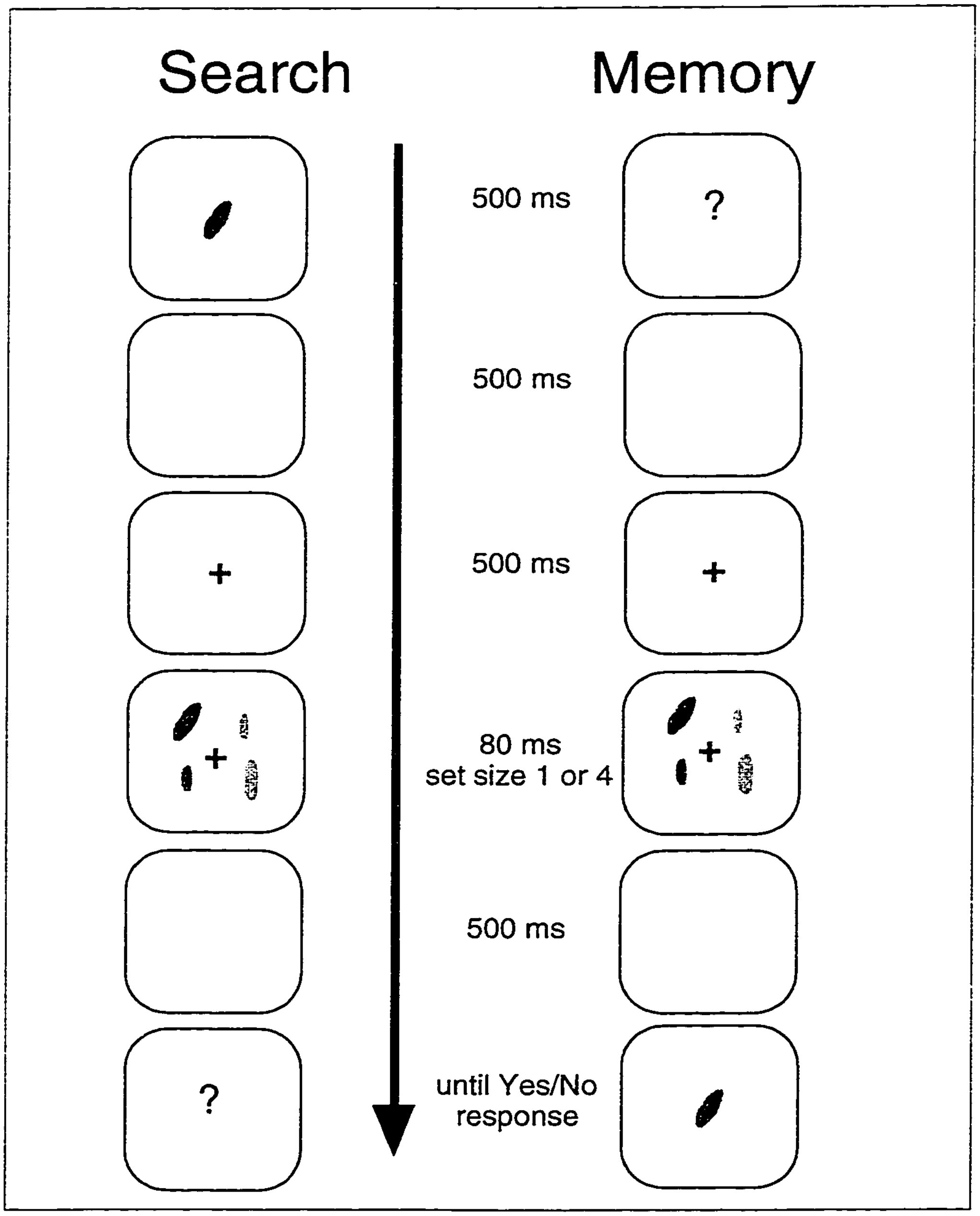

Figure 19: Combined Search and Memory Tasks

As shown on the right side of Figure 19, the display sequence on a memory trial was the same as the search trial except that the target was presented at the end rather than at the beginning of the sequence. The display sequence was the following on the memory trial: (a) a question mark was presented at the center of the display for 500 ms. (b) A blank screen was presented for 500 ms. (c) A plus-sign fixation point was presented at the center of the display for 500 ms. (d) The fixation point was joined by the study display — 1 or 4 ellipses — presented for 80 ms. (e) A blank screen was presented for 500 ms. (f) The target stimulus was presented at the center of the display as a prompt for response. There was no postcue.

As in the previous experiment, following the display sequence, the observer responded by pressing the "p" key to indicate "target was *present* (anywhere in the display set)" or the "a" key to indicate "target was *absent*." The correct response in the Figure 19 example is "p".

Trials were presented in blocks of 32, not including 4 practice trials at the beginning of each block. A session consisted of 10 blocks. Each observer participated in 8 sessions, for a total of 2560 trials. Randomization and counterbalancing were the same as the Postcue Memory experiment with the addition that task varied between blocks. Each of the 5 consecutive pairs of blocks within a 10-block session consisted of a search task block and a memory task block in random order within that pair.

<u>Theoretical predictions</u>

Based on the theory outlined above, prediction curves were generated for this task. Curves corresponding to various a-values are shown in Figure 20. The same prediction curves apply to both the search and memory tasks.

<u>Results</u>

For search, mean percent correct for set size 1 was 88.9 ± 1.1 and for set size 4 was 73.3 ± 1.0; the mean difference was 15.5 ± 0.8. For memory, mean percent correct

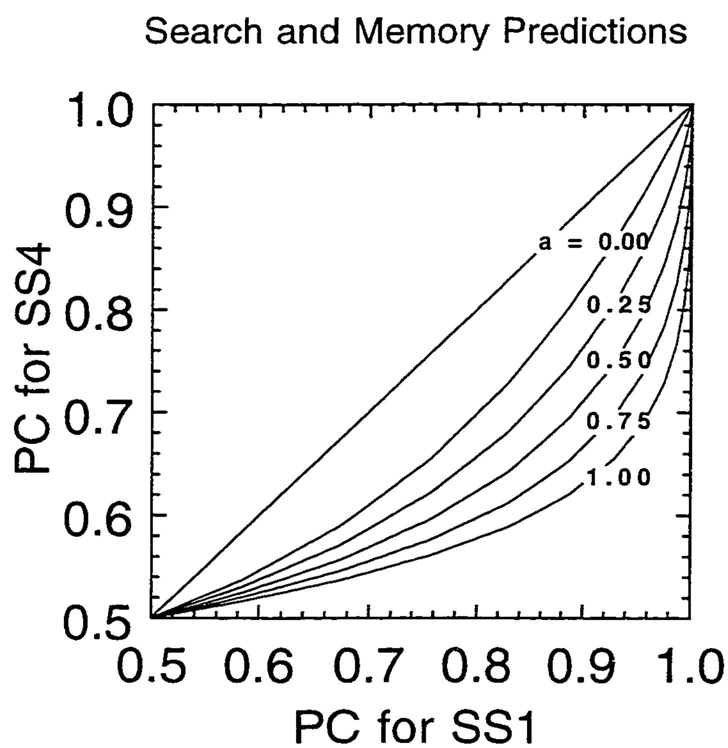

Figure 20: Combined Search and Memory Predictions

for set size 1 was 87.3 ± 0.9 and for set size 4 was 63.3 ± 1.2; the mean difference was 24.0 ± 1.2. The data are shown in Figure 21. For each task, percent correct for set size 4 is plotted against percent correct for set size 1. This yields a scatterplot of 7 data points for search (plotted using black-filled circles) and 7 data points for memory (plotted using black-filled squares), one of each for each observer. There was an effect of set size: percent correct for set size 4 was always lower than for set size 1 as indicated by the individual data points all falling below the identity line (included in Figure 21 for reference).

For each observer in each task (i.e. each data point) the percent correct results for set sizes 1 and 4 were converted to a and delta values. The mean a value for the search task was 0.31 ± 0.04 and for the memory task was 0.83 ± 0.10, a significant difference of 0.51 ± 0.08 ($t(6) = 6.06$, $p < .001$). The curves corresponding to these mean a-values are shown in Figure 21 with dashed lines. The a=0 and a=1 curves are also shown for reference. Mean delta values were 2.55 ± 0.11 for search and 2.40 ± 0.09 for memory.

Next, the search and memory task results are used to make interpretations about the capacity limits of perception and memory encoding. Because the search task was designed to measure perceptual processing capacity, the a value corresponding to perception is the same as the a value for search: 0.31 ± 0.04. This a value for perception is closer to 0 than to 1, though not consistent with a=0, unlimited-capacity perception.

