UNIVERSITY OF ALBERTA

# VISUAL SEARCH FOR MULTIPLE TARGETS: THE INFLUENCE OF STIMULUS SIMILARITY AND MEMORY LOAD ON SEARCH STRATEGY

BY

© JOCELYN B. AUBREY

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree of DOCTOR OF PHILOSOPHY.

DEPARTMENT OF PSYCHOLOGY

EDMONTON, ALBERTA

Spring, 1992

ISBN 0-315-73013-7

UNIVERSITY OF ALBERTA

RELEASE FORM

NAME OF AUTHOR: Jocelyn B Aubrey

TITLE OF THESIS: Visual Search for Multiple Targets: The Influence of Stimulus Similarity and Memory Load on Search Strategy

DEGREE: Doctor of Philosophy

YEAR THIS DEGREE GRANTED: 1992



#1, 832 Armour Road

Peterborough, Ontario, K9H 2A2

1991 12 31

UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled VISUAL SEARCH FOR MULTIPLE TARGETS: THE INFLUENCE OF STIMULUS SIMILARITY AND MEMORY LOAD ON SEARCH STRATEGY by JOCELYN B. AUBREY in partial fulfillment of the requirements for the degree of DOCTOR OF PHILOSOPHY in PSYCHOLOGY.

Dr. Allen R. Dobbs

Dr. C. Donald Heth

Dr. Michael Dawson

Dr. Peter Dixon

Dr. Roc E. Walley

Dr. R. B. Wilberg

Dr. E. Ratsoy

Dr. Charles Scialfa

1991 12 31

Abstract

The experiments in this series explored the attentional processing demands associated with the visual search of large, unorganized arrays of letters for several instances of a target letter(s). The paradigm required subjects to use a light-sensitive pen to search computerized displays for all instances of a target(s) embedded among a large number of distractors. Search through such arrays is presumed to require an overall strategy that ensures effective monitoring of parts of the array that have been scanned and facilitates the choice of what parts will be searched next. Search strategies are apparent in the pattern of target detections (i.e. vertical or horizontal sweeps over the array). Such strategic monitoring is assumed to utilize the limited capacity executive component of the working memory system. Also utilizing executive capacity are control processes such as comparison of display items to target templates in memory and rehearsal of target items. The overall hypothesis tested throughout this set of experiments was that strategic monitoring of search could be disrupted by manipulations to the processing demands of the search. When processing demands were manipulated through perceptual relationships among targets and distractors (Experiments 1 and 2) or via memory load and sequential familiarity of letters in the memory set (Experiments 3A and 3B), there was some evidence that reductions in processing demand improved strategy maintenance. Although there was limited support for the original hypothesis, the consistent finding that processing load negatively affected time and accuracy suggests an alternative

hypothesis. Because the paradigm forced the use of a strategy for monitoring the search, the attentional requirements associated with the maintenance of that strategy interfered with search efficiency (time and accuracy) when processing demands increased.

The first two experiments also tested the Duncan and Humphreys (1989) similarity model which predicts that the efficiency of search for a specific target is directly related to the amount of perceptual similarity among and between targets and nontargets. The results provided limited support for the model when alphabetic stimuli are used and the task requires search through large arrays for several instances of a target.

TABLE OF CONTENTS

## LIST OF TABLES

## LIST OF FIGURES

Countless times during our daily lives we visually scan our environment in search of particular items. Often we need to look for more than one of the same thing (blueberries on a bush; all the chocolate chip cookies on a tray of mixed cookies) or a number of different things (searching through an unordered array of various sized nuts and bolts for both nut and bolt in 6 mm and 15 mm sizes). Rapid and accurate detection is best accomplished if we scan in such a way as to avoid repetitious rescanning of areas of the visual scene. If the objects to be scanned are arranged in some kind of order, rows and columns for example, the orderly arrangement provides scan paths (e.g., back and forth, up and down) that facilitate easy monitoring of search throughout the display. When there is little order (blueberries are seldom arranged in rows on a bush) the viewer must invoke some kind of search strategy. The use of an overall search strategy ensures that the viewer can effectively monitor which parts of the scene have already been searched and where to look next.

In addition to keeping in mind a useful search strategy during the search, the viewer may have more or less difficulty finding the desired objects depending on the perceptual features of the environment. Clumps of blueberries set against green leaves are relatively easy to distinguish. On the other hand, nuts and bolts are all the same colour and although the two sizes to be found may be different (6mm and 15 mm), they are mixed in with others that are close in size to each of those required (10 mm and 18 mm, for instance). These two examples represent search problems that vary considerably in the ease with which they can be accomplished because of the differing levels of perceptual similarity between items to be found and the background in which they are embedded. Finding the appropriate two sizes of nuts and bolts is much more difficult than finding clumps of blueberries, whether or not a particular search strategy is used.

The major question of interest in this paper concerns the interaction between the attentional demands associated with the perceptual features of arrays to be searched and those utilized in the monitoring of search strategy. There has been considerable research investigating the parameters affecting the efficiency of visual search for a target among several distractors. A number of factors have been identified that influence the time to find a target and the accuracy of detection. These include number of items in the target set, number in the distractor set, featural relationships between targets and distractors and size of the array (Duncan, 1985; Madden & Allen, 1989; Pashler, 1987a, 1987b; Treisman & Gelade, 1980; Treisman & Gormican, 1988). In the majority of the studies in the relevant literature, test arrays included only one target. Those few studies in which arrays contained multiple targets required subjects to respond to only one of them (e.g., Duncan, 1980; Erickson & Schultz, 1979; Ward & McClelland, 1989). Furthermore, very little attention has been paid to the strategies people use when they search. The paradigm used in this series of studies requires subjects to search for several instances of a target or targets among a very large number of distractors (e.g. Aubrey & Jutai, 1989). The use of these kinds of arrays has considerable heuristic value for aiding our understanding of the strategic processes associated with search because the patterns of detection allow inferences to be made about strategy use. As well, such investigations are valuable for determining whether the factors that prior research has demonstrated affect the efficiency of search for one target, similarly influence search for many targets.

The remainder of this introduction will present a review, first of the rather limited literature on search strategy and a discussion of the possible cognitive processes involved in the use of such strategies. This will be followed by a summary of the literature on visual search for one target. That research has determined many of the stimulus and procedural factors that influence search efficiency and the variations in processing demands associated with manipulations of these factors.

Search strategy

Researchers in the area of human factors have attempted to determine the relationship between search strategy and the efficient detection of faults on industrial parts. The models of search strategy that have been investigated are based on assumptions about the sequence of eye fixations and their durations. At one extreme are models that treat eye fixations as being completely independent and random so that a fixation can occur anywhere in the visual field independent of where the previous fixation occurred (Krendel & Wodinski, 1960; Lamar, 1960 cited in Morawski, Drury & Karwan, 1980). At the other extreme is a model proposed by Williams (1966) that treats fixations as samples from the search field that are made without replacement. This model assumes that the sequence of fixations proceeds systematically through one scan over the array and if the target has not been detected, another systematic scan is undertaken. It is important to note that defining systematic search in this way does not imply any regular search pattern, only that any one area will not be refixated during a scan. As defined by these models, systematic search is logically more efficient than random search (Morawski et al., 1980).

Neither model appears to fit the experimental data very well, however (Megaw & Richardson, 1979). Arani, Karwan and Drury (1984) proposed a variable-memory model of visual search that assumes that search is intended to be systematic but suffers from imperfect memory. Systematic search is a special case in which memory for previous locations is perfect and random search is a case in which memory is completely lacking. The authors demonstrated that this model, which allows for some refixations within a systematic scan of the field, fits some previously published experimental data.

Although attempts to mathematically model search strategies are valuable for aiding in the determination of search efficiency, few studies have considered the influence of stimulus parameters. One study suggested that the size of the visual lobe (the peripheral area around the central fixation point from which information can be extracted) is important in determining search efficiency (Kraiss

& Knäeuper, 1982). Generally, the size of the visual lobe is small when the target is embedded in a complex background or surrounded by irregularly positioned nontargets that are highly similar to the target. A model proposed by these authors and tested against published experimental data indicated that a systematic strategy does not improve performance over a random strategy when the visual lobe area is relatively large. However, when the visual lobe size is small, the use of a systematic strategy results in shorter times and more accurate detection. In another study, although visual lobe measurements were not taken into consideration, eye movements of search through a densely packed array of homogeneous non-targets for a stimulus of high similarity, were suggestive of a random walk pattern (Scinto, Pillalamarri & Karsh, 1986). These studies of search for one target suggest that perceptual factors may be of particular importance when assessing the kind of strategies people use.

Processing demands of search strategies. The maintenance of strategies and monitoring of performance during problem-solving tasks are usually considered to be higher level cognitive processes. In two recent conceptualizations, they have been ascribed to the central executive component of Baddeley's (1986; Baddeley & Hitch, 1974) working memory model and the supervisory attentional system (SAS) modelled by Norman and Shallice (1986; Shallice, 1988). In both models, the executive or SAS, is responsible for the scheduling of active processes and monitoring of memory activities.

One role of the executive in a visual search task would be to control processes such as comparison of display items to target templates in memory, rehearsal of target items, response programing and shifting of attention to other parts of the display. Monitoring of the attention shifts associated with searching large arrays for many instances of a target also would be under executive control. This monitoring is conceptually similar to memory updating which is an important component of many working memory tasks such as mental arithmetic (Hitch, 1978) and counting (Logie & Baddeley, 1987) and is believed to be coordinated by the central executive (Morris & Jones, 1990). In a visual search task, if the array is organized in rows and columns, there is little monitoring necessary - the arrangement of the display dictates search strategy. When the items are randomly arrayed, it becomes necessary to maintain some kind of strategic control over the shifts in attention. Presumably, the execution of an overall strategic plan would lessen the memory demands associated with monitoring what part of the array had already been searched because knowing where one is now searching would provide a reminder of where one has been. Research on unordered arrays indicates that the majority of normal subjects use systematic strategies as demonstrated by detection patterns of horizontal or vertical sweeps over the array (Aubrey & Jutai, 1989; Jutai, 1989). With pathological populations that have been determined to have attentional deficits, search patterns tend to be more random (Gauthier, Dehaut & Joanette, 1989; Weintraub & Mesulam, 1988).

Stimulus parameters and search efficiency

A very extensive literature has examined the cognitive and perceptual processes involved in the visual search of various sized arrays for a target. The paradigm most commonly used involves presenting test arrays consisting of several items (usually 2-8) and requiring the subject to indicate as rapidly as possible if a particular target is present or absent. If the slope of the response time function plotted against array size is relatively flat, the searches are assumed to have taken place in parallel; all items in the array having been surveyed in one apprehension. Target detection in such cases is presumed to be 'preattentive' or 'automatic' - not requiring attention. If reaction time increases with the size of the array, it is assumed that each item in the array was inspected and a 'controlled' search had been undertaken. This is determined by comparing the slope of the positive trials in which a target was present to negative trials that did not contain a target. On average, the target can be found on positive trials by searching half of the items (serial, self-terminating search), whereas all items must be scanned on negative trials. Thus, the slope related to negative trials is steeper than that for positive trials. Determining which factors are responsible for automatic and which for controlled search has been the focus of considerable investigation.

One of the more prominent formulations that has been tested extensively is Treisman's feature integration model (Treisman, 1988; Treisman & Gelade, 1980). In this model, basic features such as colour, orientation and form are available at preattentive levels of processing, but the conjoining of these features requires focal attention. Thus, parallel positive and negative response time slopes related to search for a red circle among green circles indicate automatic search processing whereas a search for a red square among red circles and green squares results in slope differences indicative of controlled, serial scanning. The model predicts, furthermore, that the slope of the target absent trials will be twice as steep as that for the target present trials (Treisman, 1988).

Although the evidence in support of automatic processing of simple features is strong, there now exists considerable evidence to suggest that in certain situations, conjoined features also may be processed preattentively. Those items that are easily discriminable from the distractors, regardless of whether or not they are identifiable in terms of simple features or more complex combinations of simple features, are detected more rapidly than items that are perceptually similar to distractors (Dehaene, 1989; Moraglia, Maloney, Fekete & Al-Basi, 1989; Duncan, 1989; Humphreys, Quinlan & Riddoch, 1989; Treisman & Gormican, 1988). These data, in addition to demonstrations that some controlled processing takes place even when detection is apparently preattentive and automatic (Madden & Allen, 1989; Ward & McClelland, 1989), suggest that the sharp distinction between the two types of processing is no longer tenable. Duncan and Humphreys (1989), for example, have presented a model of search efficiency based on a continuum of similarity between targets and distractors that postulates a corresponding continuum of processing demand from automatic to controlled.

In addition to the research that has considered the importance of perceptual features in visual search tasks, other studies have manipulated the size of the target set. The paradigm combines memory search with visual search by requiring subjects to visually scan an array for one member of a specified set of targets that are held in memory. Early research with this paradigm demonstrated that time to find a target increased linearly with increasing size of the target set (Kaplan, Carvellas & Metlay, 1966; Neisser, Novick & Lazar, 1964). It was assumed that the increase in time was due to a serial comparison of display items to each item in memory (Sternberg, 1966). With extensive practice, however, subjects can find the target from a set of ten items as rapidly as from a single item set (Neisser et al., 1964). The development over practice of automatic detection does not happen in all cases. It depends on whether the targets and distractors are mapped in a consistent or varied manner over blocks of trials. In consistent mapping, targets are taken from the same set and distractors from a different set. After sufficient practice, targets tend to 'pop out' and are identified very rapidly, regardless of the number of items in the target set, suggesting that they have been processed automatically. In the varied mapping procedure an item may be a target on one trial and a distractor on the next. For example, the search may be for digits among letters on one trial and letters among digits on the next trial. In this case, in spite of extensive practice, there is a linear increase in detection times with increases in target set size suggesting that serial or controlled processing characterizes this kind of search task (Schneider, Dumais & Shiffrin, 1984; Shiffrin & Schneider, 1977). Although there has been some research suggesting that consistent mapping is not always required in order to achieve automatic detection (Madden, 1982; Shiffrin & Czerwinski, 1988), it is particularly important to note that in most studies, even when stimuli are consistently mapped, set size effects are found early in training (but see Jonides and Gleitman, 1976 for a contrary finding).

The preceding review of the visual search literature demonstrated that various stimulus and experimental factors have been found to exert different attentional demands on the searcher. Although the research is based primarily on a search for one instance of the target, many of the same principles should apply to a search for several instances of a target. It is just that sort of search task that is the focus of consideration in this series of studies. Contrary to other visual search tasks, the trial does not end when one target has been found. Efficient search requires that the subject monitor what parts of the array have been scanned in order not to duplicate effort. Given that maintenance of an overall search strategy requires attentional control and that processing demands vary with perceptual similarity between targets and distractors and number of memory comparisons required, it can be hypothesized that manipulations to stimulus features and task demands could interfere differentially with search strategy.

