

See discussions, stats, and author profiles for this publication at:  
<https://www.researchgate.net/publication/232505778>

# Controlled and automatic human information processing: I. Detection, search, and attention

Article *in* Psychological Review · January 1977

DOI: 10.1037/0033-295X.84.1.1

---

CITATIONS

3,635

READS

2,671

2 authors, including:



[Richard M Shiffrin](#)

Indiana University Bloomington

142 PUBLICATIONS 15,013

CITATIONS

[SEE PROFILE](#)

# Psychological Review

Copyright © 1977 by the American Psychological Association, Inc.

VOLUME 84 NUMBER 1 JANUARY 1977

## Controlled and Automatic Human Information Processing: I. Detection, Search, and Attention

Walter Schneider  
University of California, Berkeley

Richard M. Shiffrin  
Indiana University

A two-process theory of human information processing is proposed and applied to detection, search, and attention phenomena. Automatic processing is activation of a learned sequence of elements in long-term memory that is initiated by appropriate inputs and then proceeds automatically—without subject control, without stressing the capacity limitations of the system, and without necessarily demanding attention. Controlled processing is a temporary activation of a sequence of elements that can be set up quickly and easily but requires attention, is capacity-limited (usually serial in nature), and is controlled by the subject. A series of studies using both reaction time and accuracy measures is presented, which traces these concepts in the form of *automatic detection* and *controlled search* through the areas of detection, search, and attention. Results in these areas are shown to arise from common mechanisms. Automatic detection is shown to develop following consistent mapping of stimuli to responses over trials. Controlled search is utilized in varied-mapping paradigms, and in our studies, it takes the form of serial, terminating search. The approach resolves a number of apparent conflicts in the literature.

### I. General Introduction

Atkinson and Shiffrin (1968) advocated a fundamental division of human memory and information processing into (a) labile control processes and (b) learned or inherent structural components. Our present theory differs from this earlier one in numerous respects and

covers many new areas, but this basic distinction remains. There are other interesting similarities as well. Atkinson and Shiffrin demonstrated controlled processing by choosing a particular control process, rehearsal, and designing numerous tasks that would vary the rehearsal demands and extend the rehearsal capabilities to the limit. In a similar fashion, we shall here explore another mode of controlled processing, search, by designing numerous tasks that vary the search demands and test the limits of the search capabilities.<sup>1</sup>

The research and theory reported here was supported by Public Health Service Grant 12717 and a Guggenheim Fellowship to the second author, by Grant MH23878 to the Rockefeller University, and by a Miller Fellowship from the University of California at Berkeley to the first author. This report represents the equal and shared contributions of both authors.

Requests for reprints should be sent to Richard M. Shiffrin, Department of Psychology, Indiana University, Bloomington, Indiana 47401.

<sup>1</sup> One important element in the present treatment that is missing in that of Atkinson and Shiffrin (1968) is our inclusion of automatic detection tasks so that the controlled short-term and learned long-term components of processing may be contrasted and compared.

However, despite the similarities in basic aims, the experiments lie in entirely different areas, and the present treatment is to be regarded as a new and independent theory to be evaluated in its own right.

### A. A Brief Theoretical Overview

Memory is conceived to be a large and permanent collection of nodes that become complexly and increasingly interassociated and interrelated through learning. Most of these nodes are normally passive and inactive and are termed long-term store when in the inactive state. The set of currently activated nodes is termed short-term store. Long-term store is thus a permanent, passive repository for information. Short-term store is a temporary state; information in short-term store is said to be lost or forgotten when it reverts from an active to an inactive phase. Control of the information-processing system is carried out through a manipulation of the flow of information into and out of short-term store. These control processes include decisions of all sorts, rehearsal, coding, and search of short- and long-term store. Long-term store contains learned sequences of information processing that may be initiated by a control process or by environmental or internal information input but that are then executed automatically with few demands upon the capacity of short-term store.

It is unnecessary for the purposes of this paper to make very many assumptions concerning the organization and structure of long-term store. Long-term store will be treated as a very general graph with complex interrelations among nodes. Each node itself may consist of a complex set of informational elements, including associative connections, programs for responses or actions, and directions for other types of information processing. What then sets off one node from another? The node is a distinguishable entity because it is unitized—when any of its elements are activated (i.e., placed in short-term store), all of them are activated. One node, of course, may activate another node; it does not do so in all situations, however, but only when the context or the state of

the information-processing system is appropriate.

An *automatic process* can be defined within such a system as the activation of a sequence of nodes with the following properties: (a) The sequence of nodes (nearly) always becomes active in response to a particular input configuration, where the inputs may be externally or internally generated and include the general situational context. (b) The sequence is activated automatically without the necessity of active control or attention by the subject.

An automatic sequence differs from a single node because it is not necessarily unitized. The same nodes may appear in different automatic sequences, depending on the context. For example, a red traffic light might initiate a braking response when the perceiver is in a car, and a walking, halting, or traffic-scanning response when the perceiver is a pedestrian.

Since an automatic process operates through a relatively permanent set of associative connections in long-term store, any new automatic process requires an appreciable amount of consistent training to develop fully. Furthermore, once learned, an automatic process is difficult to suppress, to modify, or to ignore.

A type of automatic sequence that plays a very important role in the present paper is one that modifies ongoing controlled processing by attracting attention to a specified locus or node. In particular, when subjects in search tasks are consistently trained to recognize certain inputs as targets, these inputs acquire the ability to initiate *automatic-attention responses*. These attention responses then direct attention (i.e., will direct controlled processing) automatically to the target, regardless of concurrent inputs or memory load, and enable a correct detection to occur.