The memory task measurement, however, includes capacity limits for the joint processes of perception and memory encoding. The a value for the memory task is 0.83 ± 0.10 and is consistent with a=1, fixed-capacity ($t(6) = 1.7$, $p > 0.1$). The a value for memory encoding alone is calculated by subtracting out the contribution due to perception. The basis for this calculation is given above in the Theory section. Because the search and memory tasks had identical study displays, it is reasonable to assume that the contribution due to perception is the same in both tasks. The perceptual capacity limit as measured with the search task is $a = 0.31 \pm 0.08$. When this amount is subtracted from the memory task a

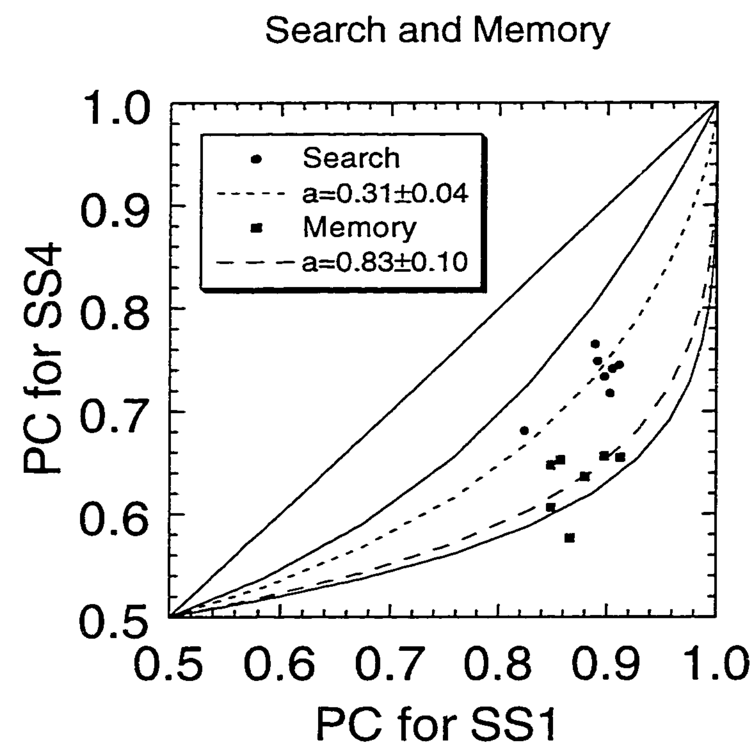

Figure 21: Combined Search and Memory Results

value, the result is the estimated a value for the memory encoding process alone: 0.51 ± 0.08. This value falls in the middle of the 0 to 1 range, not consistent with unlimited capacity at 0 nor with fixed capacity at 1. Hence by this measure, memory encoding is limited capacity, but not fixed capacity, yet the limit is significantly larger than the capacity limit of perception. The mean difference between the capacity limit parameters for memory encoding alone and perception alone is 0.20 ± 0.09 ($t(6) = 2.22$, $p < 0.05$).

Discussion

The results of this Combined Search and Memory Experiment provide the opportunity to compare capacity limits in perception and memory encoding. The two tasks were designed to isolate these processes and were matched and included within a single experiment along with well-defined stimuli and a unified theoretical context. The results are consistent with capacity limits in perceptual processing, the unlimited-capacity model prediction does not fit the data. However, the limits in perception are relatively small as measured on the a-scale — the a-values are certainly closer to 0 than to 1. In the case of memory encoding plus perception, it is a different story. The capacity limits are larger, closer to 1 than to 0, in fact, consistent with the a=1, fixed-capacity model, predictions. However, when the effect of perception is excluded, the memory encoding capacity limit parameter is a value midway between 0 and 1, indicating limited capacity memory encoding, and a limit that is larger than that of perception.

GENERAL DISCUSSION

A summary of the results from the Feature Search, the Postcue Memory, and the Combined Search and Memory Experiments are shown in Figure 22. The mean a values with standard error bars for each task are plotted in black-filled circles for search and open circles for memory. The a value corresponding to memory encoding alone based upon the Combined Search and Memory Experiment is the difference between the a values for the search and memory tasks in that experiment. This plot illustrates the major contribution of this study which was the creation of a theoretical context in which to measure and compare capacity limits in different processes using a single scale. At the grossest level of analysis, both perception and memory encoding were found to be limited-capacity on this theoretical a value scale because the estimated a values for the two processes were both greater than zero. However, because of the unified theoretical context, a more detailed level of analysis was possible. In particular, it was found that the capacity limits were greater for memory encoding than for perception. The perception a values leaned towards but were not consistent with the a=0, unlimited-capacity, account. The memory encoding plus perception a-values were consistent with the a=1, fixed-capacity, account.

Implications for Perception Research

It should be noted that in order to match the search task and the memory task study displays in the Combined Search and Memory Experiment, it was necessary to have heterogeneous distractors. This turned the search task into a kind of conjunction search, as compared to the feature search task of the first experiment. A qualitative distinction between the capacity limits of the processes involved in feature search as opposed to conjunction search was the basis for the creation of the two-stage limited-capacity perception models of Treisman, Wolfe, and others. Hence, these results provide an opportunity to evaluate their assumption.

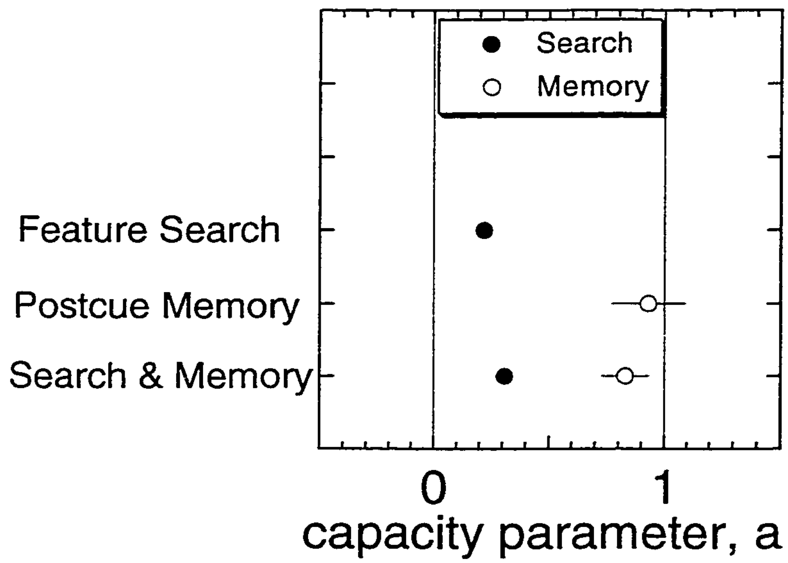

Figure 22: Results Summary

Although it is not the focus of this project, it is possible to informally compare the capacity-limit estimates of feature search and conjunction search because the a value is independent of target-distractor discriminability and the other differences between the two tasks are accounted for in the model. The a value for feature search as measured in the Feature Search Experiment was 0.22 ± 0.03 and for conjunction search as measured in the Combined Search and Memory Experiment was 0.31 ± 0.08, as shown in Figure 22. These two values are sufficiently similar to suggest that it is not necessary to assume a qualitative distinction in mechanisms at work in the two tasks such as an unlimited-capacity parallel process for feature search and limited-capacity serial process for conjunction search. Eckstein and his colleagues verified this conclusion in a set of studies focused upon this very issue. He showed that a SDT-based model with unlimited-capacity perception predicts set-size effects in visual search accuracy not only for feature search, as also shown by Palmer (1994), but also for conjunction, triple-conjunction, and disjunction search. (Eckstein, 1998; Eckstein, Thomas, Palmer, & Shimozaki, 1999).

The estimates of capacity limits in perception based upon the Feature Search Experiment results and the conjunction search task results of the Combined Search and Memory Experiment indicate that perception does have a capacity limit, albeit a relatively small one. This result does not confirm the unlimited-capacity perception hypotheses put forth by Palmer, Eckstein, and others. However, it is unlikely that it supports the specific limited-capacity hypothesis of Triesman and Wolfe. It is difficult to determine what kind of a value their models would predict, but perhaps it would be close to the a=1, fixed-capacity value. A serial process such as in their models is similar to fixed capacity because processing occurs for a fixed amount of time that is based upon the exposure duration of the stimulus, and the processing time is, in effect, divided up among the stimuli. This idea is similar to the fixed-capacity memory-encoding model of Shibuya and Bundesen (1988).