The Duncan and Humphreys' similarity model. The point of departure for manipulations to stimulus factors is the similarity model proposed by Duncan and Humphreys (1989) which predicts that the efficiency of search for a specific target is directly related to the amount of perceptual similarity among and between targets and nontargets. The evidence supporting this model is based on studies in which subjects search for one instance of a target within a relatively small number of distractor items (typically 1 to 7). The model posits an early stage of processing in which an overall perceptual description of the visual field is developed. This stage is considered to occur outside of awareness and proceeds automatically. The visual representation is segmented such that parts that can be linked together (Gestalt groupings) are formed or boundaries are defined between parts that must be described separately. One important mechanism for such segmentation involves perceptual groupings based on proximity or similarity. The segments are referred to as *structural units*.

These structural units compete for access to a limited capacity visual short term or working memory (WM). Structural units are thought of as having some 'weight' which reflects the strength with which they compete for WM access. A structural unit gains weight in direct relation to its similarity to a template representation of the target item of interest. Any change in weight to one structural unit is distributed to other units proportional to the strength of grouping between them (weight linkage). This weight linkage can result in a phenomenon that Duncan and Humphreys refer to as 'spreading suppression' in which strongly grouped nontargets can be efficiently rejected due to a shared reduction in weight. The authors summarize selection weighting as follows: "...two factors combine to determine selection weights. The first is match of each input against a template of currently needed information. Weights increase with increasing match. The second is weight linkage. Any change in weight for one input is distributed to others in proportion to the strength of perceptual grouping." (p. 446).

Of considerable importance in this account of visual selection are the similarities between targets and nontargets and those among nontargets. Increasing similarity between targets and nontargets means that there will be greater competition for WM access because of high weights for both kinds of items. The less similarity there is among nontargets, the less likelihood that spreading suppression will take place to aid in efficient, parallel rejection of groups of nontargets. Spreading suppression is of greatest benefit when grouping within a set of nontargets is stronger than grouping between nontargets and targets. To the extent that there is high similarity between targets and nontargets and reduced similarity among nontargets (little spreading suppression), visual search will be inefficient (reflected in long reaction times for detecting the target). The size of structural units will be small, possibly to the level of individual items in the display. The limited capacity

WM may be filled with highly weighted nontargets that will have to be attended to (compared to the template) in order to be rejected. Working memory will then have to be emptied and refilled again, thus attention will be shifted from item to item. By contrast, when targets and nontargets are very dissimilar and perceptual grouping among highly similar nontargets can take place, very little attention would be required to detect target items. This is because only highly weighted targets gain access to WM.

The present experiments were designed to explore the relationship between the processing demands imposed on the executive controller by stimulus factors associated with the array and those involved in the maintenance of strategic control. The experimental paradigm requires the subject to search a computer screen for all instances of a target among a large array of non-targets. A light sensitive pen is used to cancel targets once they are found. The manner in which the computer program treats targets once they are detected tends to reinforce the need for a careful strategy. Rather than blanking out a letter or replacing it with a visible marker like a white square once it has been detected, each hit is replaced with a member of the distractor set. Thus, the subject is not provided with any kind of tangible evidence that a particular part of the array has been covered. In order not to go over the same area twice (or even more often), a subject has to remember which part of the array has already been searched. The simplest way to monitor that is to have an overall strategy in mind.

The first two experiments tested the Duncan and Humphreys (1988) similarity model on large arrays when the subjects searched for many instances of one target (Experiment 1) or of two different targets (Experiment 2). The third experiment considered the size of the target set as well as categorical, rather than perceptual relationships among stimuli, as important factors. The overall hypothesis tested throughout this set of experiments is that strategic monitoring of search can be disrupted by increases in the processing demands associated with the experimental procedure. Because the central executive is assumed to have a limited capacity (Baddeley, 1986), increasing the load for one aspect of executive responsibility (e.g. control of processing) could be expected to interfere with another aspect of executive responsibility (monitoring or strategy maintenance). Whether processing demands are manipulated through perceptual relationships among targets and distractors (Experiments 1 and 2) or via memory load and categorical familiarity (Experiments 3A and 3B), it is expected that strategy maintenance will be impaired on those trials in which processing demands are greatest .

## Experiment 1

This experiment was designed to extend Duncan and Humphreys' (1989) findings to large arrays in which several instances of a target must be located. The Duncan and Humphreys model predicts that the greatest search difficulty should be encountered on the trial in which there is high similarity between targets and distractors along with low similarity among distractors. In terms of the process continuum that they have proposed, this condition should require the most controlled processing. This is because the targets cannot easily be distinguished from the distractors and presumably attention must be shifted from item to item. Furthermore, the reduced possibility of spreading suppression among non-targets that cannot be easily grouped, should decrease the likelihood that large clumps of non-targets could be quickly rejected. Target detection times should be longer and number of targets found should be fewer compared to trials on which the targets are easily distinguished from distractors that can be grouped and efficiently rejected. If strategic control is susceptible to the increased processing demands associated with difficult perceptual differentiation and reduced perceptual grouping, the same condition should show evidence of poor strategy maintenance. This should be apparent in terms of a less systematic pattern of target detections.

# Method

## Subjects

Eighteen introductory psychology students volunteered to participate in the experiment as an option in partial fulfillment of course credit. There were nine males and nine females with an average age of 20.8 years (range 18-35).

## Apparatus and Stimuli

Letter stimuli were presented on an Amdek monitor connected to an IBM XT personal computer with high resolution EGA graphics. A light pen (L-PC Lite Pen: The Lite-Pen Co., Los Angeles, CA) was used for cancelling targets. In earlier pilot research, it was found that subjects fatigued easily when they had to hold the pen and touch an upright screen. Participants, therefore, were seated directly in front of the computer monitor set into a table at a $30^{o}$ angle, an arrangement that allowed the person to rest his or her arm on the table in front of the monitor. Subjects viewed the screen from a distance of about 40 cm, although they were free to get closer by leaning forward if they wished.

The letter stimuli used in this series of experiments were designed using the Fontman font generating program. Stimulus letters were each 12 pixels by 15 pixels in size and occupied 5 mm x 7 mm of space on the screen. Letters were designed to conform in shape and outline thickness to those used by van der Heijden, Malhas and van den Roovaart (1984). These authors provided a confusion matrix for uppercase English letters that was used in the current studies for determining specific letter stimuli in each condition of the experiments. The matrix provides a measure of confusability between any two letters. To develop the matrix, the authors presented to subjects a single letter on a computer screen for a brief duration averaging 6.42 msec. A verbal identification of the letter was required. It must be acknowledged that letter identification confusion is far more likely to occur during such brief exposures than in the current experimental situation in which participants were not limited in the time available for inspecting the stimuli. The van der Heijden et al. matrix, however, provided an empirical basis for the selection of stimuli that were highly similar or highly dissimilar. Appendix 1 provides the confusability matrix for the letters used in the first two experiments (from van der Heijden et al., 1984).

Design

Target-nontarget (T-N) and nontarget-nontarget (N-N) similarities were manipulated. Factors included high and low T-N similarity and three levels of N-N similarity: high - identity (one letter), high - homogeneous (two featurally similar letters) and low - heterogeneous (two featurally dissimilar letters). To deal with the possibility that any effects found might be due simply to the uniqueness of the particular target letter, three different letters were tested in each T-N condition. All letters in the high T-N condition had the same distractors for comparable cells as did the three letters used in the low T-N condition. An effort was made to ensure that the three letters used in each condition had similar confusability ratings (van der Heijden et al., 1984).

Thus, the design was 2 (T-N Similarity) x 3 (N-N Similarity) x 3 (Letters) with all factors within subjects. Table 1 illustrates the design and stimuli used.

Subjects completed three blocks of three trials, one for each level of N-N similarity. Three orders were used for a counterbalanced presentation of the blocks. For half the subjects, trials in each block were presented in one sequence and the other half received trials presented in a different sequence. Both sequences alternated T-N high similarity trials with T-N low similarity trials. Each test session lasted approximately 50 minutes.

Table 1

Experimental design used in Experiment 1

| | | N-N Similarity | | |
|---|---|---|---|---|
| | | Identity | High | Low |
| T-N Low | (1) | X / O | X / O + Q | X / O + I |
| | (2) | K / O | K / O + Q | K / O + I |
| | (3) | H / O | H / O + Q | H / O + I |
| T-N High | (1) | P / R | P / R + E | P / R + D |
| | (2) | B / R | B / R + E | B / R + D |
| | (3) | S / R | S / R + E | S / R + D |

Note: Letter to the left of the slash (/) is the target. Letter(s) to the right are non-targets.

The experimental method to be described below was used throughout this series of experiments. Any variations will be specifically outlined for each experiment.

Procedure

Prior to a trial, subjects were told which letter(s) to search for in the upcoming array. The target letters remained on the screen for approximately twelve seconds followed by a blank screen for ten seconds before the array appeared. Subjects thus had a verbal and visual presentation of the target letter(s) and adequate rehearsal time before the trial began.

On those trials with target sets greater than one item, subjects were told to search concurrently for all instances of all members of the set. That is, they were not to search for all instances of one target letter and then all instances of the next. Speed and accuracy were stressed equally in the instructions. Subjects were told to inform the experimenter once they believed they had found all the targets, at which time the trial ended. They were never told how many targets were in the array nor were they informed of the number of targets successfully found.

Practice with the light-pen was given on an array, smaller in overall size than test arrays. Subjects searched for 10 instances of a single target letter (N) embedded within distractors drawn from the remainder of the alphabet.

On test trials, a randomly distributed array of 288 white letters was presented on a black screen. With the light pen held in the preferred hand, subjects searched for and cancelled all instances (36) of a specified target or targets. Items touched with the pen were replaced by a member of the distractor set. The screen locations for the 288 letters remained the same for all arrays. Six different patterns of target locations were determined, each of which was carefully chosen to ensure that nine target locations were evenly distributed within each quadrant of the array. These different target location patterns were balanced over trials within each experiment in order to ensure that subjects were unable to develop a memory for specific target location patterns over several trials.

Dependent Measures

Time. The program recorded time taken between letter detections and did not begin the timing until after the first one, thus the time associated with the first detection was zero. For this reason, the first hit was not included in any of the analyses. Median inter-target detection times were analyzed. Medians were chosen rather than means because, as is common with most reaction time (RT) research (i.e. McCormack & Wright, 1964), the means were positively skewed. The median represents a measure of central tendency that is less influenced by extreme scores than is the mean (Glass & Hopkins, 1984).

Number of targets detected. Often subjects would complete a search trial and then do another scan over the array looking for missed targets, usually finding two or three more before indicating that the search was over. The last three hits were deleted from the data files prior to analysis because these latter hits generally required longer than average times and their locations did not necessarily fit into the search pattern developed by the subject over the rest of the array. Therefore, when the deletion of the first untimed hit also is taken into account, all analyses were run on data files in which a maximum of 32 targets had been detected.

Systematicity of search. When people search these kinds of letter arrays for specified targets, they tend to exhibit fairly specific detection patterns. These patterns may look like horizontal or vertical sweeps over the array; sweeps which may be relatively narrow or broad in terms of their width. Occasionally, some subjects show evidence of having searched back and forth or up and down within each of the four quadrants of the array. When the individual consistently follows a particular pattern from the beginning of the search to the end, it can be classified as a highly systematic search. A highly unsystematic search pattern, on the other hand, is one in which there is no evidence of any pattern and the detections appear to have been made in a completely random order suggesting that no particular search strategy had been utilized. Between these two extremes are patterns which demonstrate varying amounts of systematicity that are not, however, consistent throughout the array. It is therefore possible to quantify a range of search systematicity from completely unsystematic to highly systematic.

In order to facilitate analysis of search systmaticity, target detections for each trial per subject were coded in three different ways. One way coded hits in terms of the quadrant of the array in which it occurred; another coded in terms of four equally spaced horizontal bands from top to bottom of the array and one coded in four vertical bands across the array. This system of coding provided a sequential list of detections within four possible locations. Thus, a sequence like 1 1 1 1 1 1 1 1 2 2 2 2 2 2 2 2 3 3 3 3 3 3 3 3 4 4 4 4 4 4 4 4 for hits coded using the horizontal index, indicates a highly systematic search pattern that was conducted in sweeps back and forth across the array. A particular trial that yielded such a highly systematic sequence on the horizontal dimension would appear to be considerably less systematic on the vertical and quadrant coding dimensions.

A descriptive, noninferential statistic, $H_2$, was used to quantify the sequential nature of target detections[1]. This statistic has been used to analyze serial dependencies in the behaviour of rats learning a bar press response (Frick & Miller, 1951; Miller & Frick, 1949) and variability in rats' sequential choices of arms in the radial maze (Loh & Beck, 1989). When used to analyze a search coded by quadrants, for example, the statistic $H_2$ provides a measure of the uncertainty of pairs of successive quadrant choices minus the uncertainty of pairs of successive individual quadrant choices; $H = (P_1 \log_2 1/P_1)$. $P_1$ represents the probability assigned to each of the four quadrant alternatives and is calculated by dividing the frequency of a quadrant choice by the frequency of all choices. In the present study, the inverse of $H_2$ was used to represent a measure of systematicity associated with each search trial. A high score on this measure indicates a very

1 I am grateful to Dr. C. H. M. Beck for providing the computer program to conduct this analysis.

predictable series of quadrant choices and thus represents a highly systematic pattern. The H2 statistic was calculated for each of the three types of coding (horizontal, vertical and quadrant) for every individual search trial. The largest of the three indices of systematicity was used as the dependent measure in this series of experiments. Appendix 2 provides examples of different search patterns as well as the associated systematicity statistics.

Data analyses

Prior to analysis, the data were inspected and corrected for problems due to difficulties with the light pen and distractors that had been chosen in place of intended targets. The light pen is not a very sensitive instrument and if not held so that the pen rests directly on the screen and perpendicular to it, the appropriate response may not occur. A loud buzzing noise indicated to the subject that the pen had not been appropriately activated whereas a 'beep' and the replacement of the letter by another letter indicated a complete hit. The occurrence of persistent buzzing was noted by the experimenter who recorded the efforts by the subject to activate the pen by noting the location within the array in which such difficulties occurred. This made it possible to determine that inordinately long times for particular hits were due to problems activating the pen rather than search difficulties. All data files (one for each trial for each subject) were inspected for these pen problems and times associated with these were replaced by the median time for that data file. In Experiment 1, 71 such times were replaced, representing .6% of the approximately 11,540 hits[2] over the whole experiment.

Subjects often chose a letter adjacent to a legitimate target and then immediately chose the appropriate target. Because the correct target was obviously intended in such cases, data files were adjusted to delete these inadvertent hits (false alarms). One hundred three such pen errors were deleted (.9% of total hits). Occasionally subjects hit a distractor but did not hit the target adjacent to it. It was assumed that the subject had really intended to hit the target letter; 23 such errors were adjusted to indicate a target hit (.2% of total hits). Actual commission errors (selection of a distractor letter) were only considered to have occurred when a distractor hit was not followed immediately by an adjacent target hit or was not adjacent to a target even if that target was not hit. Seven commission errors occurred with no subject making more than one such error over the 18 test trials. Because there were so few commission errors, no analyses were conducted on these data.

Analyses of variance (ANOVA) for repeated measures using the multivariate procedures of Systat 3.1 were conducted on all dependent measures. *Post-hoc* contrasts conducted on interactions were evaluated against alpha levels adjusted using the Bonferroni correction (Rosenthal & Rosnow, 1991). To calculate the adjusted

---

2 This figure was arrived at by calculating the average number of hits per trial and multiplying that by the number of subjects then by the number of trials; the number of deleted hits was added to this total.

alpha level, the experiment-wise alpha of .05 was divided by the number of contrasts conducted. Only those contrasts that were considered most likely to explain any particular interaction were undertaken. Note that separate variances were used in computing these contrasts.