A *controlled process* is a temporary sequence of nodes activated under control of, and through attention by, the subject. Because active attention by the subject is required, only one such sequence at a time may be controlled without interference, unless two sequences each require such a slow sequence of activations that they can be serially inter-

woven. Controlled processes are therefore tightly capacity-limited, but the cost of capacity limitations is balanced by the benefits deriving from the ease with which such processes may be set up, altered, and applied in novel situations for which automatic sequences have never been learned. Controlled-processing operations utilize short-term store, so that the nature of their limitations is determined at least in part by the capacity limitations of short-term store.

One type of controlled process that plays a very important role in this paper appears in search tasks at low levels of practice or whenever targets are inconsistently mapped to responses across trials. This controlled process is a serial comparison mechanism in which each possible target is compared in turn to each presented item until a match is found.

### B. Goals

The primary empirical goal of the research reported in the present paper is the ordering of the large bodies of literature on detection, search, and attention into a simple predictive system. It will be shown that the results of experiments in these areas are determined by three major variables: (a) the memory load, or complexity, (b) the nature of the stimulus-to-response mapping (particularly the consistency of the mapping), and (c) the amount of practice.

Our theoretical goal in the present paper is the proposal and empirical defense of a unified model for selective attention, short-term search, and detection and recognition of targets amongst multiple sensory inputs. The theory and the interpretation of the data are based upon the two qualitatively different mechanisms that we have called *automatic processing* and *controlled processing*. In detection and search tasks, controlled processing will be termed *controlled search*, and automatic processing will be termed *automatic detection*.

### C. Outline

Because the paper has a very wide scope, includes many experiments, and is quite long,

it will be published in two sections. The present paper comprises Part I, and Part II will appear in the next issue of this journal (Shiffrin & Schneider, in press).

In Part I, search and attention tasks are shown to depend on the same basic mechanisms. We shall show that divided-attention limitations appear and have characteristics due to the use of controlled search and that these limitations are bypassed when automatic detection is utilized. Quantitative models are developed and applied to the results of search tasks using reaction time measures and to the results of attention tasks using accuracy measures. Finally, the research literature is reviewed in light of the theory and data, and a number of anomalous or paradoxical findings are shown to be consistent with and predictable from the two-process theory.

The experimental paradigm that is used to produce the data, to provide the various demonstrations, and to achieve our goals is a basic visual search procedure. The subject is asked to search for the presence of one or more of a memorized set of  $n$  objects among a set of  $m$  visual inputs. When the targets and the distractors are never mixed across trials, then performance improves radically and is relatively unaffected by load. These changes are ascribed to the development of automatic detection. On the other hand, when targets and distractors are mixed from trial to trial, then a slow serial search must be used that is highly dependent on load, requires attention, and is relatively inefficient. These results are ascribed to the use of controlled search. We shall show that such findings hold both in tasks that use a small number of simultaneous inputs and reaction time as the dependent measure, and also in tasks that use many successive frames of inputs and accuracy as the dependent measure.

Part II will utilize the experimental techniques of Part I to examine perceptual learning, learning in general, categorization, and the focusing of attention. In each case, the roles of automatic detection and controlled search will be explored. A general theory of human information processing will be put forth, with particular emphasis on automatic

and controlled processing. The proposed theory will be compared with and contrasted to previous models of search and attention.

## II. Introduction to Attention and Search Tasks<sup>2</sup>

### A. Attention Tasks

In very general terms, selective attention is the control of information processing so that a sensory input is perceived or remembered better in one situation than another according to the desires of the subject. Selectivity of attention is necessary because the processing and memory system has a limited capacity. Thus, in many tasks the subject finds that he has an information overload, and he is forced to select a portion of the input information for processing, rehearsing, and coding. The reduction of performance due to an overload is termed a *selective-attention deficit*.

Selective-attention deficits may be partitioned into two classes. Speaking loosely, a *divided-attention deficit* is said to occur when the necessity to give controlled processing to additional sensory inputs, or additional memory elements, reduces performance. In practice, one must take care to rule out the possibility that the additional inputs reduce performance in ways not under the subject's control. Lateral masking, for example, can automatically reduce the quality of processing when additional inputs are presented in close proximity to those being processed.

A *focused-attention deficit*, on the other hand, is said to occur when the subject fully knows which inputs are relevant but cannot ignore the irrelevant inputs and finds performance harmed by their presence. Again, care must be taken to insure that the extra inputs reduce performance through effects upon controlled processing, rather than through automatic factors like masking.

Examples of these processes are often described in terms of listening behavior at cocktail parties, where one attempts to listen to one of many simultaneous conversations. Listening to just one of the conversations without distraction by the others is an ex-