The difference between the conclusions of Palmer and Eckstein (no capacity limit in perception) and those of the present study (small capacity limit in perception) have several possible explanations, the most notable of which are the differences in task and stimuli. While these differences are accounted for in the models which would allow for a comparison of a basic perceptual process that is independent of these differences, it is possible that no such independent process exists. Instead it may be dependent upon task and stimuli. The mixed results in the simultaneous vs. successive paradigm may be explained by such a stimulus dependency. The successive advantage was not found for simpler letter stimuli but was found for more complex word stimuli or harder-to-discriminate letter stimuli. The stimuli used in the present study, ellipses that varied along three dimensions, could be considered more complex than the 2-dimensional ellipses modeled by Eckstein et. al. (1999) or the simple feature search stimuli used by Palmer (1994). Indeed Palmer (1994) found larger set-size effects in more complex tasks than what was predicted by an unlimited-capacity model. An example of such a task was his point orientation task. The stimuli were pairs of 1 degree separated points. The target was defined by rotating the two points around one another while maintaining a constant separation distance. The distractor pairs were horizontally oriented. Relational information was necessary to accomplish the task, not just stimulus attribute information.

Furthermore, even if there were no task or stimulus dependencies, and a basic perceptual capacity limit did exist, the different estimates of this limit may not be correct because they are based upon theoretical assumptions that may not be correct. In the present theory these assumptions include normal representation distributions with equal variances, a maximum value decision variable, and orthogonality of stimulus dimensions. While these assumptions provided a reasonable place to start, it remains to be seen how varying these assumptions will affect capacity limit estimates. If the accuracy of these assumptions can be tested and improved, it may be the case that the improved theory would yield unlimited capacity perception estimates to match those of Palmer and Eckstein.

Implications for Memory-Encoding Research

The results of the present study indicate that memory-encoding is indeed quite limited in capacity which is in agreement with the bulk of previous research on the topic. The evidence for unlimited-capacity memory encoding was confined mostly to the results from dual task experiments in two modalities. It is possible that different modalities have separate perceptual and memory systems which would explain why capacity loads placed on one would not influence the capacity load of another. More specifically, and within the context of the kinds of SDT-based model presented here, while visual-stimulus representation precision may be affected by the presence of multiple stimuli within the visual field, it may not be affected by the presence of stimuli in another modality or the processing of those stimuli in the second modality. The focus of the present work is upon the capacity limits of memory encoding within the visual system. Hence these two-modality, dual-task results are not directly related to the issue at hand.

The other piece of evidence found in support of unlimited-capacity memory encoding was from a coarse localization task (McLean, et. al., 1998). In this localization task, the target was always present, and the observer decided whether it appeared on the right or left side of the display. It was already mentioned that one of the purposes of the present study was to improve upon the design of the McLean, et. al. study. Differences in results could be attributed simply to these improvements, in which case the results of the present study would be the more convincing. For example, in the localization task, if the observer adopted the strategy of attending to only one side of the display, the task reduced to a yes-no task with set sizes 1 and 2 rather than 2 and 4. Specifically, if the observer only attended to the right side of the display and if the target was not present among the 1 or 2 stimuli on that side, the correct response would be that the target had appeared on the left.

Thus the change in performance due to a set size change from 2 to 4 of the whole display would actually be the change produced in going from set size 1 to 2.

This strategy was discovered early on in the study and observers were instructed to attend to the entire display. Thus, it is possible that the localization result of McLean et. al. was accurate and the difference in results was due to real differences in the task. In particular, the kind of information needed to perform the yes-no task was different than that which was needed for the localization task. For the former, identity of the stimuli was needed, but in the latter, location information was the key. It is possible that separate systems are responsible for these two kinds of information, the so-called "what" and "where" pathways (parvocellular and magnocellular systems, e.g. Kandel, Schwartz, & Jessell, 1991, p. 446-447), and perhaps these two different pathways have different capacity limits. Some researchers have suggested that location information is encoded obligatorily by special-purpose processors that are "independent of attention" (i.e. unlimited capacity) based upon evidence that location is encoded under incidental learning conditions (for a discussion see Logan, 1998).

In the same study by McLean et. al. (1998), a postcue yes-no memory task yielded results quite consistent with the result of the Postcue Memory Experiment of the present study in spite of the differences in stimuli, controls, and modeling. Perception plus memory encoding was found in both cases to be limited in capacity, and consistent with the fixed capacity model prediction, confirming the common hypothesis supported by other research (e.g. Palmer, 1990; Shibuya & Bundesen, 1988).

Implications for Perception and Memory Encoding Comparisons

Of the research described in the introduction having to do with the comparison of perception and memory-encoding capacity, the consensus seemed to be that perception is unlimited capacity and memory encoding is fixed capacity with the exception of the results from McLean et. al. (1998). Leaving aside the localization results which were discussed in detail above, the yes-no task results of McLean et. al. indicated a relatively small capacity limit for perception, and adding the memory encoding component only increased the capacity limit estimate by a slight amount. In agreement with McLean et. al., the present study did not confirm the previous consensus on unlimited-capacity perception, but it did find the limit to be relatively small. Contrary to the result of McLean et. al., the present study found that adding the memory encoding component to the task increased the capacity limit estimate substantially, resulting in a fixed-capacity conclusion about the joint processes of perception and memory encoding — a result that does support the previous consensus.

It is important to note here that the basic result of larger set-size effects for memory than for search in the combined experiment do not depend upon the assumptions of this theory. Also, while no strong interpretation of the absolute magnitude of a particular a value can be made, the relative change in a values between the search and memory tasks in the combined experiment is meaningful and less dependent upon theoretical assumptions.

Measures of Capacity

One of the goals of the present study was to propose a capacity limit scale. Much of previous research has focused upon two possibilities: a process is either unlimited capacity or fixed capacity. However, it seems likely because of the continuous nature of physical systems, that there would be a whole spectrum of possibilities within which unlimited capacity and fixed capacity are just two points. It seems likely that capacity limits exist in each process of the visual system, the issue is really, what is the size or degree of those limits.

The a-scale presented here and is one way of measuring capacity limits. The a value represents the degree to which the precision of a stimulus representation is affected by the presence of other stimuli. The larger the a value, the more variable the representation becomes with increasing set size. The smaller the a value, the less effect has set size upon representation variability. The a measure has the additional advantage of representing the joint effect of multiple processes in a simple way. If processes occur in serial, the output of one becoming the input of another such as the memory-encoding process acting upon the perceptual representation to create the memory representation, the loss of precision of representation becomes a positive increment on the a scale. The a value of multiple serial processes is just the sum of the a values for each of the processes alone.

CONCLUSION

In conclusion, the present study provided a theoretical context for measuring and comparing capacity limits in visual perception and memory encoding. Perception appeared to have a relatively small capacity limit, whereas memory encoding appeared to have a sharper capacity limit. Together, perception and memory encoding were consistent with fixed-capacity processing. A direct comparison between a perception and a memory-encoding task with identical study displays showed that set-size effects were smaller for perception than for memory encoding.

LIST OF REFERENCES

Anderson, C.M.B. & Craik, F.I. (1974). The effect of a concurrent task on recall from primary memory.*Journal of Verbal Learning and Verbal Behavior, 13,* 107-113.

Averbach, E. & Sperling, G. (1968). Short-term storage of information in vision. In Haber, R.N. (Ed.), *Contemporary Theory and Research in Visual Perception.* (pp. 202-214). New York: Holt, Rinehart, and Winston, Inc.

Bergen & Julesz (1983). Rapid discrimination of visual patterns. *IEEE Transactions on Systems, Man, and Cybernetics, 13*, 857-863.

Boronat, C.B. & Logan, G.D. (1997). The role of attention in automatization: Does attention operate at encoding, or retrieval, or both? *Memory and Cognition, 25,* 36-46.

Brand, J. (1971). Classification without identification in visual search. *Quarterly Journal of Experimental Psychology, 23*, 178-186.

Broadbent, D.E. (1958). *Perception and Communication.* New York: Pergamon Press.

Broadbent, D.E. (1971). *Decision and Stress* . New York: Academic Press.

Duncan, J. (1980). The locus of interference in the perception of simultaneous stimuli. *Psychological Review, 87*, 272-300.

Duncan, J. (1987). Attention and reading: Wholes and parts in shape recognition—A tutorial review. In M. Coltheart (Ed.), *Attention and Performance XII: The Psychology of Reading* . Hillsdale, NJ: Erlbaum, pp. 39-61.

Duncan, J. & Humphreys (1989). Visual search and stimulus similarity. *Psychological Review, 96,* 433-458.

Eckstein, M.P. (1998). The lower visual search efficiency for conjunctions is due to noise and not serial attentional processing. *Psychological Science, 9*, 111-118.

Eckstein, M.P., Thomas, J.P., Palmer, J., & Shimozaki, S.S. (1999). A signal detection model predicts the effects of set-size on visual search accuracy for feature, conjunction, triple conjunction, and disjunction displays. *Perception and Psychophysics* (in press).

Eckstein, M.P. & Whiting, J.S. (1996). Visual signal detection in structured backgrounds. I. Effect of number of possible spatial locations and signal contrast. *Journal of the Optical Society of America, 13,* 1777-1787.

Gegenfurtner, K.R. & Sperling, G. (1993). Information transfer in iconic memory experiments. *Journal of Experimental Psychology: Human Perception and Performance, 19,* 845-866.