Results

Time. A 2 (T-N Similarity) x 3 (N-N Similarity) ANOVA for the time measure was conducted on the median inter-target times for each subject[3]. The main effect of T-N similarity was reliable, $F$ (1, 17) = 163.63, $p$ = .000, due to the longer times taken when T-N similarity was high compared to when it was low. The main effect of N-N similarity was not reliable, Wilks' Lambda = .967, multivariate $F$ (2, 16) = .276, $p$ = .76. There was a reliable interaction between T-N similarity and N-N similarity, Wilks' Lambda = .616, multivariate $F$ (2, 16) = 4.996, $p$ = .021. As Figure 1 illustrates, times appeared to decline in the T-N High condition but increased in the T-N Low condition. Two contrasts (Bonferroni alpha = .025) were computed that compared N-N Identity to N-N Low for each of the two T-N conditions. Within the T-N High condition, there was no reliable difference due to manipulations to N-N similarity, $F$ (1, 17) = 2.178, $p$ = .158. The interaction was due to longer detection times when N-N similarity was homogeneous

---

3 The separate factor of letter was not included in any analyses because differences among the various letter stimuli within each T-N condition were not of interest here. Although there were some differences among the results for letters within the T-N high condition, the pattern was consistent for all dependent measures.

(Identity) compared to when it was heterogeneous (N-N Low) within the T-N Low condition, $F$ (1, 17) = 7.638, $p$ = .013. Table 2 provides cell means for all trials in Experiment 1.

Table 2

Experiment 1: Mean Times in Seconds With (Standard Deviations)

| | N-N Similarity | | | |
|---|---|---|---|---|
| | Identity | High | Low | Means |
| T-N Low | 0.645 (.086) | 0.691 (.107) | 0.687 (.094) | 0.674 |
| T-N High | 1.195 (.213) | 1.155 (.231) | 1.134 (.221) | 1.160 |
| Means | 0.920 | 0.923 | 0.911 | |

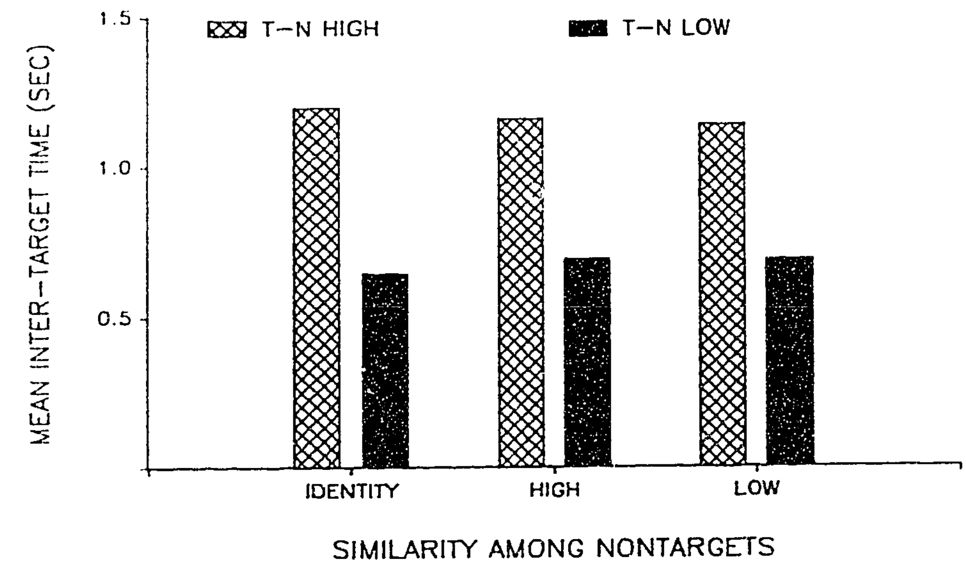



Figure 1. Experiment 1: Mean inter-target time for T-N Similarity High and Low for three levels of similarity within non-targets

Number of targets detected. A reliable main effect of T-N similarity was found, $F$ (1, 17) = 37.91, $p$ = .000 with subjects detecting fewer targets in the T-N high condition than in the T-N low condition. The effect of N-N similarity was not reliable, Wilks' Lambda = .795, multivariate $F$ (1, 17) = 2.059, $p$ = .16. The interaction was not reliable, Wilks' Lambda = .746, multivariate $F$ (1, 17) = 2.73, $p$ = .096. Cell means are provided in Table 3.

Table 3

Experiment 1: Mean Number of Target Detections With (Standard Deviations)

| | N-N Similarity | | | |
|---|---|---|---|---|
| | Identity | High | Low | Means |
| T-N Low | 31.99 (.013) | 31.96 (.107) | 31.94 (.171) | 31.97 |
| T-N High | 30.28 (1.31) | 30.83 (1.10) | 30.70 (1.10) | 30.60 |
| Means | 31.14 | 31.40 | 31.32 | |

Systematicity of Search. The interaction between N-N similarity and T-N similarity, illustrated in Figure 2, was reliable, Wilks' Lambda = .587, multivariate $F$ (2, 16) = 5.62, $p$ = .014. Inspection of the figure indicates that T-N similarity had no effect on search systematicity when non-targets were heterogeneous (N-N Low) whereas when non-targets were similar (T-N Identity and T-N High), the level of similarity between targets and non-targets was important. Two *post hoc* contrasts (Bonferonni alpha = .025) were conducted to determine the source of the interaction. A first contrast comparing the two levels of T-N similarity when non-targets were similar but not homogeneous (N-N High) indicated no reliable differences, $F$ (1, 17) = 1.196, $p$ = .29. A second contrast determined that the interaction was due to subjects' use of less systematic search strategies to find targets embedded among homogeneous distractors (N-N Identity) when T-N similarity was high than when it was low, $F$ (1, 17) = 6.453, $p$ = .021.

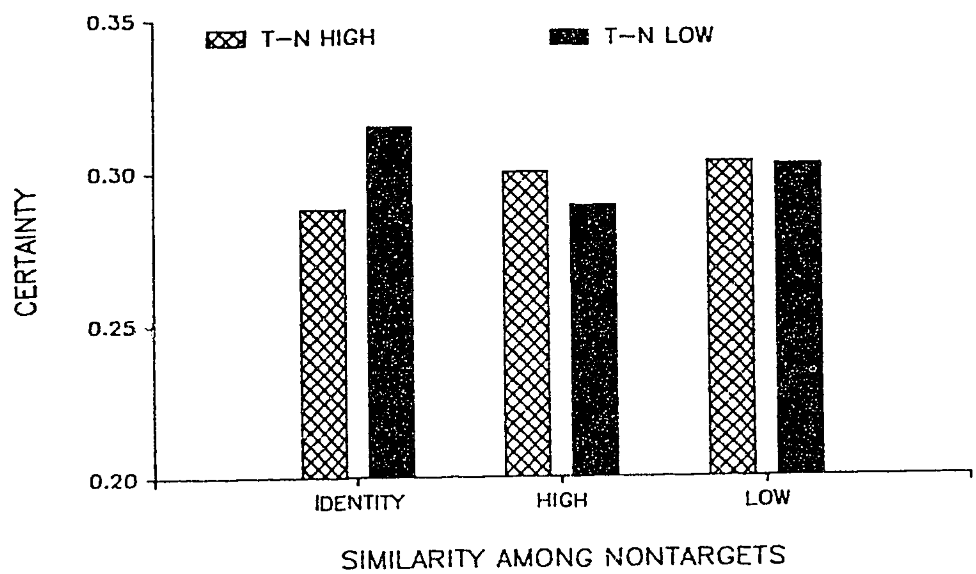



Figure 2. Experiment 1: Mean certainty scores for T-N Similarity High and Low for three levels of similarity within targets.

The main effect of T-N similarity was not reliable, $F$ (1, 17) = .388, $p$ = .54, nor was the effect of N-N similarity, Wilks' Lambda = .908, multivariate $F$ (2, 16) = .812, $p$ = .46. Cell means are provided in Table 4.

Table 4

Experiment 1: Mean Certainty Scores With (Standard Deviations)

| | N-N Similarity | | | |
|---|---|---|---|---|
| | Identity | High | Low | Means |
| T-N Low | 0.315 (.025) | 0.289 (.032) | 0.302 (.033) | 0.302 |
| T-N High | 0.288 (.045) | 0.300 (.037) | 0.303 (.037) | 0.297 |
| Means | 0.302 | 0.295 | 0.303 | |

Discussion

The main effect of T-N similarity that is apparent for both time and accuracy measures follows the predictions of Duncan and Humphreys (1989) and is consistent with data from other studies in which high similarity among targets and non-targets is associated with inefficient search for a single target (Duncan, 1989; Humphreys et al., 1989; Treisman & Gormican, 1988). The model predicts that the poorest performance should take place when there is high similarity between targets and non-targets *together with* low similarity among the non-targets. There was an interaction between T-N similarity and N-N similarity on the time measure, but the interaction was not in the direction predicted by the model. It was due to the effect of non-target heterogeneity when targets were dissimilar rather than similar to non-targets.

The paradigm employed in this series of experiments does not allow a comparison to a target absent condition for an analysis of differences in slope. In traditional visual search experiments such comparisons made it possible to draw conclusions regarding the type of processing (controlled or automatic) associated with a particular stimulus display. On the strength of prior research with the traditional paradigm it is reasonable to presume that high T-N similarity requires controlled processing. It was hypothesized in the introduction that when processing demands are high, strategic monitoring of search would be compromised also. The reliable main effect of T-N similarity that occurred on the measures of time and accuracy should therefore have been accompanied by a similar effect on search systematicity. Such a main effect did not occur although systematicity did interact reliably with N-N similarity.

The results indicated that maintaining a systematic strategy was difficult when subjects searched for targets highly similar to non-targets and appeared to be very easy when targets were dissimilar, but only in the N-N identity condition. When targets and non-targets are highly similar, both kinds of stimuli are readily activated during the search and compete for entry into working memory. This could necessitate a serial search through each item in the display. When non-targets are all the same (N-N identity), targets are easily camouflaged and greater effort may be required to distinguish them from distractors. A process may be at work here that is analogous to Duncan and Humphreys' concept of spreading suppression of weight linkage within perceptually homogeneous groups of letters. In this case, rather than allowing for easy rejection of strongly grouped non-targets that share a similar reduced weight, the groups, which may include targets along with their highly similar non-targets, are linked in terms of *increased weight*. Alternatively, when T-N similarity was low, spreading suppression may well have been the mechanism that facilitated the rejection of grouped non-targets thus expediting strategy maintenance. In both cases, the concept of perceptual grouping provides a tenable explanation for the findings. Furthermore, the results imply that homogeneity among non-targets may be more detrimental to maintenance of search strategy than heterogeneity. When nontarget homogeneity was reduced (N-N high and low conditions), strategy use was essentially the same regardless of the amount of T-N similarity.

The apparent trend here for search strategy to be less systematic when attentional demands are high, lends some support to the hypothesis that control of processing and monitoring of search strategy make demands on the same limited capacity controller. The next experiment increased processing demands further to allow additional exploration of target-nontarget relationships in terms of the Duncan and Humphreys' (1989) model.

## Experiment 2

Experiment 2 was designed to test the influence on search efficiency of manipulations to similarity relationships among letters in the target set. Duncan and Humphreys (1989) predicted that heterogeneous target items would require more elaborate templates with a resulting increased possibility that attributes of the target template would be shared by non-targets. Thus, there would be a greater likelihood that non-targets would have high weights leading to more frequent entry into working memory. Because there are considerably more non-targets than targets in the array, targets are never in close proximity. Grouping of targets, therefore, could not aid in detection. Similarly, their model predicts that heterogeneity among non-targets could be expected to be more detrimental to search than heterogeneity among targets because a reduction in grouping of non-targets means large clumps could not be easily rejected. Thus, in terms of the Duncan and Humphreys model, heterogeneity among targets is unlikely to interact with either T-N or N-N similarity.

It is probable that increasing the number of targets to be searched for from one to two would result in an overall increase in search time either because of the increased complexity of the template (Duncan & Humphreys, 1989) or because of the extra comparison process required (Sternberg, 1966). No matter which explanation is accepted, the question of interest is whether the increased processing demand is great enough to interfere with strategic control.

## Method

### Subjects

Twelve students participated in this experiment for an optional credit in an introductory Psychology course. There were four females and eight males with an average age of 20.3 years (range 18-37).

### Design

The design was 2 (T-T Similarity: High and Low) x 2 (T-N Similarity) x 3 (N-N Similarity). As can be seen in Table 5, the stimuli used as targets in Experiment 2 matched those used in Experiment 1. The same method of balancing trial presentation as used in Experiment 1 was used in this experiment.

Table 5

Experimental design for Experiment 2

| | | N-N Similarity | | |
|---|---|---|---|---|
| | | Identity | High | Low |
| T-T(2) High | T-N Low | X + K / O | X + K / O + Q | X + K / O + L |
| | T-N High | P + B / R | P + B / R + E | P + B / R + D |
| T-T(2) Low | T-N Low | X + H / O | X + H / O + Q | X + H / O + L |
| | T-N High | P + S / R | P + S / R + E | P + S / R + D |

Note: Letters to the left of the slash (/) are targets. Letter(s) to the right are non-targets.

Results

As in Experiment 1, data files were inspected and corrected for pen errors and errors of target intention. Fifty-one detection times were replaced by file median times because of difficulties with the pen. This represents 1.06% of the 4,811 hits over the whole experiment. Thirty inadvertent hits (.62%) were deleted because they were immediately followed by an adjacent correct detection. Nineteen (.4%) detections of non-targets adjacent to an undetected actual target were assumed to have been intended as correct hits and the data were adjusted accordingly. Five commission errors occurred in this data set and were not analyzed further. As in Experiment 1, 32 of the possible 36 hits have been included in the analyses.

Time. A 2 (T-T Similarity) x 2 (T-N Similarity) x 3 (N-N Similarity) ANOVA on median times resulted in a reliable main effect of T-N similarity, $F$ (1, 11) = 218.984, $p$ = .000. Subjects took longer to find targets when T-N similarity was high ($M$ = 1.66 s) than when it was low ($M$ = .66 s). The main effect of N-N similarity was reliable, Wilks' Lambda = .527, $F$ (2, 10) = 4.492, $p$ = .041. Times increased as N-N similarity decreased. There was no reliable interaction between T-N similarity and N-N similarity, Wilks' Lambda = .835, multivariate $F$ (2, 10) = .985, $p$ = .41, nor between N-N similarity and T-T similarity, Wilks' Lambda = .960, multivariate $F$ (2, 10) = .208, $p$ = 82. The main effect of T-T similarity was not reliable, $F$ (1, 11) = 1.940, $p$ = .19. The interaction between T-T similarity and T-N similarity was not reliable, $F$ (1, 11) = 3.559, $p$ = .09 nor was the three-way interaction, Wilks' Lambda = .868, multivariate $F$ (2, 10) = .757, $p$ = .49. Table 6 provides cell means for inter-target detection times in Experiment 2.