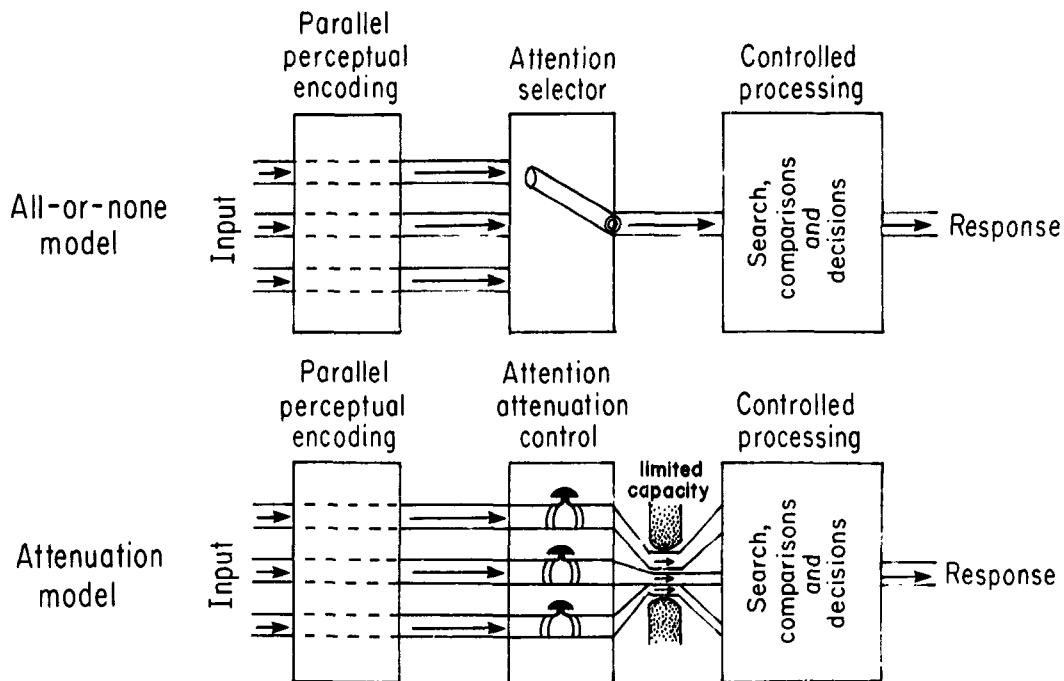
ample of the ability to focus attention. A focused-attention deficit would occur if the mention of one's name in a peripheral conversation caused one to lose track (temporarily) of the conversation on which one was focusing. Detecting one's name when mentioned in any of several simultaneous conversations is an example of the ability to divide attention. A divided-attention deficit would occur when two conversations could not be followed as easily as one.

The concept of selective attention is intimately related to that of limited capacity (see Shiffrin, 1976). If our capacity to process, decide about, and remember information were not limited, then selective attention would serve no purpose. It is because processing capacity is overloaded in numerous situations that a subset of the information arriving must be given special attention. Any selective-attention deficit, therefore, implies a corresponding capacity limitation.

Recent theories of selective attention (and most of the research) have sprung from Broadbent's important work in the 1950s (see Broadbent, 1958). In essence, he proposed an initial perceptual system that was not subject to selective processes but was followed by a limited-capacity system on which selective processes could operate. Part of the limited-capacity system consisted of a filter that selected some of the results of peripheral processing for further processing by the limited system. The essence of Broadbent's theory is that automatic processing occurs in parallel up to the level of the filter; beyond the filter, processing is continued only for those inputs selected by the filter. In Broadbent's original theory, the filter acted like an all-or-none switch that could be focused on just one input (or channel) at a time, and an appreciable time was needed to change the filter to a new channel. (A channel was a source carrying sensory input, such as either of the two ears.) Such

---

<sup>2</sup> A brief glossary of terms is included in Appendix A so that the terminology in this paper may be defined precisely despite variations in terminology throughout the literature.



### SENSORY PROCESSING MODELS

Figure 1. Two models of attentional selection during perceptual encoding. Top panel: Information from only a single source at one time can pass through the selector (e.g., see Broadbent, 1958). Bottom panel: Information from all sources passes the selector, but in attenuated fashion (e.g., Treisman, 1960).

a system is diagrammed in the top panel of Figure 1.

This theory, in its simplest form, predicts that an input on a nonattended channel will not be perceived. However, at least certain kinds of inputs are perceived even on a nonattended channel (see Treisman, 1969). Data of this kind could be explained by assuming that the filter occasionally switches to the nonattended channel for brief periods of time, thereby causing processing of at least some of the information there. However, the data show that the type of information on the nonattended channel determines its perceptibility. An all-or-none filter theory would predict that only time spent processing the nonattended channel should determine the information's perceptibility. A relevant result can be found in an experiment by Treisman and Riley (1969). Differing inputs (words) were made to arrive at the two ears (chan-

nels) simultaneously. Subjects repeated back (shadowed) the inputs in one ear, thereby focusing their attention on that ear. In addition, they were required to detect a specific target whenever it occurred on either ear. When the target was a word spoken in the same voice as the rest of the message, it was detected much less often on the nonattended ear than on the attended ear. However, when the target was a word in a different voice or was just a tone, it was detected about as well on either ear. Such results cannot be easily interpreted in terms of an all-or-none filter.

Faced with results of the kind described above, Treisman (1960, 1964) proposed a variation of Broadbent's model in which the filter did not operate in an all-or-none fashion. Rather, the filter was assumed to have a limited capacity that could be allocated by the subject to the various input

channels. This is called an attenuation model because the processing given to various channels is attenuated to degrees controlled by the subject. Such a model is depicted in the bottom panel of Figure 1.

In one form or another, these models have dominated theorizing about selective attention. However, when one attempts to specify the details of these models, many questions arise. Foremost among these questions is one concerning the locus at which the limited-capacity bottleneck first occurs. Deutsch and Deutsch (1963) proposed an extreme view in which the bottleneck occurs very late, during the decision or response stages but not during the perception or feature-extraction stages. Numerous other workers have proposed models in which the locus of the bottleneck is located at earlier stages of processing, perhaps just after the simplest features have been abstracted (see Part II).