Gleitman, H. & Jonides, J. (1976) The cost of categorization in visual search: Incomplete processing of targets and field items. *Perception and Psychophysics, 20,* 281-288.

Green, D.M. & Swets, J.A. (1966). *Signal Detection Theory and Psychophysics*. New York: John Wiley & Sons, Inc.

Irwin, D.E. & Andrews, R.V. (1996). Integration and accumulation of information across saccadic eye movements. In I. Toshio and J.L. McClelland (Eds.), *Attention and performance XVI.* Cambridge, MA: MIT Press, pp. 125-155.

Kandell, E.R, Schwartz, J.H., & Jessell, T.M. (1991). *Principles of Neural Science, Third Edition.* Englewood Cliffs, New Jersey: Prentice Hall

Kleiss, J.A. & Lane, D.M. (1986). Locus and persistence of capacity limitations in visual information processing. *Journal of Experimental Psychology: Human Perception and Performance, 12,* 200-210.

Leshowitz, B., Taub, H.B., & Raab, D.H. (1968). Visual detection of signals in the presence of continuous and pulsed backgrounds. *Perception and Psychophysics, 4,* 207-213.

Lindsay, P.H., Taylor, M.M, & Forbes, S.M. (1968). Attention and multidimensional discrimination. *Perception and Psychophysics, 4,* 113-117.

Logan, G.D. (1998). What is learned during automatization? Obligatory encoding of spatial location. *Journal of Experimental Psychology: Human Perception and Performance, 24,* 1720-1736.

Logan, G.D. & Etherton, J.L. (1994). What is learned during automatization? The role of attention in constructing an instance. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 1022-1050.

Luce, R.D. (1959). *Individual Choice Behavior.* New York: Wiley.

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

Maunsell, J.H.R. & Newsome, W.T. (1987). Visual processing in monkey extrastriate cortex. *Annual Review of Neuroscience, 10,* 363-401.

McLean, J.E., Palmer, J., & Loftus, G.R. (1998). Measuring divided-attention effects in the perception and memory encoding of letters. Unpublished manuscript.

McLeod, P., Driver, J., & Crisp, J. (1988). Visual search for a conjunction of movement and form is parallel. *Nature (London), 332*, 154-155.

Mulligan & Shaw (1981). Attending to simple auditory and visual signals. *Perception and Psychophysics, 30,* 447-454.

Murdock, B.B. (1965). Effects of a subsidiary task on short-term memory. *British Journal of Psychology, 56,* 413-419.

Nagy, A.L. & Sanchez, R.R. (1990). Critical color differences determined with a visual search task. *Journal of the Optical Society of America, 7,* 1209-1225.

Nakayama, K. & Silverman, G.H. (1986). Serial and parallel encoding of visual feature conjunctions. *Investigative Opthamology and Visual Science, 27* (Suppl. 182)

Palmer, J. (1990). Attentional limits on the perception and memory of visual information. *Journal of Experimental Psychology: Human Perception and Performance, 16*, 332-350.

Palmer, J. (1994). Set-size effects in visual search: The effect of attention is independent of the stimulus for simple tasks. *Vision Research 34*,, 1703-1721.

Palmer, J. (1995). Attention in visual search: Distinguishing four causes of a set-size effect. Current *Directions in Psychological Science 4,* 118-123.

Palmer, J., Ames, C.T., & Lindsey, D.T. (1993). Measuring the effect of attention on simple visual search. *Journal of Experimental Psychology: Human Perception and Performance, 19,* 108-130.

Palmer, J. & McLean, J.E. (1995). Imperfect, unlimited-capacity, parallel search yields large set-size effects. Unpublished manuscript.

Pashler, H.E. (1987). Detecting conjunctions of color and form: Reassessing the serial search hypothesis. *Perception and Psychophysics, 41,* 191-201.

Pashler, H.E. (1987). Target-distractor discriminability in visual search. *Perception and Psychophysics, 41,* 285-292.

Pashler, H.E. (1988). Familiarity and visual change detection. *Perception and Psychophysics, 44,* 369-378.

Pashler, H.E. (1998). *The Psychology of Attention.* Cambridge, MA: MIT Press.

Pashler, H.E. (1993). Dual-task interference and elementary mental mechanisms. In D.E. Meyer and S. Kornblum (Eds.), *Attention and performance XIV.* Cambridge, MA: MIT Press, pp. 246-264.

Pelli, D. G. (1987) On the relation between summation and facilitation. *Vision Research, 27,* 119-123.

Posner, M.I., Snyder, C.R.R., & Davidson, B.J. (1980) Attention and the detection of signals. *Journal of Experimental Psychology: General, 109,* 160-174.

Shaw, M. (1980). Identifying attentional and decision-making components in information processing. In R.S. Nickerson (Ed.), *Attention and performance VIII.* Hillsdale, NJ: Erlbaum, pp. 277-296.

Shaw, M. (1984). Division of attention among spatial locations: A fundamental difference between detection of letters and detection of luminance increments. In H. Bouma and D.G. Bouwhuis (Eds.), *Attention and performance X* . Hillsdale, NJ: Erlbaum, pp. 109-121.

Shiffrin, R.M. & Gardner, G.T. (1972). Visual processing capacity and attentional control. *Journal of Experimental Psychology, 93* , 72-82.

Shibuya H. & Bundesen, C. (1988). Visual selection from multielement displays: measuring and modeling effects of exposure duration. *Journal of Experimental Psychology: Human Perception and Performance, 14,* 591-600

Shiu, L. & Pashler, H. (1994). Negligible effect of spatial precuing on identification of single digits. *Journal of Experimental Psychology: Human Perception and Performance, 20,* 1037-1054.

Silverstein, C. and Glanzer, M. (1971). Concurrent task in free recall: Differential effects of LTS and STS. *Psychonomic Science, 22,* 367-368.

Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs: General and Applied, 74,* 1-29.

Sperling, G. (1963). A model for visual memory tasks. *Human Factors, 5,* 19-31.

Sperling, G. (1969). A model for visual memory tasks. In R.N. Haber (Ed.), *Information-Processing Approaches to Visual Perception.* New York: Holt, Rinehart, and Winston, Inc., pp. 18-31.

Sperling, G. (1969). Successive approximations to a model for short-term memory. In R.N. Haber (Ed.), *Information-Processing Approaches to Visual Perception.* New York: Holt, Rinehart, and Winston, Inc., pp. 32-37.

Steinman, S.B. (1987). Serial and parallel search in pattern vision. *Perception, 16,* 389-399.

Tanner, W. P., Jr. (1956). Theory of recognition. *Journal of the Acousical Society of America, 28*, 882-888.

Tanner, W.P., Jr. (1961). Physiological implications of psychophysical data. *Annals of the New York Academy of Science, 89*, 752-765.

Townsend, J.T. & Ashby, F.G. (1983). *The Stochastic Modeling of Elementary Psychological Processes,* Cambridge, UK: Cambridge University Press.

Treisman, A.M. & Sato, S. (1990). Conjunction search revisited. *Journal of Experimental Psychology: Human Perception and Performance, 16,* 459-478.

Watson, A.B. (1979).Probability summation over time. Vision *Research, 19*, 515-522.

Wolfe, J.M., Cave, K.R., & Franzel, S.L. (1989). Guided Search: An alternative to the feature integration model for visual search. *Journal of Experimental Psychology: Human Perception and Performance, 15*, 419--433

Wolfe, J.M. (1994). Guided Search 2.0: A revised model of visual search. *Psychonomic Bulletin and Review, 1,* 202-238.

APPENDIX A: MEMORY STORAGE CONTROL EXPERIMENT

This control experiment was designed to determine whether or not observers' performance was limited by short-term memory storage capacity. Asymptotic performance for set size 4 was measured using long duration, high contrast stimuli. Near perfect performance would indicate that the limit on storage capacity was not reached with set size 4. All observers participated in this experiment first before any others reported in this manuscript. It was decided in advance that if an observer's performance was not at or above 95% correct, they were to be excluded from the rest of the study.

Method.

The methodology of the Memory Storage Control Experiment was the same as the Memory task part of the Combined Search and Memory Experiment with the following exceptions.

Stimuli. As in the Combined Search and Memory Experiment, two values along each of the 3 dimensions (contrast, orientation, and size) were chosen to define the 8 stimuli. However, the differences between the two values were larger than in the Combined Search and Memory Experiment to make the stimuli highly discriminable. Contrast was -25% or -100%. Orientation was vertical (0 degrees) or horizontal (90 degrees). Size in arc min of major axis length by arc min of minor axis length was 30 by 10 or 60 by 20, hence aspect ratio was 3 in both cases.