Table 6

Experiment 2: Mean Times in Seconds With (Standard Deviations)

| | | N-N Similarity | | | |
|---|---|---|---|---|---|
| | | Identity | High | Low | Means |
| T-T (2) High | T-N Low | 0.582 (.122) | 0.649 (.135) | 0.710 (.178) | 0.647 |
| | T-N High | 1.610 (.313) | 1.770 (.347) | 1.779 (.456) | 1.720 |
| T-T (2) Low | T-N Low | 0.615 (.091) | 0.683 (.136) | 0.702 (.128) | 0.667 |
| | T-N High | 1.426 (.342) | 1.647 (.553) | 1.743 (.431) | 1.605 |
| | Means | 1.058 | 1.187 | 1.234 | |

Number of targets detected. Subjects found fewer targets when T-N similarity was high ($M$ = 27.03) than when T-N similarity was low ($M$ = 31.81), $F$ (1, 11) = 32.577, $p$ = .000. The interaction between T-T similarity and T-N similarity was reliable, $F$ (1, 11) = 5.205, $p$ = .043. The main effects of N-N similarity, Wilks' Lambda = .632, multivariate $F$ (2, 10)= 2.913, $p$ = .10 and T-T similarity, $F$ (1, 11) = 2.325, $p$ = .16 were not significant. Also not reliable were the other three interactions: N-N x T-N, Wilks' Lambda = .676, multivariate $\underline{F}$ (2, 10) = 2.399, $\underline{p}$ = .14; N-N x T-T, Wilks' Lambda = .666, multivariate $\underline{F}$ (2, 10) = 2.512, $\underline{p}$ = .13; and T-N x N-N x T-T, Wilks' Lambda = .692, multivariate $\underline{F}$ (2, 10) = 2.226, $\underline{p}$ = .16. Table 7 provides cell means.

Table 7

Experiment 2: Mean Number of Targets Detected With (Standard Deviations)

| | | N-N Similarity | | | |
|---|---|---|---|---|---|
| | | Identity | High | Low | Means |
| T-T (2) High | T-N Low | 31.99 (.029) | 31.99 (.029) | 31.92 (.289) | 31.97 |
| | T-N High | 27.08 (4.23) | 27.25 (4.14) | 24.25 (6.30) | 26.19 |
| T-T (2) Low | T-N Low | 31.75 (.622) | 31.75 (.452) | 31.42 (.515) | 31.64 |
| | T-N High | 27.25 (3.22) | 28.67 (1.76) | 27.67 (3.03) | 27.86 |
| | Means | 29.52 | 29.92 | 28.82 | |

Systematicity of search. The ANOVA on the systematicity measure yielded no significant differences on any factor. There were no reliable main effects: T-T, $\underline{F}$ (1, 11) = .678, $\underline{p}$ = .43; T-N, $\underline{F}$ (1, 11) = 1.54, $\underline{p}$ = .24; and N-N, Wilks' Lambda = .842, multivariate $\underline{F}$ (2, 10) = .938, $\underline{p}$ = .42. None of the interactions were reliable: T-T x T-N, $\underline{F}$ (1, 11) = .022, $\underline{p}$ = .89; T-T x N-N, Wilks' Lambda .733, multivariate $\underline{F}$ (1, 11) = 1.824, $\underline{p}$ = .21; T-N x N-N, Wilks' Lambda = .895, multivariate $\underline{F}$ (2, 10) = .584, $\underline{p}$ = .58; and T-T x T-N x N-N, Wilks' Lambda = .999, multivariate $\underline{F}$ (2, 10) = .005, $\underline{p}$ = .10. Table 8 provides cell means.

Table 8

Experiment 2: Mean Certainty Scores With (Standard Deviations)

| | | N-N Similarity | | | |
|---|---|---|---|---|---|
| | | Identity | High | Low | Means |
| T-T (2) High | T-N Low | 0.300 | 0.280 | 0.281 | 0.287 |
| | | (.067) | (.039) | (.047) | |
| | T-N High | 0.303 | 0.297 | 0.298 | 0.299 |
| | | (.047) | (.047) | (.032) | |
| T-T (2) Low | T-N Low | 0.292 | 0.305 | 0.278 | 0.292 |
| | | (.058) | (.056) | (.060) | |
| | T-N High | 0.296 | 0.326 | 0.298 | 0.307 |
| | | (.039) | (.043) | (.051) | |
| | Means | 0.298 | 0.302 | 0.289 | |

Discussion

T-T similarity interacted with T-N similarity on the accuracy measure, apparently because subjects had difficulty finding two targets among similar distractors, but only when the targets themselves were highly similar. Because manipulations to target-target similarity did not influence detection time or search systematicity, it is probably reasonable to conclude that T-T similarity had no substantial influence on search efficiency. The reliable effect of T-N similarity for both time and accuracy is not surprising given the strength of the same effect in Experiment 1. The interaction between T-N similarity and N-N similarity failed to reach significance for all three dependent measures. Thus, the results of this experiment as well as the first one, do not support the extreme position of the Duncan and Humphreys model that predicts the greatest search difficulty when there is high similarity between targets and distractors together with low similarity among distractors. Other research that has demonstrated the negative effects of heterogeneity among non-targets has used stimuli for which the shared and conjoined features can be more precisely defined than is possible with the different letter stimuli used in the present series (i.e. colour patches, Duncan, 1989; letter 'T' among Ts of various orientations, Humphreys et al., 1989). Thus, disparate amounts of similarity for different conditions within the present set of experiments may have made it difficult to test the model sufficiently.

The main effect of N-N similarity on the time measure is in accord with other research suggesting that decreasing similarity among non-targets disrupts the grouping process, thereby reducing the possibility that large sets of non-targets can be rapidly rejected (Humphreys et al., 1989). It is not clear why decreases in N-N similarity significantly influenced search time in this experiment but not in the first study. If the two experiments are considered together, however, there appears to be an interaction between N-N similarity and number of items in the target set. Reductions in similarity among non-targets only influence search time when there are two items in the target set regardless of whether or not targets are perceptually similar to non-targets. Such an interpretation is merely speculative, however, and must be considered with caution as an analysis of an interaction between two experiments is not possible.

Search systematicity was not affected by any of the manipulations in this experiment. This is somewhat surprising given the reliable interaction between T-N similarity and N-N similarity that occurred in the first experiment. The same manipulations were present here except that two targets were to be found rather than one. Presumably, the addition to the target set increased processing demands - the average intertarget time in Experiment 2 ($\underline{M}$ = 1.16 s) was longer than in Experiment 1 ($\underline{M}$ = .92 s). If the additional memory load did not increase processing demands enough to interfere with strategy maintenance, was there some change in the perceptual features of the test displays that might account for the loss of an effect that was reliable in the previous experiment? The nature of the particular stimuli used in this research may provide a possible explanation. The letters P, B and S were chosen because they have been found to be easily confused with the distractors R, E and D (van der Heijden et al., 1984), although the relative amount of confusability varies. The interaction between T-N and N-N on the systematicity measure in

Experiment 1 appeared to be due to the difficulty associated with distinguishing Ps, Bs or Ss from homogeneous displays of Rs. (See Appendix 3 for data on letters used in Experiment 1). Given that interaction, it was reasonable to expect the same result in Experiment 2. In the second experiment, however, two different letters, P and B or P and S were embedded among the Rs. Because targets were never in close proximity to one another, changes in grouping of targets is not a probable explanation for the differing results. In Experiment 1 it was demonstrated that the letter P was consistently more difficult to find than the letters B or S (see Appendix 3). Subjects found fewer Ps than either Bs or Ss when searching a homogeneous display of Rs. Thus, it is possible that in Experiment 2, when searching a homogeneous display of Rs, subjects had more difficulty finding half the targets (Ps) than the other half (Bs or Ss). If that were so, any effect of T-N similarity may have been diluted. In Experiment 2, the search for P and B yielded mean targets of 14.92 (P) and 16.33 (B) while the search for P and S yielded means of 13.50 (P) and 17.67 (S). (Means are based on a possible maximum of 18 for each letter).

The results of Experiments 1 and 2 replicated other research that demonstrates the influence on search difficulty of perceptual similarity between targets and distractors, but provides limited support for the impact of nontarget heterogeneity. Indeed, the data suggest that heterogeneity only exerts an influence when the target set is larger than one. Although there is some evidence for interference between the processing demands of the search and maintenance of strategy, it would seem that the task demands imposed by the perceptual and memory load manipulations of Experiments 1 and 2 were not great enough to tax the full extent of the executive controller's capacity. The last experiment in this series was designed to impose more extensive processing demands in order to determine if strategic control can be compromised.

## Experiments 3A and 3B

The first two experiments demonstrated that perceptual similarity between targets and distractors influences the attentional demands of visual search through large arrays for several targets. Increasing the number of items in the target set from one to two increased the time to find targets, possibly due to a corresponding increase in the number of comparisons to the targets held in memory each time a stimulus item was attended. Further additions to the size of the target set should correspondingly add to the processing demands by increasing the number of required comparisons and also by making greater demands in terms of rehearsal of the memory set. It could be expected, therefore, that the processing demands associated with large memory sets should interfere with search time and accuracy and probably with maintenance of search strategy, particularly if the targets and non-targets are perceptually similar.

In addition to discrimination on the basis of perceptual features, however, targets and non-targets have been shown to be easily differentiated in terms of categorical dimensions such as alphanumeric or semantic class (see Rabbitt, 1978 for a review).

One such categorization is in terms of sequential familiarity. The use of adjacent digits (Egeth, Marcus & Bevan, 1972) or letter sequences like ABCD (Thomas, Waugh & Fozard, 1978) as target sets has resulted in faster search times than target sets of less familiar sequences. When perceptual similarities are high between non-targets and sequenced targets, high input weights of both targets and non-targets should result in heavy competition for entry into working memory. Once in WM, however, the comparison and rehearsal demands should be reduced for familiar sequences and it is this reduction in processing that results in faster search times. Similarly, the scheduling responsibilities of the executive controller should be lessened with familiar target sequences and thus, even if the size of the target set is large, strategic control should not be impaired. Experiment 3A tested search efficiency and strategy use for several target set sizes using familiar sequences of letters and nonsense strings of letters. non-targets included all the remaining letters of the alphabet. Limitations in the amount of perceptual grouping of non-targets and considerable featural overlap between targets and members of the distractor set ensured that controlled search through the array was necessary.

If familiar sequences do indeed facilitate search, a question of methodological importance for research with large target sets is whether people spontaneously reorganize strings of letters into meaningful sets (i.e. reorganize the string 'bmi' to 'ibm'). If they do so, they might be expected to search for such letters faster than if the string could not be reorganized. In order to determine if people do spontaneously reorganize letter sequences into more familiar strings, subjects in Experiment 3B received the same familiar sequences and nonsense strings as used in 3A. In that experiment, however, the familiar sequence was presented in a different, unfamiliar order.

Fifty undergraduates, participating in a different experiment, rated the familiarity and ease of memory of several letter strings that included alphabetic sequences like ABCDEF and familiar acronyms like GST[4] and RSVP. Ratings were made on unorganized versions of the same strings and on strings that were believed to have no sequence value at all. Familiar sequences with the highest ratings and unfamiliar strings with the lowest ratings were used as stimuli.

## Method

### Subjects

Twelve introductory Psychology students participated in each of the two experiments. There were six females and six males with an average age of 20.2 years (range 18-25) in Experiment 3A. In Experiment 3B seven females and five males with an average age of 20.3 years (range 18-24) participated.

---

4 GST is the acronym for the recently introduced Goods and Services sales tax that has been the focus of considerable public discussion over the past two years.

Design and procedure

Each experiment employed a 2 (String Familiarity) x 5 (Target Set Size) design with both factors within subjects. In Experiment 3A, half the subjects were presented with a block of trials consisting of a string of familiar sequenced letters followed by a block of random letter trials. The other half of the subjects received the reverse order. Within each block, three trial orders were used. The procedure followed that used in the previous experiments with one exception. After each trial, subjects were asked to write down the letters that had been searched for in the trial. This provided an indication of whether or not subjects had spontaneously reorganized the letters. The procedure and stimuli in Experiment 3B were the same as in 3A except that the familiar string of letters was presented in an unfamiliar order. Table 9 illustrates the designs and the stimuli used.

Table 9

Letter stimuli and design: Experiment 3

| | Experiment 3A | | Experiment 3B | |
|---|---|---|---|---|
| Set Size | Familiar Organized | Random | Familiar Unorganized | Random |
| 2 | VD | XH | DV | XH |
| 3 | IBM | KQV | MBI | KQV |
| 4 | RSVP | PXTU | PVSR | PXTU |
| 5 | LMNOP | NWUED | NPMLO | NWUED |
| 6 | ABCDEF | CJMSFY | DFAECB | CJMSFY |

Results: Experiment 3A

As in the first two experiments, data files were inspected and corrected for pen errors and errors of target intention. Seventy-five detection times were replaced by file median times because of difficulties with the pen. This represents 1.9% of the approximately 3,956 hits over the whole experiment. Twenty-eight nontarget detections (.7%) were deleted because they were immediately adjacent to a correct detection and were thus assumed to be unintentional hits. Four (.1%) detections of non-targets adjacent to an undetected actual target were assumed to have been intended as correct hits and the data were adjusted accordingly. Five commission errors occurred in this data set and were not analyzed further. As in Experiments 1 and 2, 32 of the possible 36 hits have been included in the analyses.

It is possible that subjects who received the block of familiar letter sequences prior to random letter strings might have been biased to attempt to try to find something familiar about the random strings. To test this possibility, order of trial blocks was included as a factor in all analyses. In those cases in which order had no effect, the analysis was re-run excluding that factor.

Time. A 2 (Block Order) x 2 (String Familiarity) x 5 (Target Set Size) ANOVA on the median times indicated that order had no influence on detection times. In an analysis excluding order, the interaction between set size and string familiarity was found to be reliable, Wilks' Lambda = .014, multivariate $F$ (4, 8) = 138.29, $p$ <

.000. *Post hoc* contrasts (Bonferonni alpha = .025) indicated that the difference between familiar and random strings was reliable for target sets of two letters, $F$ (1, 11) = 24.57, $p$ = .000, and sets of six letters, $F$ (1, 11) = 184.25, $p$ = .000. The following two main effects should be interpreted with caution in light of the reliable interaction that is illustrated in Figure 3. Subjects found targets faster when the string of letters formed a familiar sequence than when they searched for a random string of letters, $F$ (1, 11) = 35.77, $p$ = .000. The number of letters in the target set significantly influenced search times, Wilks' Lambda = .082, multivariate $F$ (4, 8) = 22.54, $p$ = .000. Times increased linearly as set size increased from 2 to 6 letters. Cell means are provided in Table 10.