Evidence suggesting that the locus for the bottleneck occurs late in processing was collected in a series of experiments by the second author and several colleagues. Shiffrin and Gardner (1972) carried out the first, prototypical studies, based in part on earlier work by Eriksen and Spencer (1969). In two different conditions, four characters were presented in the corners of an imaginary square. In both conditions, the subject had to decide which of the letters "T" or "F" had been presented as one of the four characters. In the simultaneous condition, the four characters were presented simultaneously for a duration of  $t$  msec, and they were preceded and followed by masking stimuli. In the sequential condition, the four stimuli were presented successively in a known, specified spatial order; each stimulus was presented for  $t$  msec, and each was preceded and followed by masks. The sequential condition allowed attention to be focused on each character in turn, while the simultaneous condition required sharing of processing capacity among the four stimuli. The results, surprising at the time, showed that performance in the simultaneous condition was at least as good as that in the sequential condition. Thus, performance depended on the exposure duration of a single stimulus,

not on the number of stimuli presented simultaneously. These results are difficult to reconcile with attention models proposing that more and better processing is given to the stimuli receiving greater allocation of attention.

The results of these and other experiments led Shiffrin (see Shiffrin, 1975, for a review) to agree with Deutsch and Deutsch (1963) that the locus of selective attention occurred late in processing, at some point in the memory-search, decision, and response stages. The basic idea was that automatic, uncontrolled, feature abstraction, called *systemic processing*, resulted in a series of features or codes' being placed in short-term memory shortly after presentation of an input. These codes were viewed as temporary activations of features normally present but inactive in long-term memory. The great majority of the activated features, those that could not be given special rehearsal treatment, were thought to be lost (i.e., inactivated) very quickly, within a few hundred milliseconds or less. Thus, the limitation on the subject's capacity lay in the limited rate at which the subject could search and decide about the active features before they were lost from short-term memory. The same limitation was seen as causing attention deficits when new inputs continuously arrived at such speeds that search and decisions about one set of inputs could not be completed before new inputs were presented. In this model, selective attention operated through the subject's control over the order in which inputs were searched and decisions were carried out.

Thus, when the present experimental program was undertaken, we evaluated the state of attention research as follows. Numerous studies had suggested that the locus of divided-attention deficits occurred rather late in the information-processing sequence. Many other studies had shown what sorts of paradigms would and would not produce attentional limitations (see Part II). However, there was available no overall theory able to predict where limitations would appear. Furthermore, the theorizing was largely qualitative. Quantitative theories of the nature of the attentional process were few, and these

few were applied only to very specific paradigms. In particular, the time lost while carrying out one process had often been implicated as a cause of performance degradation for a competing process. Yet general quantitative models of the timing mechanisms had not been produced.

As a result of these considerations, we were led to two basic goals guiding the series of studies to be reported in this paper. First, we wished to discover what characteristics of the detection and search processes led to attentional limitations. With this information, it would be possible to predict which situations would produce attentional limitations and which situations would not. Second, for those situations producing limitations, we wished to discover what the quantitative and qualitative rules of search were that determine the nature and amount of the attentional limitation.

### B. Search Tasks

In the preceding section, we suggested that search of short-term memory might be intimately involved with selective attention. In this section, some of the results and theorizing about short-term search are introduced.

A series of objects in short-term memory can be searched for any number of characteristics, but the typical task involves comparing one set of characters with another to decide whether there is an element common to the two sets. Sternberg (1966) presented a now-standard paradigm known as high-speed memory search. In this paradigm, a set of characters (typically 1-6 in number) termed the *memory set* (or the *positive set* in Sternberg's terminology) is presented and presumably entered in short-term memory. Then, a single character called the *input item* (or the *probe*) is presented (usually visually), and the subject gives one response if the input item is in the memory set (a positive response) and another if it is not (a negative response). Usually, the set of possible input items not in the memory set is well-defined; if so, this set is called the *distractor set* (*negative set* in Sternberg's terminology). In Sternberg's experiment, errors occurred infrequently, and

reaction time (also known as response latency) was the dependent measure.

Sternberg found that reaction time was a linear function of memory-set size and that it increased about 40 msec for each additional item in the memory set. Further, the memory-set-size functions for inputs that were in the memory set and inputs that were not were parallel. Sternberg proposed a serial comparison process in which each input item was compared in turn to each member of the memory set. He further assumed that the search was exhaustive: In an exhaustive search, the input item is compared to all members of the memory set before a decision is made on whether a match has been found (even if a match actually occurs early in the comparison process). This model predicts linear, parallel set-size functions. A terminating model, in contrast, assumes that search stops whenever a match occurs; such a model would predict that functions for positive trials would have one half the slope of functions for negative trials. A model in which comparisons are made against the members of the memory set independently and in parallel would not predict linear functions.

Very similar results occur in visual search. In the prototype of this paradigm (e.g., Atkinson, Holmgren, & Juola, 1969), a single memory item was placed in short-term memory, and then the characters comprising the visual display (or frame) were presented simultaneously. The subject had to decide whether or not the memory-set item was present in the visual frame. Again, linear, parallel, set-size functions occurred, with slopes of 40 msec/item. Naturally, an exhaustive, serial comparison process of the kind proposed by Sternberg could predict these results.