Design and Procedure. The display sequence on a trial was the same as the memory trial of the Combined Search and Memory Experiment except that the study display always consisted of a set of 4 stimuli. Trials were presented in 4 blocks of 24 with 4 practice trials at the beginning of each block. Each of the 8 stimuli served as the target 3 times per block, in random order.

Results

Mean percent correct was 96.6 ± 0.4. Individual scores for the 7 observers were 95, 96, 96, 96, 97, 98, and 98 percent correct.

Discussion

All observers met the a priori performance criterion of 95% or better on this task. Hence all 7 observers were retained, 6 of whom participated in all of the other experiments. The remaining observer, the author, did not participated only in the Combined Search and Memory Experiment. This result supported the notion that for these kinds of stimuli, a set size of 4 does not impose performance limitations due to a memory storage capacity limit. Other researchers have found similar results (Irwin and Andrews, 1996; Luck and Vogel, 1997).

APPENDIX B: DISCRIMINABILITY CONTROL EXPERIMENTS

These three control experiments were designed to measure discriminability thresholds for each observer so as to define stimuli for experiments that would satisfy two goals: First, the stimuli would be approximately equally discriminable from each other to avoid a confound with set size and to avoid various guessing strategies. Second, the stimuli would be well-defined in psychological space for the purposes of mathematical modeling (see Appendix C).

These experiments were essentially feature search experiments with set size 1. The same base stimulus served as the target in each trial. A single stimulus was presented after the target and the observers indicated whether it was the same or different from the target. The three experiments differed from one another in the types of distractors: either varying in contrast, orientation, or size, from the base stimulus. Because the three experiments were nearly identical, they are reported together below. They are referred to individually when necessary as the Contrast Experiment, the Orientation Experiment, and the Size Experiment.

Method.

The methodology of the Discriminability Control Experiments were the same as the Feature Search Experiment with the following exceptions.

Observers. The same six observers participated in these experiments as did in the Feature Search Experiment. The author also participated.

Stimuli. For each of the three experiments, the stimuli were 6 ellipses: a target/base and 5 types of distractors. The target ellipse was the same as in the Feature Search Experiment (-25% contrast, vertical orientation, 30 arc min major axis length, 10 arc min minor axis length, aspect ratio 3).

The 5 distractors differed from the target in one dimension: contrast, orientation, or size, for the Contrast Experiment, the Orientation Experiment, and the Size Experiment, respectively. The contrast increments from the base/target stimulus defining the 5 distractors in the Contrast Experiment were -5, -10, -15, -20, and -25%. In the Orientation Experiment, the increments were 3, 4, 6, 9, and 13 degrees of rotation from the vertical. In the Size Experiment, the increments were 2, 4, 6, 8, and 12 arc min in major axis length with corresponding increments of minor axis length to maintain the aspect ratio as close to 3 as possible given the pixelation of the display. These increment values were chosen to span the performance range. See Figure 23 for a schematic illustration of the stimuli.

Design and Procedure. The Design and Procedure applies to each experiment individually. As shown in Figure 24, the display sequence on a trial was identical to the Feature Search Experiment, except that set size was 1 in the 80 ms study display. Observers responded with the "p" or "a" key as before to indicate target present or absent (i.e. the study stimulus was the same as or different from the target). The correct response in the Figure 24 example is "a" because the target was absent from the study display.

Trials were presented in blocks of 30, not including 4 practice trials at the beginning of the block. A session consisted of 10 blocks. Each observer participated in 2 sessions, for a total of 600 trials. Distractor varied between blocks; there were 2 blocks for each distractor. Each of the 2 consecutive groups of 5 blocks within a 10-block session consisted of a block for each of the 5 distractors in random order within that 5-block group. Target status (present or absent) varied randomly within a block with the target present on half of the trials.

Analysis and Results

A sample data set from observer T in the Size Experiment is shown in Figure 25. Percent correct is plotted against discriminability (size increment, in units of major axis length increment, in this case) yielding 5 data points corresponding to the 5 distractors which were chosen to span the performance range from chance to perfect. Percent correct increased with increasing discriminability.