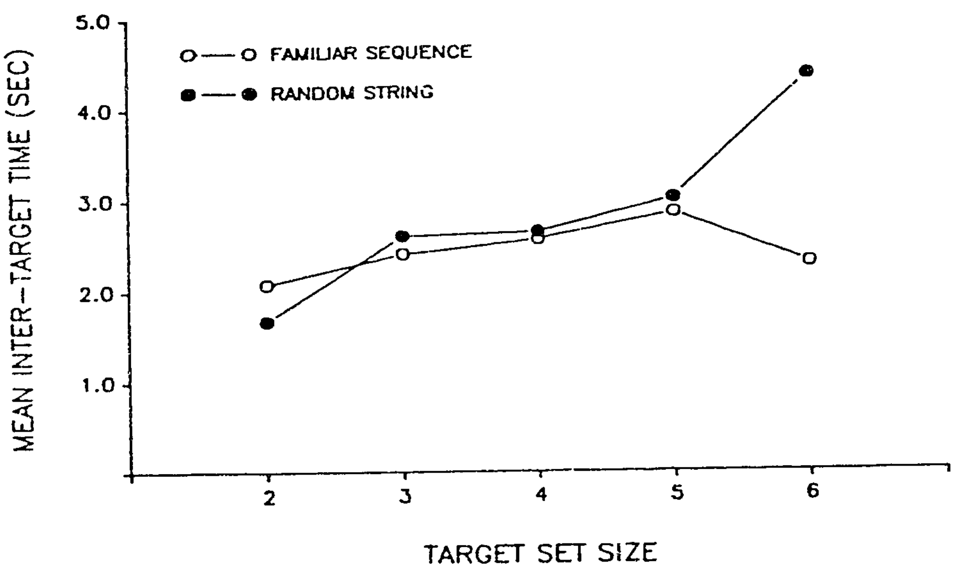



Figure 3. Experiment 3a: Mean inter-target times for familiar sequences and random strings of letters over five target set sizes.

Table 10
Experiment 3A: Mean Times in Seconds With (Standard Deviations)

| | Familiar Organized | Random | Means |
|---|---|---|---|
| Set Size | | | |
| 2 | 2.075 (.449) | 1.668 (.535) | 1.872 |
| 3 | 2.408 (.542) | 2.604 (.823) | 2.506 |
| 4 | 2.565 (.754) | 2.645 (.661) | 2.605 |
| 5 | 2.849 (.944) | 3.008 (.960) | 2.929 |
| 6 | 2.293 (.752) | 4.358 (1.09) | 3.326 |
| Means | 2.438 | 2.857 | |

Number of targets detected. The 2 (Block Order) x 2 (String Familiarity) x 5 (Target Set Size) ANOVA revealed no order effects. An analysis collapsed over order indicated that familiarity interacted significantly with set size, Wilks' Lambda = .210, multivariate $F$ (4, 8) = 7.51, $p$ = .008. *Post hoc* contrasts (Bonferonni alpha = .017) indicated that the familiarity factor significantly influenced accuracy for the three letter set, $F$ (1, 11) = 13.80, $p$ = .003 and for six letters, $F$ (1, 11) = 14.72, $p$ = .003. As in the analysis of time, the significant effect in the two letter set was in the other direction, $F$ (1, 11) = 9.066, $p$ = .012. The interaction, illustrated in Figure 4, should be taken into account when considering the significance of the following main effects. More targets were detected when the letters were presented in a familiar sequence than when the letters constituted a random string, $F$ (1,11) = 4.94, $p$ = .048. There was a main effect of set size, Wilks' Lambda = .219, multivariate $F$ (4, 8) = 7.142, $p$ = .009. The cell means are presented in Table 11.

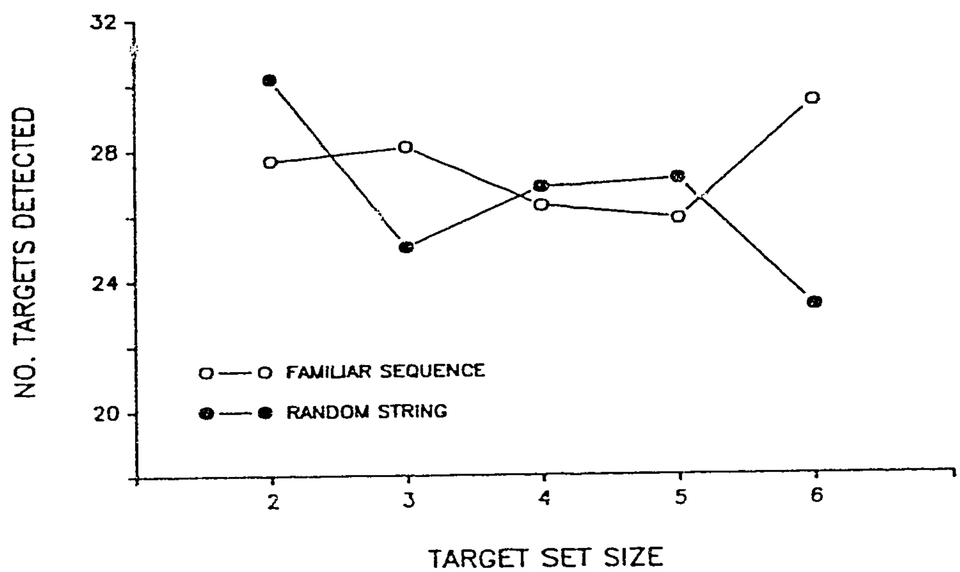



Figure 4. Experiment 3A: Mean number of targets detected for familiar sequences and random strings of letters over five target set sizes.

Table 11
Experiment 3A: Mean Number of Target Detections
With Standard Deviations)

| | Familiar Organized | Random | Means |
|---|---|---|---|
| Set Size | | | |
| 2 | 27.67 (3.42) | 30.17 (3.19) | 28.92 |
| 3 | 28.08 (4.52) | 25.00 (3.59) | 26.54 |
| 4 | 26.25 (3.89) | 26.83 (3.19) | 26.54 |
| 5 | 25.83 (5.37) | 27.08 (3.18) | 26.46 |
| 6 | 29.42 (1.56) | 23.17 (5.49) | 26.30 |
| Means | 27.45 | 26.45 | |

Systematicity of search. Because order did not influence search strategy, a 2 (Familiarity) x 5 (Set Size) ANOVA was conducted. Although the multivariate procedure did not reveal any significant effects, the results of the univariate analysis that accompanies the multivariate output showed a reliable interaction between familiarity and set size, $F$ (4, 44) = 2.699, $p$ <.05. Figure 5 illustrates the interaction. *Post-hoc* contrasts (Bonferonni alpha = .025) were conducted comparing random and familiar strings on set sizes two, $F$ (1, 11) = 1.355, $p$ = .27 and six, $F$ (1, 11) = 6.756, $p$ = .025. The latter contrast, just reached significance and thus provides marginal support for the hypothesis that maintenance of strategic search is impaired when processing demands are high. The main effects of familiarity, $F$ (1, 11) = 2.223, $p$ =.16 and set size, Wilks' Lambda = .752, multivariate $F$ (4, 8) = .658, $p$ = .64, were not reliable. Table 12 presents cell means.

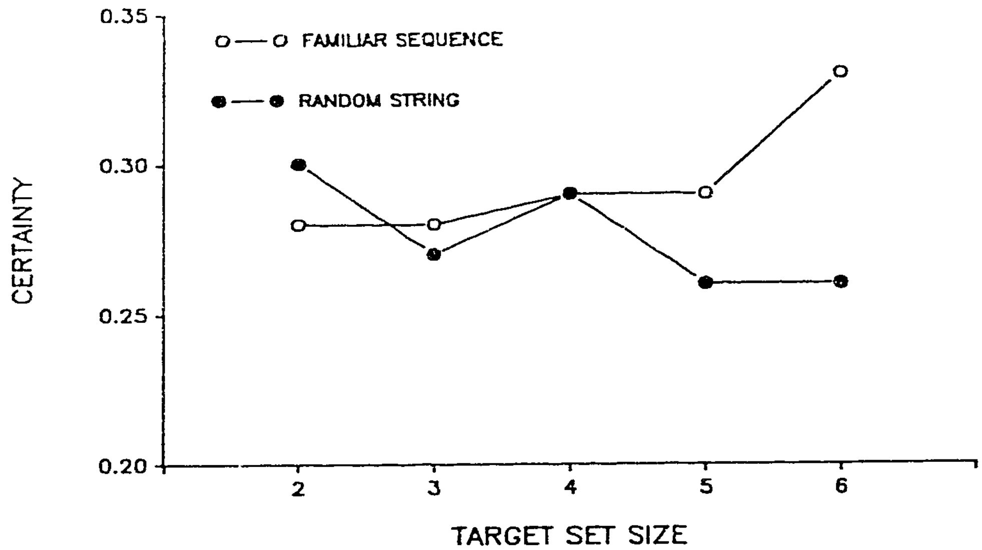



Figure 5. Experiment 3A: Mean certainty scores for familiar sequences and random strings of letters over five target set sizes.

Table 12

Experiment 3A: Mean Certainty Scores with (Standard Deviations)

| | Familiar Organized | Random | Means |
|---|---|---|---|
| Set Size | | | |
| 2 | 0.276 (.076) | 0.297 (.063) | 0.287 |
| 3 | 0.279 (.064) | 0.266 (.070) | 0.273 |
| 4 | 0.286 (.070) | 0.289 (.073) | 0.288 |
| 5 | 0.291 (.074) | 0.272 (.067) | 0.282 |
| 6 | 0.328 (.051) | 0.260 (.082) | 0.294 |
| Means | 0.292 | 0.277 | |

Discussion

The results of this experiment lend some support to the hypothesis that the familiarity of a string of letters can substantially reduce the processing demands inherent in large target sets. There was a consistent increase in processing difficulty (increased time) as target set size increased and the advantage of familiarity was valuable only when set size reached six. The unusual reversal in time for set size two may have something to do with the particular letter sets used. The letters V and D comprised the familiar set, while the random set included X and H. When the set is so small, it is unlikely that familiarity would make a major difference to rehearsal or comparison time and the ease of finding Xs and Hs over Vs and Ds may well have been a function of the perceptual relationships among targets and distractors in the stimulus array. The same argument could be advanced to explain the reliable effect of familiarity on set size six. A major flaw in the design of Experiments 3A and 3B occurred because perceptual features were not equated between the random and familiar sets. Given that the letters in the largest set for both conditions contained angular, rounded and square features, it is probably unlikely that perceptual feature dissimilarity could explain completely the advantage of familiarity found for set size six. Furthermore, the reversal in direction of the effect that is associated with set size six compared to size two argues for something more than perceptual features as the cause. The logical explanation is the increase in processing demands associated with memory rehearsal and comparisons.

The pattern of results associated with accuracy is similar to that associated with detection time. At set size two, random strings were easier than familiar, a result that is consistent with the faster time to find the items. Familiarity exerted an effect for set sizes three and six and, as with the time results, the advantage of familiarity is likely to explain the differences. The advantage of familiarity was not very strong for the systematicity measure. Generally, search strategy did not change substantially over increasing set size regardless of string familiarity until set size six and the effect was only marginally significant. Overall, the results of this experiment indicate that familiarity of the string of letters in the target set can reduce the processing demands associated with large set sizes although that conclusion must be tempered due to design weaknesses.

Results: Experiment 3B

Twenty-nine detection times were replaced by file median times because of difficulties with the pen, representing .7% of the approximately 3,878 hits over the whole experiment. Thirty nontarget detections (.7%) were deleted because they were immediately adjacent to a correct detection and were thus assumed to be unintentional hits. Fourteen (.4%) detections of non-targets adjacent to an undetected actual target were assumed to have been intended as correct hits and the data were adjusted accordingly. Twelve commission errors occurred in this data set and were not analyzed further.

Time. A 2 (Block Order) x 2 (Unorganized Familiar/Random Letter String) x 5 (Target Set Size) ANOVA showed results similar to those found in Experiment 3A. The interaction between string familiarity and set size, illustrated in Figure 6, was reliable, Wilks' Lambda = .178, multivariate $F$ (4, 7) = 8.1, $p$ = .009. *Post hoc* contrasts (Bonferonni alpha = .025) indicated that the interaction was due to familiarity differences for the two letter set, $F$ (2, 10) = 9.62, $p$ = .005 and the six letter set, $F$ (2, 10) = 9.82, $p$ = .004 with the same reversal in direction of the effect between the smallest and largest set as was seen in the previous experiment. The influence of familiarity on set size was responsible for the following main effects. Subjects found the familiar unorganized letters more quickly ($M$ = 2.35) than a random string of letters ($M$ = 2.75), $F$ (1, 10) = 14.60, $p$ = .003. The length of the letter string significantly influenced inter-target detection times, Wilks' Lambda = .059, multivariate $F$ (4, 7) = 28.14, $p$ = .000.

A reliable main effect of order was found, $F$ (1, 10) = 5.467, $p$ = .041. When subjects received the block of unorganized familiar strings before the block of random strings, they took more time per target detection ($M$ = 2.93 s) than if the trials were in the reverse order ($M$ = 2.18 s). Because there were no interactions with order, it is difficult to determine why the order effect occurred. One way to consider the effect is in terms of number of subjects who spontaneously reorganized the letter strings as evidenced by their recall of target letters following each trial. Of the twelve subjects in this experiment, all but one showed evidence of reorganization on one or more trials of unorganized familiar sequences. Because the six item string is the one that showed the greatest difference between the familiarity conditions, spontaneous reorganization was scored on the basis of recall on the six letter familiar sequence trial. Subjects received a score of 1 if they reordered the letters DFAECB to ABCDEF and a score of 0 if they did not do so. Of those receiving a block of unorganized familiar strings followed by a block of random strings, only two of the six subjects reorganized the string to ABCDEF compared to four of the six receiving the other order. The order effect found in this experiment might have been explainable in terms of the number of subjects receiving each order who spontaneously reorganized the six letters. However, because the number of observations in each cell is so small, a statistical analysis of these differences would not be appropriate. Furthermore, it must be noted that, although half the subjects (6) in the experiment did not reorganize the string into its well-known order, five of those six organized the letters into some kind of meaningful set, usually two pronounceable trigrams.

The results of the time analysis suggest that subjects attempted spontaneously to reorganize long strings of letters into meaningful strings and that this facilitated target detection time. It is unclear, however, why detection times should be shorter for random strings when they were presented before the unorganized familiar strings rather than after. The latter trial order might have been expected to influence time by biasing subjects to look for ways in which to reorganize the letter strings. Table 13 provides cell means for all trials within the two presentation orders.

Table 13
<u>Experiment 3B: Mean Times in Seconds With (Standard Deviations)</u>

| | Order One | | Order Two | | |
|---|---|---|---|---|---|
| | Familiar Unorgan. | Random | Familiar Unorgan. | Random | Means |
| Set Size | | | | | |
| 2 | 1.603<br>(.520) | 1.362<br>(.343) | 2.213<br>(.548) | 1.735<br>(.525) | 1.728 |
| 3 | 1.693<br>(.441) | 2.343<br>(.598) | 2.752<br>(.864) | 2.510<br>(.457) | 2.325 |
| 4 | 2.057<br>(.487) | 2.220<br>(.707) | 2.915<br>(.602) | 2.772<br>(1.15) | 2.910 |
| 5 | 2.242<br>(.412) | 2.370<br>(.682) | 2.760<br>(.726) | 3.108<br>(1.17) | 2.620 |
| 6 | 2.058<br>(.366) | 3.797<br>(.902) | 3.190<br>(.753) | 5.258<br>(1.853) | 3.576 |
| Means | 1.931 | 2.418 | 2.766 | 3.077 | |

Note: Order One = block of random letters followed by unorganized letter strings; Order Two = block of unorganized letter strings followed by block of random letters.