The simple picture painted by the preceding results suggests that subjects search short-term memory serially and exhaustively at a rate of about 40 msec per item; but this is far from the whole story. The main alternative findings we wish to discuss at this point are those showing that search in some conditions shows either no dependence, or a much reduced dependence, on memory-set size. Eggerth, Jonides, and Wall (1972), for

example, showed that search speed for a "4" in a background of "C" distractors did not depend on the number of characters in the display. Jonides and Gleitman (1972) showed that this result held when subjects searched for a digit among letters or a letter among digits but not when they searched for a letter among letters or a digit among digits. Jonides and Gleitman introduced controls to show that this pattern could not have arisen due to a gross physical cue discriminating the letter and digit sets. In particular, they showed that a shape difference between the set of letters and the set of numbers was unlikely to have been the cause of the flat set-size function.

In memory search, Swanson and Briggs (1969), Simpson (1972), and Kristofferson (1972b), among others, have shown that extended practice with a memory set that does not change across trials reduces the dependence of reaction time on memory-set size and causes the set-size function to become curvilinear with a slope that tends toward zero as the set size increases. In hybrid, visual, memory-search tasks, Neisser, Novick, and Lazar (1963), Neisser (1974), and Sperling, Budiansky, Spivak, and Johnson (1971) have shown that given an unchanging, multi-item memory set, search for *any* item of that set can proceed only as fast as search for the *particular* item of that set that is searched for the most slowly when that item alone comprises the memory set. In these studies, there was a categorical distinction between memory-set items and distractors (letters vs. numbers). However, Briggs and Johnsen (1973) showed that use of an unchanging memory set over trials considerably flattened the set-size functions even when both the memory set and the distractors consisted of letters. In summary, then, there are many results showing search to be almost independent of the number of items held in memory or presented visually.

It is natural to suppose that the search process underlying the flat set-size functions is one that allows attentional limitations to be bypassed, and to suppose that the search process underlying the linear set-size functions with slopes of about 40 msec per item is

one that results in divided-attention deficits. But what are the two underlying search mechanisms and what gives rise to them?

To answer this question, one must determine the distinguishing properties of the two types of search studies. A number of task variations seem to lead to flat set-size functions in at least some studies. These include tasks in which the memory-set items differ from the distractors by a simple prominent physical characteristic (such as color in visual search or voice in memory search) or by a categorical distinction. They also include tasks in which long practice is given with unchanging memory sets and distractor sets.

We feel that the evidence suggests two very different types of search, the type adopted seeming to depend on the nature of the training in each task. When memory-set items and distractor items are consistently mapped to responses across trials, an efficient search technique can be learned that eventually gives rise to flat set-size functions. When the mapping is varied from trial to trial, such learning cannot occur and the subject is forced to carry out serial search at about 40 msec per comparison. A number of two-process search models with some resemblance to this one have been proposed previously (see Atkinson & Juola, 1973, 1974; Corballis, 1975; Shiffrin & Geisler, 1973), but discussion of other models will be deferred until Part II.

To sum up our introduction to search processes, then, there seem to be at least two types of search patterns. One is serial in nature with comparisons taking about 40 msec each. The other is more nearly parallel in nature and is less dependent on the number of alternative distractors present in short-term memory during the search.

Our researches into search processes are guided by the same goals guiding our explorations of attentive processes. First, we wished to establish that two qualitatively different search processes existed and to discover the conditions under which the two processes could develop. Second, we wished to delineate the quantitative and qualitative nature of each of these search mechanisms.

If these goals were achieved, we hoped to show that selective attention limitations resulted from the limitations of the two search processes.

### III. Empirical and Theoretical Investigations of Basic Processes in Search and Attention

The goals mentioned in the preceding section were instrumental in guiding our choice of an experimental task. To allow a wide latitude for variations in load, a search task was chosen in which both the memory-set size and the display size could be varied. The task was designed so that the categorical relationship between the targets and the distractors could be varied and also so that the mixing over trials of the items from the memory set and the items from the distractor set could be varied. Thus we could anticipate that in some of the conditions, serial search would be utilized, and that in others, parallel, automatic detection might be utilized. Finally, we decided to utilize two basic variants of the search task. In the first, many input displays would be presented sequentially, and the subject's detection accuracy would be the dependent measure. Such a task would relate to studies in the attention literature. In the second, only a single display of simultaneous items would be presented, accuracy of response would be high, and the reaction time to produce a response would be the dependent measure. Such a task would relate to typical studies in the search literature. Eventually, the results of these two basic task variants could be contrasted, compared, and fit with a common quantitative model.

#### A. A Multiple-Frame Attention and Search Task

The first study to be described is a search task using accuracy as a measure. However, we wished to study conditions in which accuracy would be limited by the rate of search rather than by the perceptibility of the inputs. Thus, rather than present stimuli at perceptual threshold, we decided to present stimuli well above threshold, but in such great numbers, and in such rapid succession,

that errors would occur owing to the limited rate of search through the inputs (at least in those conditions requiring limited serial search). The remaining details of the paradigm were chosen to mimic the usual conditions in search studies measuring reaction times—we hoped to balance the current extreme emphasis on studies utilizing reaction time measures only. Finally, and perhaps most important, we wished to design a paradigm that would incorporate in a single study on common subjects the entire range of variables and relations that had previously appeared in separate studies using different subjects.