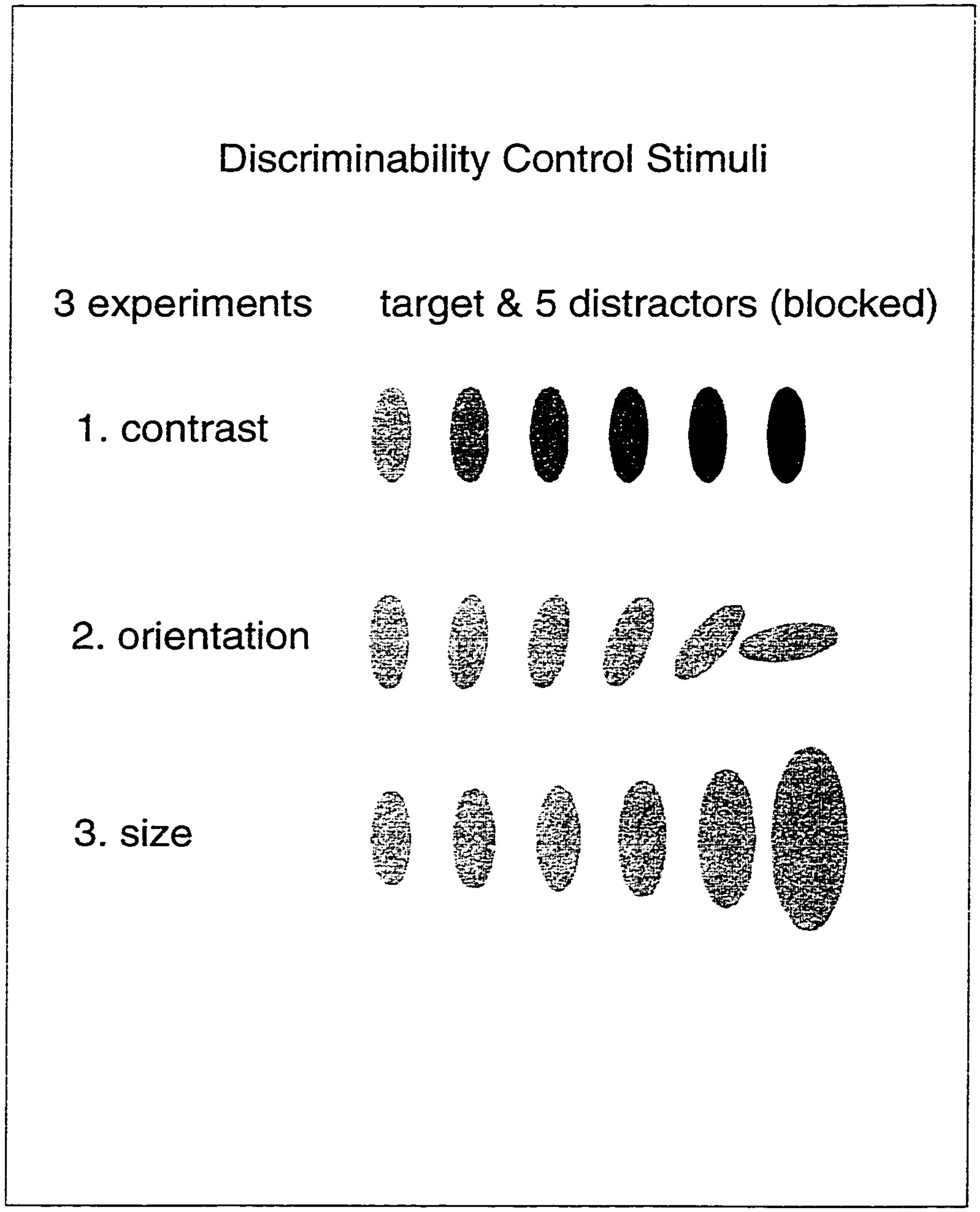

Figure 23: Discriminability Control Stimuli

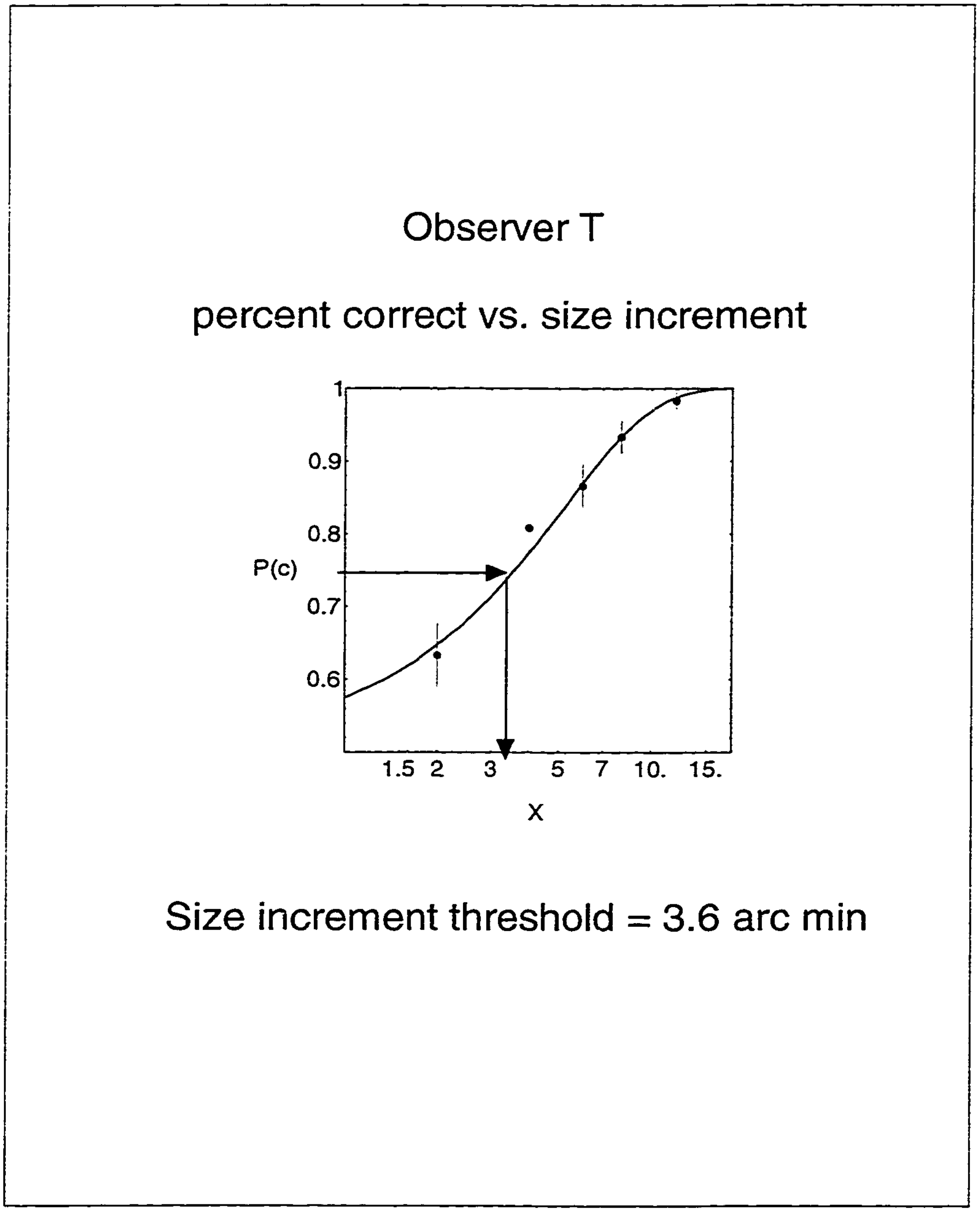

Figure 25: Sample Discriminability Control Data Set and Threshold Estimate

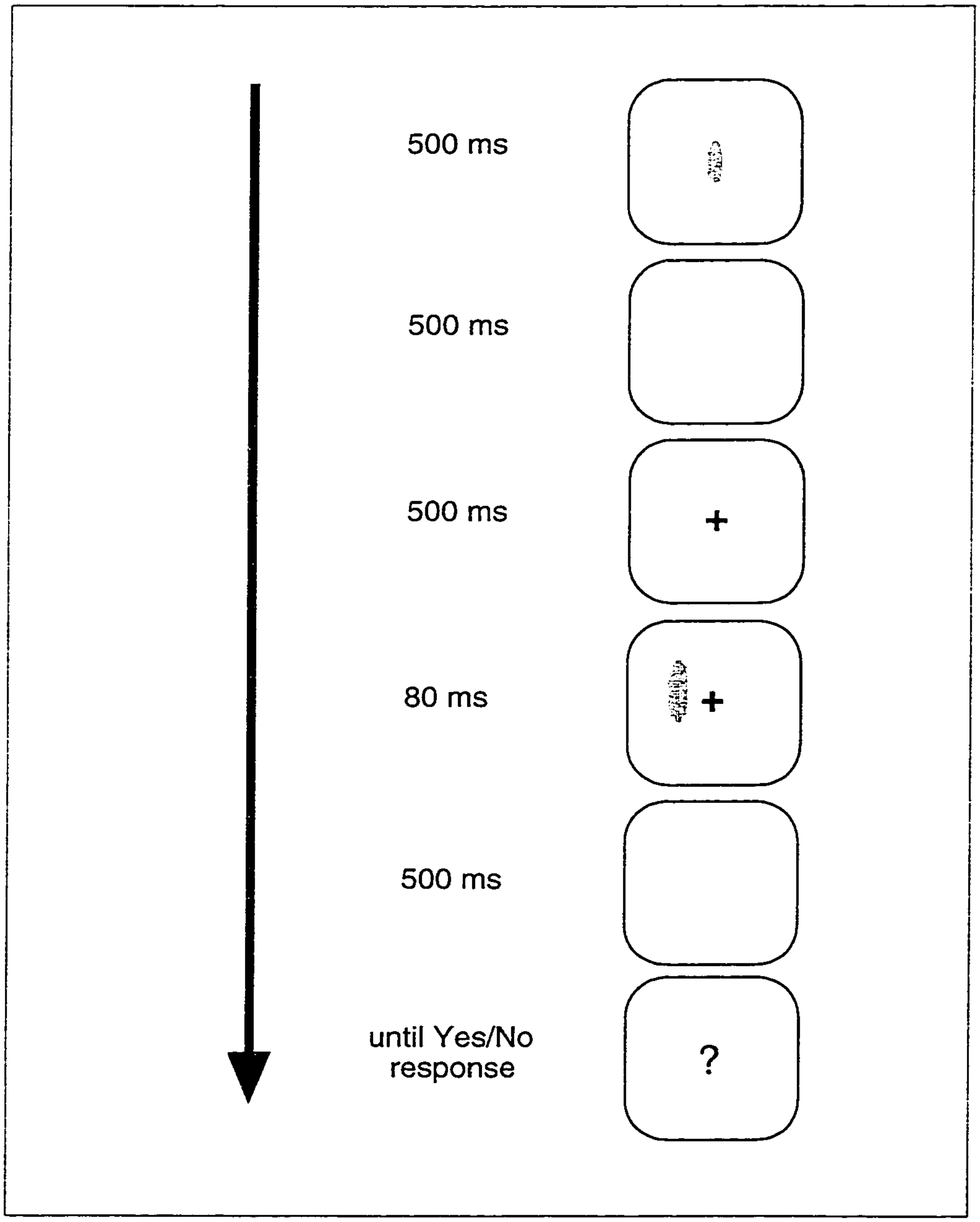

Figure 24: Discriminability Control Task

Psychometric functions were fit to each data set, shown for Observer T in Figure 25 as the smooth S-shaped curve. The form of the psychometric function was assumed to follow the d' power law, often used in psychophysics (Pelli, 1987). This allows for a nonlinear, and in particular, a power function, mapping of the stimulus increment variable onto the internal representation of the stimulus increment variable. In particular, given a stimulus increment variable, x, probability correct estimates, p_e, were calculated as follows, where F is the standard normal cumulative distribution function:

$$p_e = F((x/t)^S z_{crit})$$

or, equivalently,

$$p_e = P(z \le (\tfrac{x}{t})^s z_{crit})$$
$$= \tfrac{1}{2\pi} \int_{-\infty}^{(\frac{x}{t})^s z_{crit}} e^{\frac{-y^2}{2}} dy$$

where t and s are parameters of the psychometric function, the threshold and steepness, respectively, to be described further below, and z_{crit} is the z-score corresponding to threshold criterion performance, in this case, 75%. The x-scale corresponds to the external representation of the stimulus increment and the z-scale to the internal representation.

An optimization procedure due to Watson (1979) was used to find the values of t and s that produced the best fitting probability correct estimates according to a maximum likelihood objective function based upon binomially distributed errors. Lower and upper asymptotic performance was assumed to be 50% (chance) and 100% (perfect), respectively.

The parameter, t, represented the stimulus increment that yielded 75% correct performance, referred to as the discriminability threshold: 3.6 arc min for the psychometric function in Figure 25. The parameter, s, was a measure of steepness of the psychometric function. It was expected that an s of 1 would fit the data based upon results of previous pedestal experiments (Leshowitz, Taub, and Raab, 1968), and it was found that this was the case. When both t and s were estimated simultaneously, mean s values were 0.98 ±

0.04, 1.01 ± 0.07, and 0.95 ± .05 in the Contrast, Orientation, and Size Experiments, respectively. Hence the value of s was set to 1 and the thresholds were estimated alone, producing the following results:

In the Contrast Experiment, the mean threshold was a contrast increment of -9.6 ± 0.3%. In the Orientation Experiment, the mean threshold was an orientation increment of 3.7 ± 0.1. In the Size Experiment, the mean threshold was a major axis length increment of 4.0 ± 0.2 arc min.