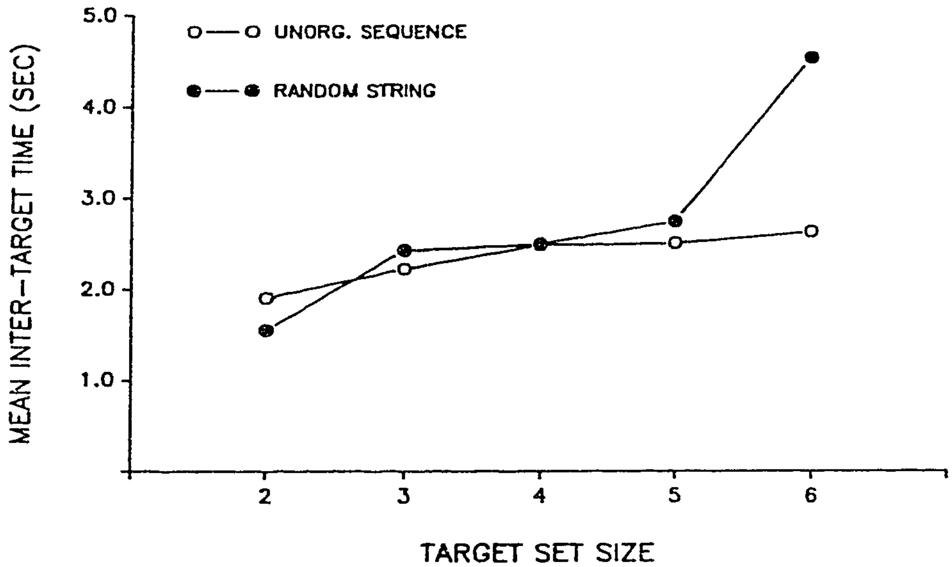



Figure 6. Experiment 3B. Mean inter-target times for unorganized sequences and random strings of letters over five target set sizes.

Number of targets detected. Because order had no effect, the data were collapsed across this factor. The interaction between familiarity and set size was reliable, Wilks' Lambda = .083, multivariate $F$ (4, 8) = 22.03, $p$ = .000 and is illustrated in Figure 7. *Post hoc* contrasts were conducted on set sizes 2, 4, 5 and 6 (Bonferonni alpha = .012). The results indicated that the interaction was due to the poorer target detection associated with the random string of six letters compared to the unorganized six letter sequence, $F$ (1, 11) = 36.765, $p$ = .000. String familiarity was reliable, $F$ (1, 11) = 9.596, $p$ = .010; search for random strings resulted in fewer detections than search for unorganized sequences. Target set size exerted a reliable influence, Wilks' Lambda = .087, multivariate $\underline{F}$ (4, 8) = 28.11, $\underline{p}$ = .000. With increasing set size, the number of detections decreased. Means and standard deviations are provided in Table 14.

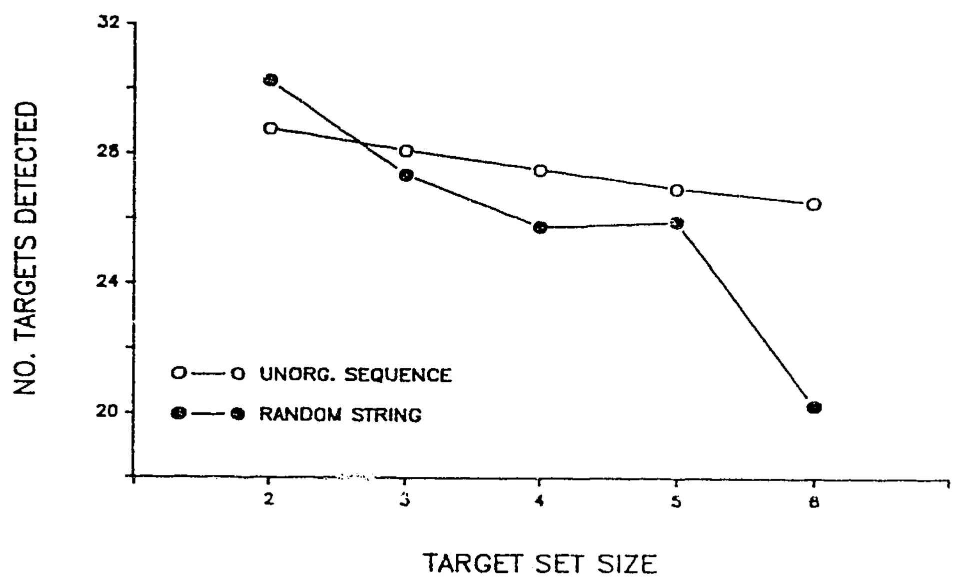



Figure 7. Experiment 3B: Mean number of targets detected for unorganized sequences and random strings of letters over five target set sizes.

Table 14
Experiment 3B: Mean Number of Target Detections With (Standard Deviations)

| | Familiar Unorganized | Random | Means |
|---|---|---|---|
| Set Size | | | |
| 2 | 28.75 (3.33) | 30.25 (2.01) | 29.50 |
| 3 | 28.08 (2.88) | 27.33 (3.09) | 27.71 |
| 4 | 27.50 (2.65) | 25.75 (4.86) | 26.63 |
| 5 | 26.92 (3.58) | 25.92 (3.12) | 26.42 |
| 6 | 26.50 (4.10) | 20.25 (3.75) | 23.38 |
| Means | 27.55 | 25.90 | |

Sytematicity of search. No reliable effects were found in the analysis of this measure. Familiarity, $F$ (1, 11) = 2.136, $p$ = .17; set size, Wilks' Lambda = .705, multivariate $F$ (4, 8) = .837, $p$ = .54 and familiarity x set size, Wilks' Lambda = .850, multivariate $F$ (4, 8) = .354, $p$ = .84. Table 15 presents cell means.

Table 15
Experiment 3B: Mean Certainty Scores with (Standard Deviations)

| | Familiar Unorganized | Random | Means |
|---|---|---|---|
| Set Size | | | |
| 2 | 0.248 (.046) | 0.269 (.053) | 0.259 |
| 3 | 0.265 (.046) | 0.273 (.068) | 0.269 |
| 4 | 0.233 (.054) | 0.256 (.041) | 0.245 |
| 5 | 0.247 (.061) | 0.258 (.058) | 0.253 |
| 6 | 0.260 (.075) | 0.253 (.051) | 0.257 |
| Means | 0.251 | 0.262 | |

Discussion

The results of Experiment 3B indicate that subjects spontaneously attempted to reorganize strings of letters into meaningful sequences in order to reduce the memory demands of the task. This finding has important consequences for research designed to investigate the effects of memory load on visual search. Unless the spontaneous reorganization of letter strings is a focus of the research, incorrect interpretations of data could be made if stimulus strings that can be reorganized are used.

As with Experiment 3A, the finding that random letters took less time than familiar unorganized letters when set size was two, may have been due to perceptual factors. In this experiment, however, the effect was only significant for the time measure. For the same reasons as expressed in the discussion following the previous experiment, the effect of familiarity with a large set of six letters is more likely due to the considerable reduction in the processing demands associated with a familiar string of letters than to perceptual factors. An inspection of Figures 3 and 6 shows that Experiment 3B nicely replicated the results of 3A, although the benefit associated with familiarity of the largest target set was not as great in this experiment.

Contrary to the results of Experiment 3A, the advantage of familiarity did not influence search systematicity. Part of the explanation for this may be related to the way in which subjects reorganized the six letter unorganized sequence; either as a familiar alphabetic sequence or as two pronounceable trigrams. Nearly half the subjects chose the latter method of remembering the letter string. This may have made memorization easier, but it is not clear if the rehearsal of the trigram was in terms of two sets of three separate letters (i.e D-E-F C-A-B) or as three words. If the latter rehearsal method was used, it is conceivable that some kind of decomposition into letters would have to be achieved at the comparison stage of the search. This extra computation would require greater processing control than if the rehearsal was letter by letter and thus could interfere somewhat with strategic monitoring. Consequently, the dramatic increase in systematicity associated with the familiar string of six letters in Experiment 3A, may have been masked by the different reorganizing strategies employed by subjects in this experiment. Taken together, however, Experiments 3A and 3B provide limited evidence for interference in the maintenance of systematicity as the processing demands increased with the size of the target set.

## General Discussion

At the outset of this series of experiments it was assumed that subjects would generally adopt a systematic search strategy when presented with an unstructured array of a large number of letters and instructed to find all the instances of a target. This assumption was based on detection patterns noted in pilot and other research. Maintaining such a strategy while at the same time searching for targets was assumed to make demands on the same limited capacity system as controlled the various processes involved in shifting attention over the array. The experiments reported here did not explicitly test that assumption. Rather, they were designed to determine if the same perceptual and memory load factors that influence selective attention during visual search would impact on strategy use. An examination of the average time required to find each target provided evidence for the relative amounts of controlled processing associated with various conditions. The results of this series of studies provided some, albeit weak, evidence for interference with strategic monitoring on those trials in which processing control demands were high (long RTs). Breakdowns in the maintenance of systematic strategies were seen in Experiment 1 when a target highly similar to the distractor was to be found among homogeneous distractors and in the data of Experiment 3A when the target set consisted of six random letters. To the extent that the same factors that influenced reaction time and are presumably related to the amount of processing involved also influenced strategy use, the assumption of a shared controller is tenable.

With any research it is important to consider the consistency of the findings across subjects in the study. Did all subjects show the same declines in systematicity when processing demands were heavy? The answer is definitely not; if they had, the effects would have been much more definitive. There was considerable between subject variability in terms of choice of strategies and their use. Such variability in strategy use is not uncommon in problem solving research and suggests that individual differences are important in this regard. There is other research to indicate that such individual difference factors as age (Sanford, 1973) or spatial ability (Vincente, Hayes & Willeges, 1987) may be important to consider in future research with this paradigm.

There also was considerable within-subject variability. Subjects did not maintain the same kind of strategy across all trials in an experiment. They might have searched horizontally on two or three trials and then switched to a vertical strategy for a few trials. The new strategy may have been as systematic as the first one but because it meant that a different pattern characterized the search, it was difficult to determine if the new strategy was being followed as consistently as the previous one. The certainty measure that was ultimately used allowed one score to represent relative systematicity regardless of the particular strategy employed. (See Appendix 4 for a discussion of the various attempts to measure systematicity.) Although the certainty measure was determined to be the best of those measures of systematicity that were considered, it is still less than perfect. An arbitrary decision was made to divide the array into four portions for each of the three types of search strategies. The possibility remains that five or six portions may have been more appropriate. Considerable effort is still required to find the best way of measuring search strategy in order to determine how easily people are able to adhere to a chosen plan.

Is it necessarily more efficient to have an overall strategy? (For the sake of argument, let efficiency be defined as the most number of targets found in the least amount of time). The answer may be sometimes yes and sometimes no. The eye-movement data cited in the introduction indicated that systematic strategies were most efficient in those circumstances in which the task involved a search through densely packed homogeneous non-targets for a highly similar target (Scinto et al., 1986). When discrimination was easy, systematic strategies were not more efficient than random ones (Kraiss & Knäeper, 1982). In Experiments 1 and 2, highly dissimilar targets virtually 'popped out' of the display. It is possible that a systematic search strategy was not really necessary for efficient target detection under those circumstances. It is not known if the detection patterns that resulted were an accurate reflection of the visual searching of the subjects (eye-movement measurements would have been useful in this regard). Perhaps the detection patterns reflected a perceptual-motor plan related to movement of the pen.

In the present research, systematicity was seen to break down when targets were not easily discernible from non-targets (Experiment 1: T-N High/N-N Identity condition). Longer detection times and poorer accuracy were associated with the same experimental condition that produced a less systematic pattern. These data support the eye-movement research in providing evidence that random search strategies are not very efficient. On the other hand, as processing demands increased, particularly with large memory loads (e.g., random conditions, Experiments 3A and 3B), efficiency declined (longer times, fewer target detections) but there was only marginal evidence that random search patterns characterized those trials. The perceptual relationships between targets and distractors were constant over the five trials; items were densely distributed and there was considerable featural overlap between targets and distractors. The eye-movement literature suggests that under those circumstances systematic search strategies are the ones most likely to be effective. Indeed, subjects appeared to maintain consistency in strategy use from trial to trial, but the increased processing associated with larger target sets impaired efficiency (time and accuracy). Considered in that way, it is reasonable to assume that maintenance of search strategy utilizes the same limited capacity attentional controller as that required to monitor the memory rehearsal and comparison processes of the task.

The findings from the present series may be better understood in terms of an overall profile of search efficiency that includes time, accuracy and strategy. The time and accuracy results provide compelling evidence that when attentional processing demands are high, performance suffers. The results from this paradigm are particularly interesting regarding the accuracy measure. Previous studies that investigated search for one target seldom report the large effects of accuracy seen in this research. The original hypothesis that processing load impacts on strategy maintenance may need to be revised. Instead, <u>because</u> an overall strategy must be maintained, the manipulations to load on the attentional controller may result in deficits in time and accuracy. Further research is required to investigate how the same processing demands associated with target set size or perceptual discriminability affect time and accuracy when maintenance of an overall strategy is less necessary. This could be done by providing the strategy as part of the display (e.g., structured displays or coloured bands imposed over the array) or by training in the use of strategies.

The primary purpose of this research was to consider strategy use in relation to processing demands. A second goal was to test the Duncan and Humphreys (1989) similarity model with this paradigm (Experiments 1 and 2). The model posits a continuum in search efficiency from relatively automatic to primarily controlled processing depending on the amount of similarity between targets and within the array of non-targets. With this paradigm, completely automatic processing is not possible because, at the very least, attention must be shifted from target to target and this requires attentional control. It is possible, however, to make assumptions regarding the relative contribution of both types of processing; i.e. mostly automatic or mostly controlled.

Duncan (1989) has provided four hypotheses that summarize the model. It would be useful at this time to consider each of these hypotheses in light of the data provided by the present research. First, it is assumed that "as long as targets are sufficiently unlike non-targets, search performance is always independent of array size, irrespective of nontarget homogeneity" (p. 458). This implies that as long as there is low T-N similarity, the search for targets should be relatively automatic. This appeared to be the case in the current research. In both Experiments 1 and 2, when targets were perceptually dissimilar from non-targets, search time was very fast regardless of variations to nontarget homogeneity. The second part of the hypothesis did not appear to hold when two targets were to be searched for, however. Experiment 2 showed that heterogeneity among non-targets increased the amount of controlled processing even when targets were highly dissimilar to targets.