These constraints led to the adoption of a paradigm that is a variant of one introduced by Sperling et al. (1971) and that is related to a study by Sternberg and Scarborough (Note 1). It also has similarities to the visual search tasks introduced by Neisser (1963, 1974). The subject is required to view a succession of frames of visual characters and to search for the possible occurrence of any one of another set of characters held in short-term memory (see Figure 2). The attention demands of the task are changed by varying (a) the time for which each visual frame is presented, (b) the number of characters presented in each frame, and (c) the number of characters held in short-term memory. When a member of the memory set does appear in the sequence of frames, it is called a *target*; the other characters appearing in the frames are called the *distractors*. For each condition, the frame time is adjusted so that performance is above chance but below complete accuracy. This paradigm, although it uses accuracy as a dependent measure, incorporates many conditions found in earlier search experiments that have used reaction time measures.

To insure that the two basic types of search results appeared in the experiment, the relationship between the memory set and the distractors was varied. Sperling et al. (1971) showed that a number could be located in a series of frames of letters with very small effects of memory-set size and frame size. Proceeding from this result, we decided that one set of our conditions should utilize num-

bers as memory-set items and letters as distractors (and vice versa for other subjects). Another set of otherwise identical conditions used letters (numbers for other subjects) as both memory-set items and distractors. We have termed the number-letter and letter-number conditions *consistent mapping* (CM) conditions, because across trials, the memory-set items (the numbers, say) never appear in the visual display except as targets, and the distractors are never used in the memory set. The number-number (or letter-letter) conditions are termed *varied mapping* (VM) conditions, because the memory-set items on one trial are distractors on other trials, and vice versa. In both VM and CM conditions, however, a new memory set is presented to the subject before each trial. In the CM conditions, of course, the subject may ignore the particular letters or numbers given as a memory set prior to each trial, since any input in a category differing from the distractor category will of necessity be a target.

In this first experiment, we did not attempt to focus the subject's attention in any specified manner. Rather, the subject was instructed to divide his attention as equally and efficiently as possible across all inputs and memory-set items. Subjects were extremely well-practiced and all conditions were blocked, since we wished to discover the limits on the subjects' potential efficiency in these tasks.

### 1. Method

The basic paradigm is illustrated in Figure 2. A sequence of 20 frames was presented on each trial. Each frame consisted of four elements arranged in a square around a central fixation dot. The elements presented could be digits, consonants, or random dot masks. The time from the onset of one frame to the onset of the next is termed the *frame time* (denoted  $f$ ). The frame was presented for all but the first 15 msec of the frame time, and no character or mask was ever presented in the same display position in two successive frames.

The subject's task was to detect any member of the memory set that appeared in the sequence of frames. Targets appeared randomly on one half of the trials. They could appear in any frame except the first three and the last two. The subject was to press a key when he thought he detected a target and to press a different key at trial's end if he did

not detect a target. The frame time was kept constant across all frames of each trial, and the basic dependent variable was the accuracy of the subject's response.

Three primary independent variables were manipulated. The number of characters per frame was set to 1, 2, or 4, in order to explore the connection between selective attention and visual search. This variable is denoted *frame size* (abbreviated  $F$ ) and was held constant during all frames making up a trial. When frame size was less than 4, all noncharacter positions were filled by masks. Character positions were chosen randomly for each frame, with the constraint that no character or mask could appear in the same display positions in successive frames.

The number of characters in the memory set (presented in advance of the trial) was either 1 or 4; it was varied in order to explore the connection between selective attention and memory search. This variable is denoted *memory-set size* (abbreviated  $M$ ). A different memory set was presented in advance of each trial.

Finally, and most important, the relationship between the memory set and the distractor set was varied as described above in order to contrast the two basic search processes, *consistent mapping* and *varied mapping*. In both conditions, five distractors were chosen for each trial, and the distractors for each of the 20 frames of that trial were drawn from these five items.

Four subjects were used for Experiment 1. Two paid undergraduate students (one female) were assigned to a CM condition in which they were to search for digits among consonants. These subjects were also run in a VM condition in which they were to search for consonants among consonants. Two other subjects, one a paid female undergraduate and the other the experimenter, were assigned to a CM condition in which they were to search for consonants among digits. These subjects were also assigned to a VM condition to search for digits among digits. It should be noted that the memory-set items in the CM conditions were never distractors in either condition. However, the distractors in the CM conditions took on both roles in the VM conditions: On some trials they were memory-set items, and on other trials they were distractors. The four subjects had at least 10 hours of practice in pilot work in Experiment 1 before data were collected. All subjects had corrected, 20/20 vision and viewed the displays binocularly.

Each trial began with the presentation of the memory set. Subjects had as much time to memorize the items as they wished, and they started the trial by pushing a start button. At trial's end, a feedback tone was given to subjects to indicate an error. Subjects were encouraged to take frequent rests and were instructed to maintain the highest possible level of accuracy.

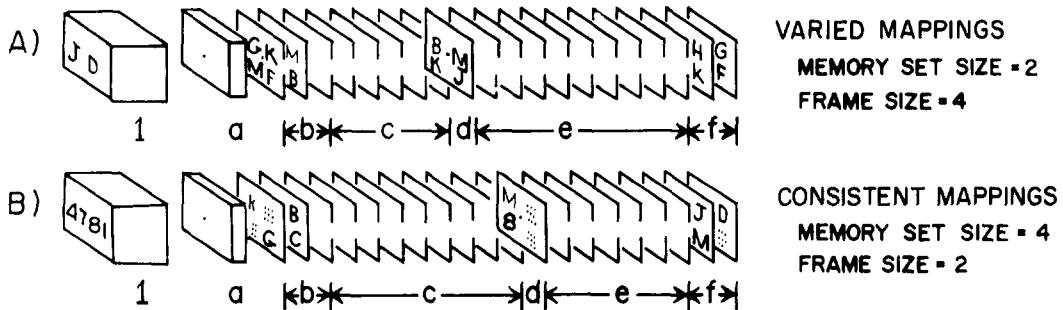


Figure 2. Two examples of a positive trial in the multiple-frame search paradigm of Experiment 1: (A) varied mapping with memory set = (J, D), and (B) consistent mapping with memory set = (4, 7, 8, 1). (1: presentation of memory set; a: fixation dot goes on for .5 sec when subject starts trial; b: three dummy frames that never contain target; c: distractor frames; d: target frame; e: distractor frames; f: dummy frames that never contain target. Frame time is varied across conditions.)