Discussion

If these estimates are correct, a distractor with contrast -34.6% (-25% of the target plus a mean increment of -9.6%) should be correctly judged as distinct from the target by these observers 75% of the time. Similarly, a distractor with orientation 3.7 degrees (0 of the target plus a mean increment of 3.7 degrees) should be correctly judged as distinct from the target by these observers 75% of the time. Finally, the same would be expected of a distractor with major axis length 34 arc min (30 arc min of the target plus a mean increment of 4 arc min) and corresponding minor axis length to keep the aspect ratio as close to 3 as possible given the pixelation of the display (2 arc min per pixel).

These mean threshold values were used as a discriminability unit along the stimulus dimension to define the stimuli in the main experiments of this study. In one experiment stimuli were defined based on a multiple of 1.5 times the threshold values. In the other two experiments, a multiple of 2 was used. It was reasonable to assume that multiplying thresholds would affect discriminability in the same way on each dimension because the psychometric functions were verified to be approximately the same shape on each dimension within the range of interest.

The mean was used rather than tailoring the stimuli to each individual observer because observers were quite consistent with one another as evidenced by the small variance in their threshold estimates and because the resolution of the display prevented further refinement of precision (e.g. size can only be adjusted in pixels, not fractions of pixels).

APPENDIX C: MATHEMATICAL DETAILS OF THEORY

The mathematical details required for generating quantitative predictions of the Theory are outlined in this Appendix. The mathematical model differs a bit for each of the different tasks, stimuli, and stimulus sampling schemes. The model for each experiment will be discussed in turn. The basic assumptions as described in the Theory section are the same for each. The modeling goal in each case is to come up with the expressions for the hit and false alarm probabilities, and then solve for the criterion by setting H = 1 - FA which yields percent correct, the observable to be predicted.

Modeling of Feature Search Experiment

For the Feature Search Experiment, the modeling is the simplest. The distractors are homogeneous and the target and distractor vary only along one dimension at a time (the three types of distractors are blocked), thus the representation of a stimulus can be reduced down to a one-dimensional random variable instead of a three-dimensional random variable. The decision variable, V, is just equal to the value of the random variable, rather than the max of the three values along the three dimensions. The target representation is distributed normally with mean 0 and variance = n^a. The distractor representation is the same but shifted to a higher mean which will be referred to as d for discriminability. The instances of this one-dimensional model with a=0 and a=1 are identical to the unlimited-capacity and fixed-capacity models, respectively, used by Palmer et. al. (1993).

Let n be the set size. The hit and false alarm probabilities are defined as described in the Theory section, where c represents the criterion:

$$H = 1 - P(V_T > c, V_{D1} > c, V_{D2} > c, \ldots V_{Dn-1} > c)$$

$$FA = 1 - P(V_{D1} > c, V_{D2} > c, \ldots V_{Dn} > c)$$

Where V_T is the decision variable which in this one-dimensional case is equal to the representation corresponding to the target stimulus and $V_{D1}, \ldots V_{Dn}$ are the representations corresponding to the n different distractors. It is assumed that V_T, V_{D1}, ...

V_{Dn} are independent random variables, which is to say that the value of each individual random variable representation which is sampled from its distribution does not depend upon the value of the other samples, hence the joint probability is the product of the individual probabilities:

$$H = 1 - P(V_T > c)\, P(V_{D1} > c)\, P(V_{D2} > c) \ldots P(V_{Dn-1} > c)$$

$$FA = 1 - P(V_{D1} > c)\, P(V_{D2} > c) \ldots P(V_{Dn} > c)$$

Because all of the distractors come from the same distribution in this one-dimensional case, the equation can be simplified to

$$H = 1 - P(V_T > c)\, [P(V_D > c)]^{n-1}$$

$$FA = 1 - [P(V_D > c)]^{n}$$

Thus for set size 1

$$H = 1 - P(V_T > c)$$

$$FA = 1 - P(V_D > c)$$

and for set size 4

$$H = 1 - P(V_T > c)\, [P(V_D > c)]^{3}$$

$$FA = 1 - [P(V_D > c)]^{4}$$

Given n, a, and d which define the distributions of the V's (V_T is normal with mean 0 and variance=n^a and V_D is normal with mean d and variance=n^a), c is determined by solving

$$H = 1 - FA$$

based upon the equal bias assumption (observers are given bias feedback and come close, so this assumption is reasonable). Equal bias with normal distributions means that the criterion is centered between the two distributions, in other words, probability correct is the same on target present trials (the hit probability H) as it is on target absent trials (the correct rejection probability which is 1-FA). Thus, once c is determined, the probability correct prediction is just

$$PC = H = 1 - FA.$$

Modeling of Postcue Memory Experiment

For the Postcue Memory Experiment, the modeling is a bit more complex. In this case, the distractors are heterogeneous and sampled randomly with replacement from the full set of 8 stimuli. Thus the three-dimensional representation is necessary (refer back to Figure 9). The decision variable, V, is the maximum of the three dimensions: $V = \max(x,y,z)$ (refer back to Figure 11). Let n be the set size. The target representation distribution is always placed at the origin for simplicity in the modeling, hence it 3-D normal with mean 0 and variance = n^a. The distractor representations have the same variance, but the means are at the different corners of the cube with side length d. Thus one mean is at (0,0,d), another is at (0,d,0), and another is at (0,d,d) and so forth.

Because there is a postcue, only the representation corresponding to that one stimulus that is cued is evaluated in the decision. That stimulus will be the target or a distractor. The hit and false alarm probabilities are defined as follows, where c represents the criterion:

$$H = P(V_T < c)$$

$$FA = P(V_D < c)$$

The distribution for V_T is known at this point and hence the hit probability could be calculated, given c, but there are 7 different distractors that could be in the cued location and so each possibility must be taken into account when determining the distribution of V_D. Of these 7 possibilities, there are 3 groups. Let DI be a distractor representation from the first group that are the 3 at the closest corners of the cube, one side length away from the target. These have means at (d,0,0), (0,d,0), and (0,0,d). Let DII be a distractor representation from the second group that are the 3 are 1.4d from the target with means (d,d,0), (d,0,d), and (0,d,d). Lastly, let DIII be a distractor representation from the third group in which there is only 1 and it is 1.7d from the target with mean (d,d,d). Each of the 7 are equally likely, hence

$$FA = (3/7)\ P\ (V_{DI} < c) + (3/7)\ P\ (V_{DII} < c) + (1/7)\ P\ (V_{DIII} < c)$$

Substituting the max functions for the V's, we get

$$FA = (3/7)\ P((max(x_{DI}, y_{DI}, z_{DI}) < c)$$

$$+ (3/7)\ P((max(x_{DII}, y_{DII}, z_{DII}) < c)$$

$$+ (1/7)\ P((max(x_{DIII}, y_{DIII}, z_{DIII}) < c)$$

Assuming the values one each dimension are independent from one another

$$FA = (3/7)\ P(x_{DI} < c)\ P(y_{DI} < c)\ P(z_{DI} < c)$$

$$+ (3/7)\ P(x_{DII} < c)\ P(y_{DII} < c)\ P(z_{DII} < c)$$

$$+ (1/7)\ P(x_{DIII} < c)\ P(y_{DIII} < c)\ P(z_{DIII} < c)$$

Now the subscripted x, y, and z's in the equation above are one-dimensional normal random variables with mean of either 0 or d and variance = n^a. Thus, the equation simplifies to the following

$$FA = (3/7)\ [P(u_0 < c)]^2\ P(u_d < c) + (3/7)\ P(u_0 < c)\ [P(u_d < c)]^2 + (1/7)\ [P(u_d < c)]^3$$

where u_0 is 1-D normal with mean 0 and variance = n^a and u_d is 1-D normal with mean d and variance = n^a. Now, given the hit probability from above simplified in the same way:

$$H = P\ (V_T < c) = [P(u_0 < c)]^3$$

and given n, a, and d which define the distributions of the random variables, c is determined as in all of the other models, by assuming equal bias and solving

$$H = 1 - FA$$

which yields probability correct

$$PC = H = 1 - FA$$

It is of note that in this model, the only effect of set size is in the variance of the distributions. The set size, n, does not enter into the equations for FA and H as it did in the Feature Search Model, because only the representation of the postcued stimulus is evaluated in the decision process. This explains why the set-size effect is zero when a = 0 and hence representation variance = var_i = constant for all set sizes.

Modeling of Combined Search and Memory Experiment

For the Combined Search and Memory Experiment, the modeling is again, a little different. In this case, the distractors are heterogeneous, sampled randomly without replacement from the set of 7 non-target stimuli.