The second hypothesis states that "even increasing target-nontarget similarity has relatively little effect when nontarget homogeneity is at a maximum, in particular when non-targets are all identical" (p. 458). This hypothesis predicts that there should be no search time differences between T-N high and T-N low on the trial in which non-targets were homogeneous (N-N identity). That was clearly not the case; there was a considerable difference in search time regardless of whether the target set contained one or two letters. In fact, in Experiment 1 when targets and non-targets were *dissimilar*, there was a slight, but reliable, increase in time when non-targets were heterogeneous compared to when they were homogeneous. To be fair, Duncan added the parenthetical caveat that "(of course, performance must ultimately suffer as targets become indistinguishable from non-targets)". The data from Experiment 1 showed that the letter P was more difficult to find than either B or S when embedded among Rs. One could hardly argue, however, that P is 'indistinguishable' from R. To be sure, P is particularly difficult to find because of the absence of a defining feature (Treisman & Gormican, 1988; Treisman & Souther, 1985). Looking for an S or a B among Rs is a search for particular curves, looking for a P is a search for the absence of an oblique line. P is less distinguishable than either B or S but even the least indistinguishable of the three (S) caused subjects to take considerably more time than if the target was unlike the distractors. Clearly, Duncan's caveat is important. Determination of the point at which similarity between targets and non-targets begins to seriously interfere with search among homogeneous distractors should be a focus of future research.

The third hypothesis predicts that search performance should be worst "when target-nontarget similarity is high but nontarget-nontarget similarity is low" (p. 458). As discussed previously, the data from the present research indicated that this combination of stimulus factors did not seriously interfere with search efficiency. There was evidence that it might become important when two different targets are to be found rather than one. The fact that nontarget heterogeneity also affected target selection in the T-N low condition, makes it difficult to accept that Duncan's third hypothesis holds even for two targets.

The fourth hypothesis posits that "between these extremes variations in target-nontarget and nontarget-nontarget similarity trace out a continuous surface of search difficulty" (p. 458). The use of three N-N similarity conditions in this research was an attempt to provide an assessment of the continuum but it did so only on that one factor. A better assessment might have been possible if T-N similarity had also included a middle condition between the high and low extremes.

Overall, the research provides limited support for the Duncan and Humphreys model when alphabetic stimuli are used and the task requires search through large arrays for several instances of a target. It is unlikely that experimental paradigm differences explain the discrepancies with the model. However, before that possibility can be ruled out, further research should be undertaken with the multi-target procedure using stimuli that can be more easily quantified on a similarity continuum (i.e. colour patches, Duncan, 1989). The research suggests that the utility of the similarity model may not extend beyond what appears to be a fairly proscribed set of circumstances in which the relationships among the stimuli are well delineated. Research is required to explore more fully the extent to which variations in perceptual similarity among targets and non-targets influence visual search.

An alternative explanation to that provided by Duncan and Humphreys for the effect of T-N similarity may be found in Treisman and Gormican's (1988) 'group scanning' hypothesis. This model posits that subjects scan groups of items in parallel and that the size of the group depends upon the discriminability of the target from distractors; this is reminiscent of the visual lobe model of Scinto et al. (1982). When the target is highly discriminable, the size of the group is relatively large; perhaps as many as eight items (Pashler, 1987b). In a recent paper, Treisman (1991) provided an explanation for the distractor heterogeneity effect. When targets and non-targets are highly discriminable, a target is easily detected because the activity from the map that codes the feature is easily separated from the pooled activity of all other feature maps (non-targets). If the non-targets are heterogeneous, the presence of separate pooled activity maps for each distractor will slow down processing. Homogeneous distractors may be detrimental to search when they share similarity to targets, but, because they can be spatially grouped, the size of the beam of attention can be adjusted to the size of the cluster. Clusters will necessarily be smaller if distractors are heterogeneous and do not group together and attention will have to be shifted more frequently. (This explanation is not very different from Duncan and Humphreys' grouping hypothesis).

Previous studies on visual search for one target among a varying number of distractors have been conducted primarily to develop a thorough understanding of the extent to which stimuli are processed prior to becoming the focus of attention. The debate has been between those who maintain that only simple features are preattentively processed and thus attention enters the system early (e.g. Treisman, 1988) and those who contend that considerably more information is pre-processed with attention occurring later (e.g. Duncan, 1984). The paradigm used in the present research does not allow any conclusions to be drawn relative to the early vs. late discussion. It is, however, predicated on assumptions about automatic (pre-attentive) and controlled (attentive) processing that have been developed over the course of the extensive research history tnat sought to understand the locus of the attentional bottleneck. The research reported here provides further confirmation that perceptual relationships among targets and distractors and also memory load manipulations influence the attentional processing demands associated with visual search. More importantly, the paradigm has allowed assumptions to be made about the strategic monitoring of search that has not been possible with more traditional procedures.

The paradigm allows questions to be considered regarding larger issues of attentional control that go beyond investigation of the factors that influence the immediate focus of attention to those that relate to ongoing attentional monitoring. Thus, the theoretical framework provided by the concept of working memory is a useful one for considering the processes that underly completion of the task. More specifically, the concept of an executive controller that is responsible for monitoring of all cognitive processes as well as maintenance of an ongoing search strategy can be explored. The findings from the current research add to the growing body of literature that has demonstrated that there are limitations to the capacity of the executive component of working memory (i.e. Baddeley, 1986; Morris & Jones, 1991). Further research with this paradigm is necessary to determine more specifically how strategy maintenance and other processing demands interact to tax the limited capacity executive controller.

References

Arani, T. T., Karwan, M. H., & Drury, C. G. (1984). A variable-memory model of visual search. Human Factors, 26, 631-639.

Aubrey, J. B., & Jutai, J. (1989, September). Effects of memory load on visual search of verbal and nonverbal stimulus displays. Paper presented at the 4th annual Joseph R. Royce Conference, Department of Psychology, University of Alberta, Edmonton.

Baddeley, A. (1986). Working memory. Oxford: Clarendon.

Baddeley, A. D., & Hitch, G. (1974). Working memory. In G. H. Bower (Ed.), The psychology of learning and motivation. Advances in research and theory, (pp. 47-89). New York: Academic.

Dehaene, S. (1989). Discriminability and dimensionality effects in visual search for featural conjunctions: A functional pop-out. Perception & Psychophysics, 46, 72-80.

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

Duncan, J. (1985). Visual search and visual attention. In M. I. Posner & O. S. M. Marin (Eds.), Attention and performance XI (pp. 85-106). Hillsdale, NJ: Erlbaum.

Duncan, J. (1989). Boundary conditions on parallel processing in human vision. Perception, 18, 457-469.

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

Egeth, H. E., Marcus, N., & Bevan, W. (1972). Target-set and response-set interactions: Implications for the study of human information processing. Science, 176, 1447-1778.

Eriksen, C. W. & Schultz, S. W. (1979). Information processing in visual search: A continuous flow conception and experimental results. Perception & Psychophysics, 25, 249-263.

Frick, F. C., & Miller, G. A. (1951). A statistical description of operant conditioning. American Journal of Psychology, 64, 20-36.

Gauthier, L., Dehaut, F., & Joanette, Y. (1989). The Bells Test: A quantitative and qualitative test for visual neglect. International Journal of Clinical Neuropsychology, 11, 49-54.

Glass, G. V. & Hopkins, K. D. (1984). Statistical methods in education and psychology. Englewood Cliffs, NJ: Prentice-Hall.

Hitch, G. J. (1978). The role of short-term working memory in mental arithmetic. Cognitive Psychology, 10, 302-323.

Humphreys, G. W., Quinlan, P. T., & Riddoch, M. J. (1989). Grouping processes in visual search: Effects with single- and combined-feature targets. Journal of Experimental Psychology: General, 118, 258-279.

Jutai, J. W. (1989). Spatial attention in hypothetically psychosis-prone college students. Psychiatry Research, 27, 207-215.

Kaplan, I. T., Carvellas, T., & Metlay, W. (1966). Visual search and immediate memory. Journal of Experimental Psychology, 71, 488-493.

Kraiss, K. F., & Knäeuper, A. (1982). Using visual lobe area measurements to predict visual search performance. Human Factors, 24, 673-682.

Krendel, E. S., & Wodinski, J. (1960). Visual search in unstructured fields. In A. Morris & E. P. Horne (Eds.), Symposium on visual search techniques, (pp. 151-169). Washington, DC: National Academy of Sciences.

Logie, R. H., & Baddeley, A. D. (1987). Cognitive processes in counting. Journal of Experimental Psychology: Learning, Memory and Cognition, 13, 310-326.

Loh, E. A., & Beck, C. H. M. (1989). Rats treated chronically with the benzodiazepine, diazepam or with ethanol exhibit reduced variability of behavior. Alcohol, 6, 311-316.

Madden, D. J. (1982). Age differences and similarities in the improvement of controlled search. Experimental Aging Research, 8, 91-98.

Madden, D. J., & Allen, P. A. (1989). Amount and duration of attentional demands during visual search. Perception & Psychophysics, 45, 577-585.

McCormack, P. D., & Wright, N. M. (1964). The positive skew observed in reaction time distributions. Canadian Journal of Psychology, 18, 43-51.

Megaw, E. D., & Richardson, J. (1979). Target uncertainty and visual scanning strategies. Human Factors, 21, 303-315.

Miller, G. A., & Frick, F. C. (1949). Statistical behavioristics and sequences of responses. Psychological Review, 56, 311-324.

Moraglia, G., Maloney, K. P., Fekete, E. M., & Al-Basi, K. (1989). Visual search along the colour dimension. Canadian Journal of Psychology, 43, 1-12.

Morawski, T. B., Drury, C. G., & Karwan, M. H. (1980). Predicting search performance for multiple targets. Human Factors, 22, 707-718.

Morris, N., & Jones, D. M. (1991). Memory updating in working memory: The role of the central executive. British Journal of Psychology, 81, 111-121.

Neisser, U., Novik, R., & Lazar, R. (1963). Searching for ten targets simultaneously. Perceptual and Motor Skills, 17, 955-961.

Norman, D. A., & Shallice, T. (1986). Attention to action. In R. J. Davidson, G. E. Schwartz, & D. Shapiro (Eds.), Consciousness and self-regulation: Advances in research and theory (Vol. 4), (pp. 1-18). New York: Plenum.

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

Pashler, H. (1987b). Target-distractor discriminability in visual search. Perception & Psychophysics, 41, 285-292.

Rabbitt, P. (1978). Sorting, categorization, and visual search. In E. C. Carterette & M. P. Friedman (Eds.), Handbook of perception, Vol. IX. Perceptual processing, (pp. 85-134). New York: Academic.

Rosenthal, R., & Rosnow, R. L. (1991). Essentials of behavioral research (2nd ed.). New York: McGraw-Hill.

Sanford, A. J. (1973). Age-related differences in strategies for locating hidden targets. Gerontologia, 19, 16-21.

Schneider, W., Dumais, S. T. & Shiffrin, R. W. (1984). Automatic and control processing and attention. In R. Parasuraman D. R. Davies (Eds.), Varieties of attention, (pp. 1-27). Orlando, FA: Academic.

Scinto, L. F. M., Pillalamarri, R., & Karsh, R. (1986). Cognitive strategies for visual search. Acta Psychologica, 62, 263-292.

Shallice, T. (1988). From neuropsychology to mental structure. Cambridge: Cambridge University Press.

Shiffrin, R. M., & Czerwinski, M. P. (1988). A model of automatic attention attraction when mapping is partially consistent. Journal of Experimental Psychology: Learning, Memory and Cognition, 14, 562-569.

Shiffrin, R. M. & Schneider, W. (1977). Toward a unitary model for selective attention, memory scanning, and visual search. In S. Dornic (Ed.), Attention and performance, VI, (pp. 413-440). Hillsdale, NJ: Erlbaum.

Sternberg, S. (1966). High-speed scanning in human memory. Science, 153, 625-654.

Thomas, J. C., Waugh, N. C., & Fozard, J. L. (1978). Age and familiarity in memory scanning. Journal of Gerontology, 33, 528-533.

Treisman, A. (1988). Features and objects: The fourteenth Bartlett Memorial Lecture. The Quarterly Journal of Experimental Psychology, 40, 201-237.

Treisman, A. (1991). Search, similarity, and integration of features between and within dimensions. Journal of Experimental Psychology: Human Perception and Performance, 17, 652-676.

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

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

Treisman, A. & Souther, J. (1985). Search asymmetry: A diagnostic for preattentive processing of separable features. Journal of Experimental Psychology: General, 114, 285-310.

Van Der Heijden, A. H. C., Malhas, M. S. M., & Van Den Roovaart, B. P. (1984). An empirical interletter confusion matrix for continuous-line capitals. Perception & Psychophysics, 35, 85-88.

Vicente, K. J., Hayes, B. C., & Williges, R. C. (1987). Assaying and isolating individual differences in searching a hierarchical file system. Human Factors, 29, 349-359.

Ward, R., & McClelland, J. L. (1989). Conjunctive search for one and two identical targets. Journal of Experimental Psychology: Human Perception and Performance, 15, 664-672.

Weintraub, S., & Mesulam, M-M. (1988). Visual hemispatial inattention: Stimulus parameters and exploratory strategies. Journal of Neurology, Neurosurgery, and Psychiatry, 51, 1481-1488.

Williams, L. G. (1966). Target conspicuity and visual search. Human Factors, 8, 80-92.

Appendix 1

Confusion Matrix for Experiment 1

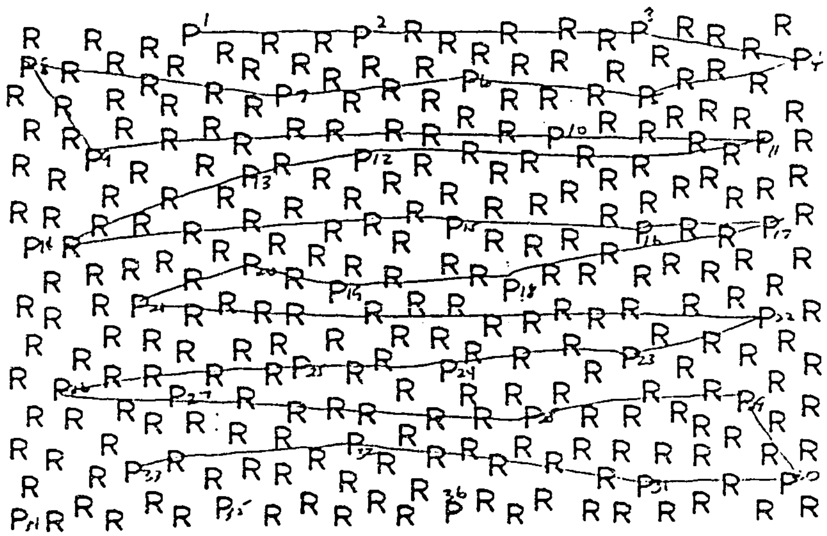
| Stim. | Response H | K | L | O | Q | X |
|---|---|---|---|---|---|---|
| H | | .015 | .013 | .007 | .006 | .002 |
| K | .010 | | .011 | .000 | .001 | .173 |
| L | .022 | .014 | | .005 | .005 | .004 |
| O | .001 | .001 | .003 | | .142 | .001 |
| Q | .003 | .002 | .007 | .252 | | .001 |
| X | .002 | .086 | .007 | .001 | .000 | |