Five independent variables were manipulated: the mapping (CM or VM); memory-set size ( $M = 1$  or  $4$ ); frame size ( $F = 1, 2$ , or  $4$ ); frame time,  $f$ , (3 different times per condition, as shown in Figure 3); and target present or absent—a total of 72 conditions. All variables except the last one were manipulated only between blocks. Thus, there were 36 types of blocks.

The first 30 trials for the first block of the session, and the first 15 trials for each succeeding block, were run for practice; the data were not recorded. Following the practice trials, there were 120 trials in each block, 60 with target present and 60 with target absent, in random order. When the 14 experimental sessions were completed, there were a total of 60 observations in each of the 72 conditions for each subject. Various additional and technical details of the procedure are given in Appendix B.

## 2. Results and Discussion

The results are shown in Figure 3, averaged across the four subjects and broken down by all conditions and frame times. The results for individual subjects varied in overall level of accuracy but showed the same pattern across conditions. (Further discussion of results for individual subjects can be found in Section III.E. of the present paper.) The dependent variables are the probability of a correct detection of a target (called a *hit*) and the probability of a false detection when a target is not in fact present (called a *false alarm*). Because each point is based on 240 observations, the standard error of the mean for each point is at most .032 (restricting our inferences to this set of four subjects

only). All of the trends and results we shall discuss are highly statistically significant, unless we specify otherwise. Also, performance level may be measured in two ways. First, the level is defined for any point by the hit and false alarm probabilities. Second, performance is measured by the frame time needed to reach a given level of accuracy—the longer the frame time needed, the worse the performance.

Note first that performance varied widely across conditions. In fact, it was necessary to use two different scales for the CM and VM conditions in order to graph the results in Figure 3. A horizontal spacing that represents 200 msec in the right-hand panel represents only 40 msec in the left-hand panel. (The data are presented this way, rather than, say, on a logarithmic scale, because simple versions of the models that will be discussed predict performance to be linear with frame time.)

Consider first the VM data in the right-hand panel: (a) For any condition, performance increases monotonically as frame time increases. (b) Performance decreases monotonically as the product of memory-set size and frame size increases (i.e., as  $M \times F$  increases). (c) The  $M = 4, F = 1$  condition is more difficult than the  $F = 4, M = 1$  condition. (d) The false alarm rate is consistently low across conditions—most errors are misses of targets rather than false alarms.

Several conclusions are suggested by these

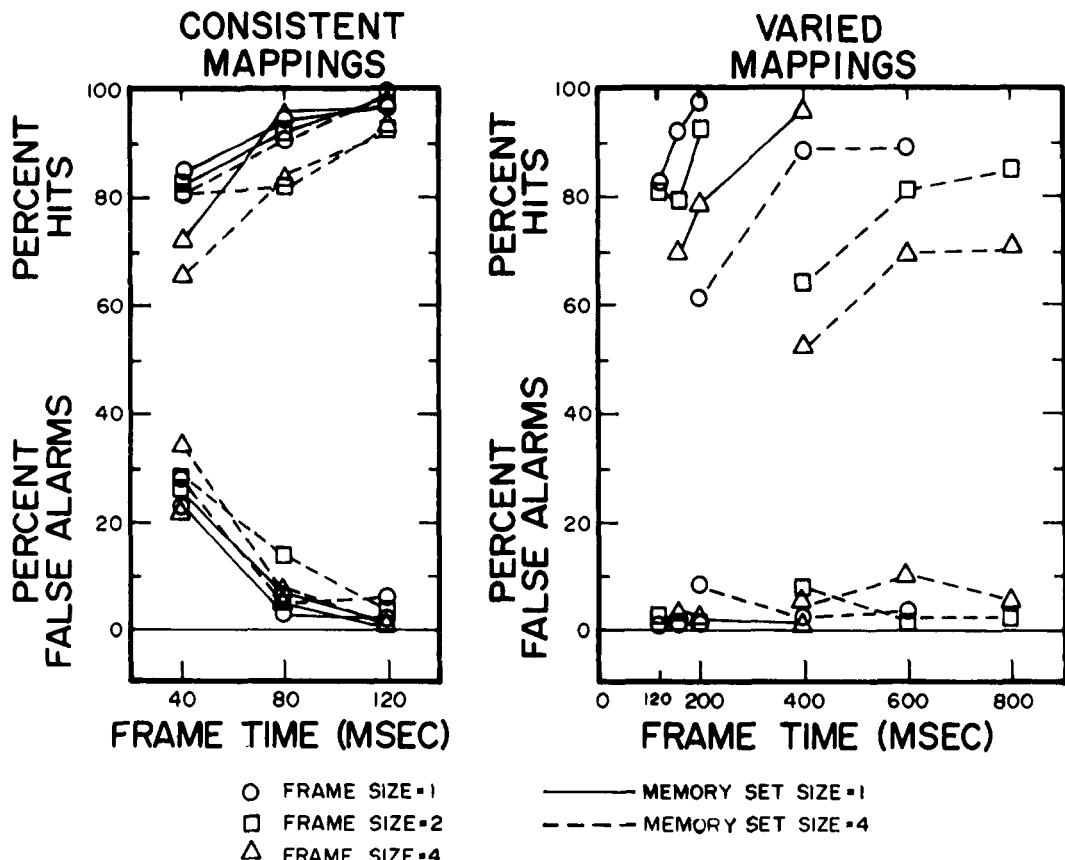


Figure 3. Data from Experiment 1: hits and false alarms as a function of frame time for each of the 12 conditions.