Set Size 1 Case

The set size 1 case is identical to the Postcue Memory Model because in that model only one representation, that of the postcued stimulus, is ever being evaluated in the decision process, which is the same situation as the set size 1 condition in this experiment. Thus the hit and false alarm probabilities are, as above,

$$H_1 = [P(u_0 < c)]^3$$

$$FA_1 = (3/7)\,[P(u_0 < c)]^2\, P(u_d < c) \;+ (3/7)\, P(u_0 < c)\,[P(u_d < c)]^2 + (1/7)\,[P(u_d < c)]^3$$

where the subscripts of 1 on H and FA refer to the set size and the random variables u_0 and u_d are normal with variance = n^a and means 0 and d, respectively.

Set Size 4 Case

The set size 4 case is more complex because the probabilities are conditioned upon the sample of distractors and there are many possible samples: When the target is present, there are 7 choose 3 = 35 combinations of 3 distractors. When the target is absent, there are 7 choose 4 = 35 combinations of 4 distractors. Furthermore, there are 3 groups of distractors as described in the Postcue Model above: I. those that are d away from the target in representation space, II. those that are 1.4d away, and III. the one that is 1.7d away.

To begin, the hit and false alarm probabilities for set size 4 can be expressed in the same high-level way as all the other models:

$$H = 1 - P\,(V_T > c, V_{D1} > c, V_{D2} > c, V_{D3} > c)$$

$$FA = 1 - P\,(V_{D1} > c, V_{D2} > c, V_{D3} > c, V_{D4} > c)$$

Now the individual distractors are not independent from one another, so these joint probabilities are not the product of the individual probabilities.

False Alarm Probability. First consider the false alarm probability. Let DI, DII, and DIII indicate distractors from the three groups as defined above. There are 35 possible combination of 4 distractors. Those combinations break down into 7 kinds: 1 of C1=(DI, DI, DI, DIII), 1 of C2=(DII, DII, DII, DIII), 3 of C3=(DI, DI, DI, DII), 3 of C4=(DI, DII, DII, DII), 9 of C5=(DI, DI, DII, DIII), 9 of C6=(DI, DII, DII, DIII), and 9 of C7=(DI, DI, DII, DII). Let Ei be the event that, given Ci is the kind of distractor combination sampled, all of the decision variables corresponding to the distractor representations in that combination are greater than c. Then

$$FA = 1 - [\ (1/35)\ P(E1) + (1/35)\ P(E2) + (3/35)\ P(E3) + (3/35)\ P(E4) + (9/35)\ P(E5) + (9/35)\ P(E6) + (9/35)\ P(E7)\ ]$$

Now each of these event probabilities must be determined. Consider E5 first because C5 has at least one of each of DI, DII, and DIII.

$$P(E5) = P(V_{DI} > c,\ V_{DI} > c,\ V_{DII} > c,\ V_{DIII} > c)$$

The joint probability can be split up because independence now applies

$$P(E5) = P(V_{DI} > c)\ P(V_{DI} > c)\ P(V_{DII} > c)\ P(V_{DIII} > c)$$
$$= [P(V_{DI} > c)]^2\ P(V_{DII} > c)\ P(V_{DIII} > c)$$

Consider the first factor

$$P(V_{DI} > c)$$
$$= P((\max(x_{DI}, y_{DI}, z_{DI}) > c)$$
$$= 1 - P(x_{DI} < c,\ y_{DI} < c,\ z_{DI} < c)$$
$$= 1 - P(x_{DI} < c)\ P(y_{DI} < c)\ P(z_{DI} < c)$$
$$= 1 - [P(u_0 < c)]^2\ P(u_d < c)$$
$$= 1 - F_0^2\ F_d$$

$= PI$

The subscripted x, y, and z's in the equations above are one-dimensional normal random variables with mean of either 0 or d and variance = n^a. For the DI distractor representation, two have means at 0, the third at d. As above, u_m represents a 1-D normal random variable with mean m and variance = n^a. F_m represents the cumulative normal distribution with mean m and variance n^a, evaluated at c. PI is a symbol to be used for simplifying notation.

Using this notation and applying the same method to simplify the other factors of the event probabilities

$$P(V_{DII} > c)$$
$$= 1 - F_0 F_d^2$$
$$= PII$$

and

$$P(V_{DIII} > c)$$
$$= 1 - F_d^3$$
$$= PIII$$

Now the event probabilities are reduced to products of the easy-to-compute probabilities PI, PII, and PIII:

$$P(E1) = PI^3\ PIII$$
$$P(E2) = PII^3\ PIII$$
$$P(E3) = PI^3\ PII$$
$$P(E4) = PI\ PII^3$$
$$P(E5) = PI^2\ PII\ PIII$$
$$P(E6) = PI\ PII^2\ PIII$$
$$P(E7) = PI^2\ PII^2$$

Finally, these expressions for the event probabilities are substituted into the false alarm probability for set size 4 repeated below.

$$FA = 1 - [\ (1/35)\ P(E1) + (1/35)\ P(E2) + (3/35)\ P(E3) + (3/35)\ P(E4) + (9/35)\ P(E5) + (9/35)\ P(E6) + (9/35)\ P(E7)\ ]$$

Hit Probability. Next consider the hit probability. The target is present and there are 3 distractors as well to make up a set of 4. There are 35 possible combination of 3 distractors. Those combinations break down into 7 kinds, the complements of those combinations of four defined above: 1 of C1=(DII, DII, DII), 1 of C2=(DI, DI, DI), 3 of C3=(DII, DII, DIII), 3 of C4=(DI, DI, DIII), 9 of C5=(DI, DII, DII), 9 of C6=(DI, DI, DII), and 9 of C7=(DI, DII, DIII). Let Ei be the event that, given Ci is the kind of distractor combination sampled, all of the decision variables corresponding to the distractor representations in that combination *as well as the target representation* are greater than c. Then

$$H = 1 - [\ (1/35)\ P(E1) + (1/35)\ P(E2) + (3/35)\ P(E3) + (3/35)\ P(E4) + (9/35)\ P(E5) + (9/35)\ P(E6) + (9/35)\ P(E7)\ ]$$

Now each of these event probabilities must be determined. Consider E7 first because C7 has at least one of each of DI, DII, and DIII.

$$P(E7) = P(V_T > c,\ V_{DI} > c,\ V_{DII} > c,\ V_{DIII} > c)$$

The joint probability can be split up because independence now applies

$$P(E7) = P(V_T > c)\ P(V_{DI} > c)\ P(V_{DII} > c)\ P(V_{DIII} > c)$$

The probabilities of each of the factors except the first one was determined above. With the same method applied to the first factor

$$\begin{aligned} & P(V_T > c) \\ & = P((\max(x_T, y_T, z_T) > c) \\ & = 1 - P(x_T < c,\ y_T < c, z_T < c) \\ & = 1 - P(x_T < c)\ P(y_T < c)\ P(z_T < c) \\ & = 1 - [P(u_0 < c)]^3 \end{aligned}$$

$$= 1 - F_0^3$$

$$= P0$$

The subscripted x, y, and z's in the equations above are one-dimensional normal random variables with means 0 and variance = n^a because they represent the target which it at the origin. As above, u_m represents a 1-D normal random variable with mean m and variance = n^a. F_m represents the cumulative normal distribution with mean m and variance n^a, evaluated at c. P0 is a symbol to be used for simplifying notation.

Now the event probabilities are reduced to products of the easy-to-compute probabilities P0, PI, PII, and PIII:

$$P(\underline{E1}) = P0\ PII^3$$

$$P(\underline{E2}) = P0\ PI^3$$

$$P(\underline{E3}) = P0\ PII^2\ PIII$$

$$P(\underline{E4}) = P0\ PI^2\ PIII$$

$$P(\underline{E5}) = P0\ PI\ PII^2$$

$$P(\underline{E6}) = P0\ PI^2\ PII$$

$$P(\underline{E7}) = P0\ PI\ PII\ PIII$$

Finally, these expressions for the event probabilities are substituted into the false alarm probability for set size 4 repeated below.

$$H = 1 - [\ (1/35)\ P(\underline{E1}) + (1/35)\ P(\underline{E2}) + (3/35)\ P(\underline{E3}) + (3/35)\ P(\underline{E4}) + (9/35)\ P(\underline{E5}) + (9/35)\ P(\underline{E6}) + (9/35)\ P(\underline{E7})\ ]$$

<u>Probability Correct Predictions</u>

The hit and false alarm probabilities are functions of the criterion, c, as well as n, a, and d which define the distributions of the random variables. The criterion, c, is determined as in all of the other models, by assuming equal bias, normal distributions, and solving

$$H = 1 - FA$$

which yields probability correct

$$PC = H = 1 - FA.$$

BIOGRAPHICAL NOTE

Jennifer E. McLean received her M.S. in applied mathematics in 1993 and her Ph.D. in psychology in 1999, both from the University of Washington.