Confusion Matrix for the Additional Letters in Experiment 2

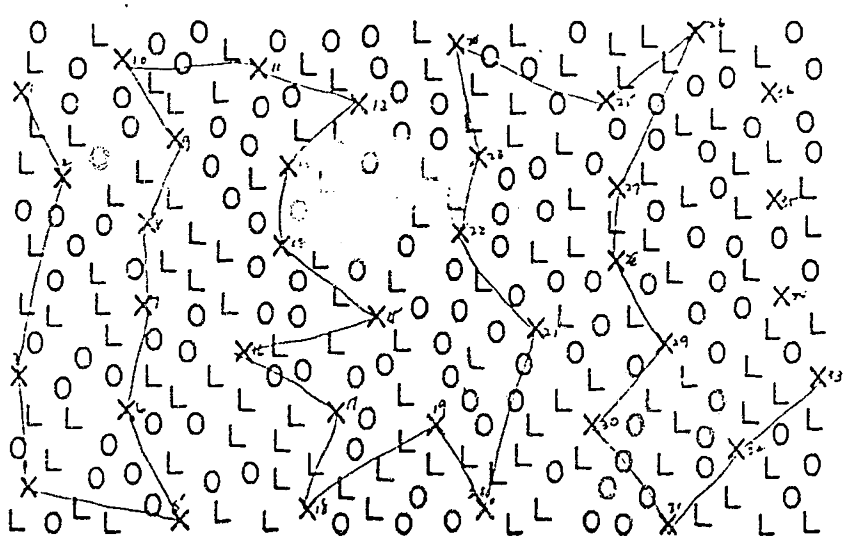
| Stim. | Response B | D | E | P | R | S |
|---|---|---|---|---|---|---|
| B | | .053 | .013 | .012 | .092 | .087 |
| D | .035 | | .004 | .060 | .018 | .014 |
| E | .042 | .006 | | .055 | .067 | .023 |
| P | .023 | .032 | .018 | | .088 | .021 |
| R | .022 | .019 | .015 | .131 | | .044 |
| S | .068 | .010 | .074 | .017 | .101 | |

Appendix 2

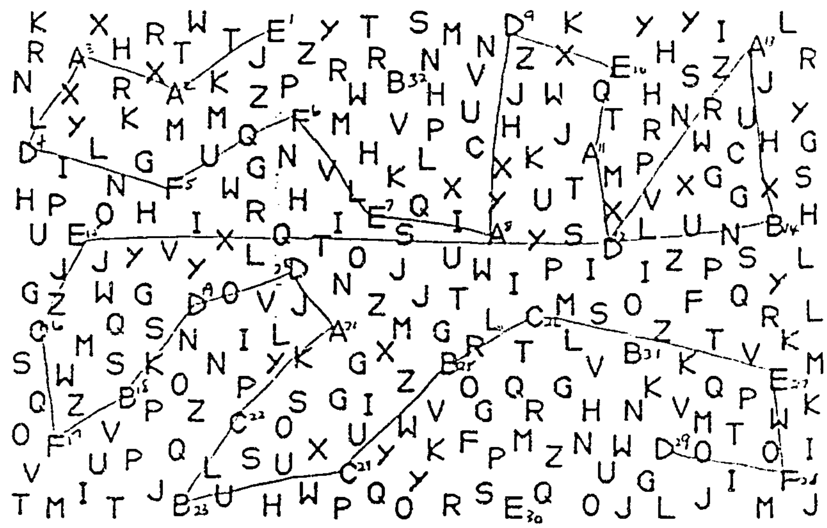
Examples of Search Patterns and Systematicity Scores



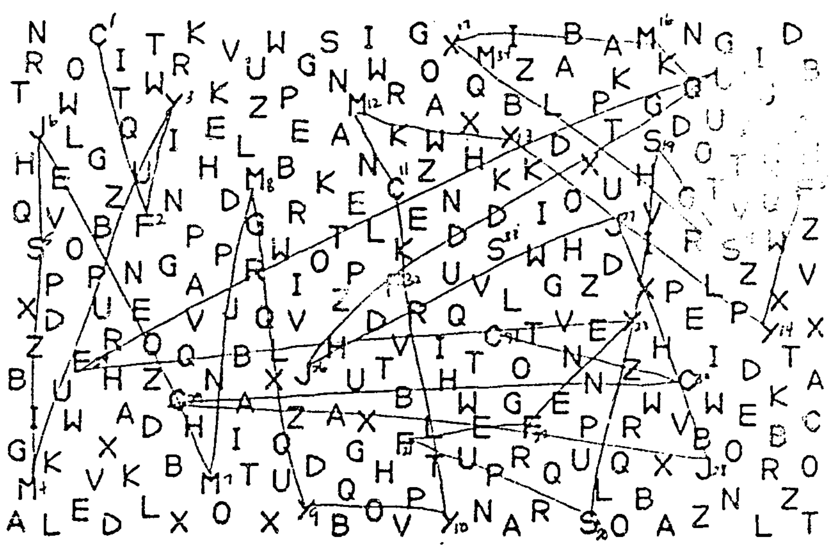
Quadrant = .19, Vertical = .18, Horizontal = .30



Quadrant = .31, Vertical = .34, Horizontal = .22



Quadrant = .36, Vertical = .27, Horizontal = .23



Quadrant = .08, Vertical = .14, Horizontal = .10

# Appendix 3

## Experiment 1: Analyses including letter as a factor

In order to determine if specific letters had a differential effect on search, relevant contrasts were conducted. Differences were found only among the letters in the T-N high condition for the dependent measures of time and accuracy and so contrasts were conducted only in that condition (Bonferroni alpha = .005) for each measure.

Time. Significantly more time was taken to find the letter 'P' ($M$ = 1.29 s) compared to the letters 'B' ($M$ = 1.12 s), $F$ (1, 17) = 28.45, $p$ = .000, or 'S' ($M$ = 1.08 s), $F$ (1, 17) = 38.775, $p$ = .000, in the N-N identity condition when non-targets consisted of a homogeneous array of 'Rs'.

Accuracy. Fewer targets were found when subjects searched among a homogeneous display of Rs (N-N identity) for a 'P' ($M$ = 28.44) than when the target was a 'B' ($M$ = 31.00), $F$ (1, 17) = 18.095, $p$ = .001, or 'S' ($M$ = 31.39), $F$ (1, 17) = 24.351, $p$ = .000.

strategy no matter what the particular strategic style chosen by each subject.

In the pilot research, subjects completed a maximum of six trials and seldom changed strategies from one trial to the next. That is, a person who searched horizontally tended to search that way for all trials. In the current research, subjects completed 10, 12 or 18 trials. Presumably, subjects often got tired of using a particular strategy because frequently individual subjects would switch strategic methods, from searching horizontally to searching vertically and possibly back again. This meant that deviations in Y-Change scores from trial to trial could have been due to a complete switch in search strategy rather than to a breakdown in the use of a particular strategy. The person may have been as systematic in maintaining the new strategy as he or she was when using the old strategy. The data, however, would indicate a large deviation that could be misinterpreted to reflect poor strategy maintenance. For this reason, Y-Change was not used as a measure of search systematicity. Mean Y-Change scores were calculated for Experiments 1 and 2 and are illustrated in Figures A4-1 and A4-2.

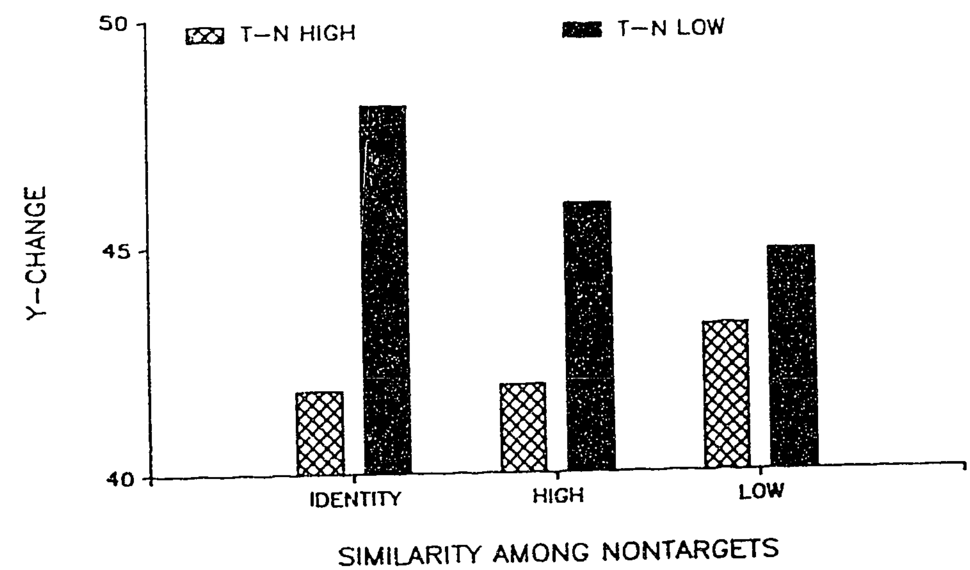



Figure A4-1. Experiment 1: Mean Y-Change Scores for T-N Similarity High and Low for three levels of similarity within non-targets.

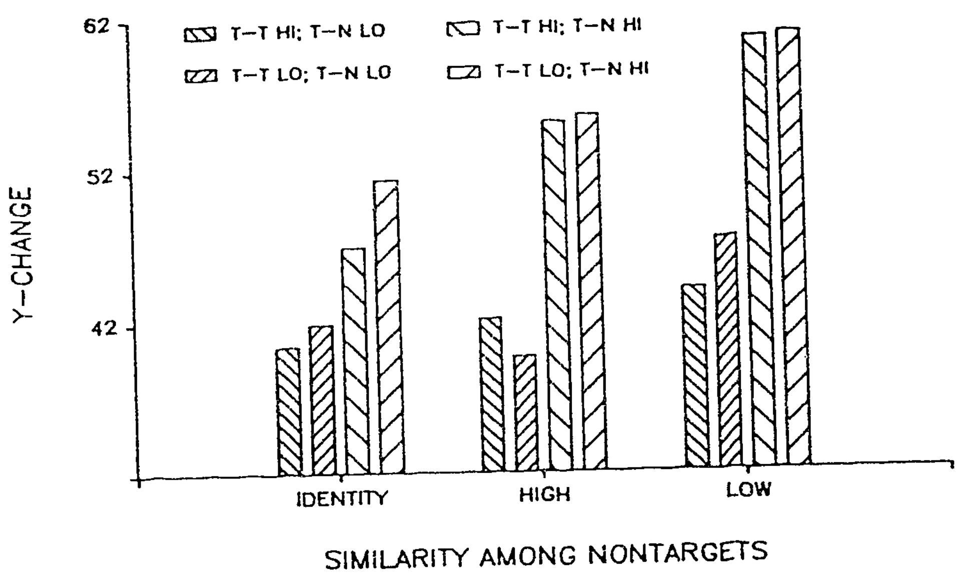



Figure A4-2. Experiment 2: Mean Y-Change scores for T-N Similarity High and Low for three levels of similarity within non-targets.

# Appendix 4

## Search systematicity: Alternate methods attempted

At least two ways of assessing search systematicity were attempted before the decision was made to use the Certainty measure reported in the text. An explanation of these measures and the reasons for abandoning them follow.

### Y-Change Score

Each letter on the screen is associated with a particular position on a grid and is designated with the X and Y coordinates specific to the lower left corner of the letter. If a person uses a horizontal search strategy that sweeps back and forth across the array in a relatively narrow band, the average change in the Y coordinate associated with successive hits is fairly small. (A Y-Change measure is computed by calculating the absolute difference in Y coordinates between each successive pair of hits and determining the average over total hits). This Y-Change measure is larger, of course, if the band width is wider. Deviations in the average Y-change measure from trial to trial are indicative of deviations in search strategy.

Pilot research had indicated that even for subjects who did not use horizontal search strategies, the Y-Change measure was sensitive to changes in maintenance of search strategy. A person using a vertical strategy would have a much larger Y-Change score than a horizontal searcher, but deviations from the usual vertical strategy resulted in an even larger Y-Change score. Thus, the Y-Change measure could demonstrate trial by trial changes in search

Absolute Difference X-Y

In order to take into account the possibility that subjects might change strategy from trial to trial without necessarily becoming less systematic, an X-Y absolute difference measure was computed. The logic behind this measure was that if a subject used a horizontal strategy, the Y-change measure would be fairly small compared to the X-change measure and conversely, for a vertical searcher, the X-change measure would be small compared to the Y-change measure. Therefore, if the absolute difference between the two measures was large, it could be assumed that the person was using either a horizontal or a vertical strategy. A small difference would reflect a more random style that was neither horizontal or vertical. The screen display was not square so that the number of units on the X axis was greater than on the Y axis. Y-Change scores were, therefore, multiplied by 1.75 in order to make the two measures comparable. The absolute difference between the two measures was calculated for each subject and the average X-Y Absolute Difference score was computed for each trial in Experiment 1.

Inspection of Figure A4-3 indicates a main effect of T-N similarity using the X-Y Absolute Difference score such that larger differences were found when T-N Similarity was high. This suggests that subjects used highly systematic strategies when there was high similarity between targets and non-targets compared to when T-N Similarity was low. This finding was contrary to the hypothesis of the research and at the time was difficult to interpret. For this reason, as well as the fact that the measure did not adequately deal with other kinds of strategies (e.g., quadrant searches), the measure was abandoned in favour of the Certainty statistic reported in the text.

Now that the research is complete, the effect that was so clear in Figure A4-3 is easier to understand. As discussed previously, the use of systematic strategies may well be more likely when perceptual aspects of the display make discrimination difficult than when it is easy. The X-Y Absolute Difference measure should be explored more fully in future research.

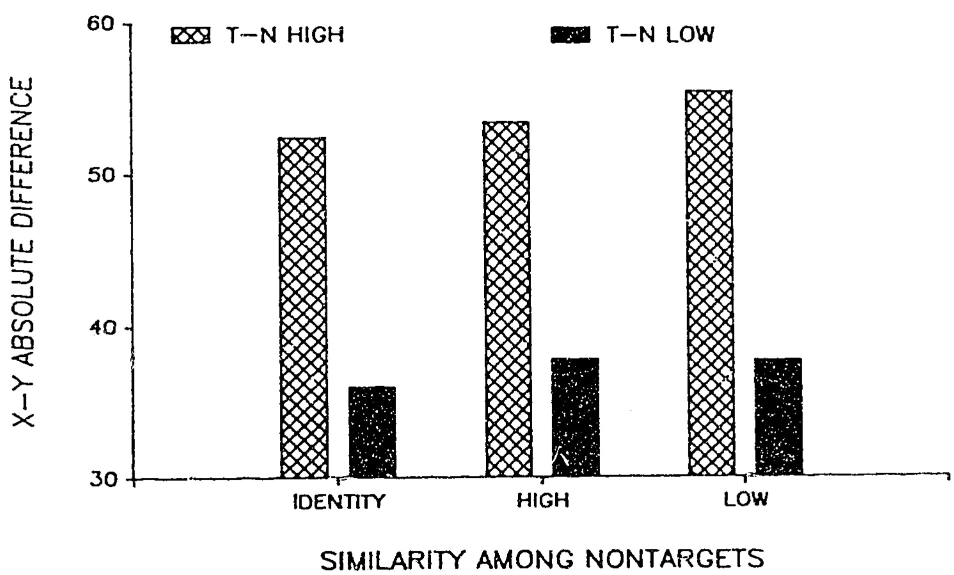



Figure A4-3. Experiment 1: X-Y Absolute Difference scores for T-N Similarity High and Low for three levels of similarity within non-targets.