VM data. To begin with, there is an extremely large divided-attention effect: Increasing the load on the short-term search process (by increasing  $M$  or  $F$ ) results in a drastic increase in the difficulty of the task. From the curves in Figure 3, the estimated frame time needed to reach a given performance level (say, .70) appears to range from about 60–800 msec, depending on the load placed on the search process. Such a result is certainly among the largest selective-attention effects ever to be reported and leaves no doubt that the attentional deficits are large enough to be explained on the basis of search limitations. (When thinking of this study in attentional terms, the  $M = 1, F = 1$  condition should be considered to be the most "focused," and the  $M = 4, F = 4$  condition the most "divided.")

The fact that performance is monotonically related to  $M \times F$  suggests a search process that takes longer to complete as the total number of potential comparisons increases. A serial search process appears to be a likely candidate for a model, but we shall temporarily defer discussion of models.

The low false-alarm rate in all conditions suggests that the processing of the features of the individual characters is quite accurate. Inaccurate processing would lead to confusions among characters, as in the model of Gardner (1973); these confusions would cause some background characters to appear as targets and would thereby lead to false alarms. There is some further informal evidence that processing of characters is complete in the slower VM conditions, since at  $f = 800$  msec, subjects can read aloud ac-

curately every character in every frame (even though their hit rate reaches only 70%).

Provisionally, then, it seems reasonable to suppose that in the VM conditions the subject searches each new frame when it appears, detecting the target if the search happens to locate it before the next frame appears. The search in these conditions is highly dependent upon load, that is, upon memory-set size and visual frame size. We shall term the search used in these conditions *controlled search*, although the reasons for this choice of labels will not become fully evident until we have considered several other experiments.

Consider next the CM conditions shown in the left-hand panel of Figure 3. A rather different picture emerges in this case: (a) The levels of performance are much better in the CM conditions than in the VM conditions. Even the easiest VM condition ( $M = 1, F = 1$ ) is more difficult than any of the CM conditions (including  $M = 4, F = 4$ ). (b) There are very small effects of memory-set size and frame size, and the direction of any effects is not monotonic with load (i.e., with  $M \times F$ ). (c) False alarms begin to increase as frame time drops and are especially pronounced at  $f = 40$  msec. (d) Performance increases monotonically as frame time increases.

There are a number of implications of these findings. The fact that performance is much better in these CM conditions than in the VM conditions and the fact that memory size and frame size have only small effects, at most, in the CM conditions, but extremely large effects in the VM conditions, suggest that a different search process is operating in the two cases. Because CM performance does not depend on load, we will denote the process used to carry out the task as *automatic detection* rather than *automatic search*. The rationale behind the use of the term *automatic* will become clearer after several additional experiments have been discussed.

The fact that false alarms take a dramatic upturn in the CM conditions when frame time drops toward 40 msec suggests that perceptual deficits begin to arise at these rapid presentation rates. At slower rates, the

process of encoding features from each display is presumably close to errorless, and performance decrements are presumably due to limitations upon the rate of search and decision. However, at frame times of 40 msec, feature abstraction is apparently incomplete, or in error. Thus, the features derived from a given location either will be insufficient to identify uniquely the character present in that location or will incorrectly denote a character other than the one actually presented. In either event, the hit rate should drop and the false-alarm rate should rise.<sup>8</sup>

The hypotheses that the rate of search limits performance in the VM conditions and that there are two different search and detection processes, called controlled and automatic, are of course not the only possible interpretations of the data from Experiment 1. The differences between the CM and VM conditions could result from a number of other factors that must be explored.

---

<sup>8</sup> The increase in false alarms at  $f = 40$  msec could be due to the misperception of distractors as targets, if processing is error prone. On the other hand, according to one version of the incomplete-processing hypothesis, false alarms would be due to guessing based on an incomplete set of correctly encoded features. Call this the "sophisticated guessing" hypothesis. Either explanation of false alarms would seem to require many more false alarms (and/or fewer hits) when 80 characters are presented ( $F = 4$ ) than when 20 characters are presented ( $F = 1$ ).

The data show, however, that neither the hit rate nor the false-alarm rate changes appreciably when  $F$  changes from 1 to 4. Hence, neither the incomplete-processing model nor the sophisticated-guessing model seems to be capable of handling the results. We suggest therefore that "pure guessing" on a proportion of the nondetection trials may explain the results. Guesses are not made in the VM conditions, since characters are completely processed and the subjective impression is that all characters are clearly seen. When frame time drops toward 40 msec in the CM conditions, then some characters are incompletely encoded. Because there is a subjective impression that characters are not seen clearly, and because the subject knows that targets appear on one half of the trials, the subject is willing to guess that a target is present on some trials when no target is presented. Such guessing causes the number of false alarms to rise.

First, the large differences among the VM conditions might result from the changes, across conditions, in the number and positioning of masks adjacent to, preceding, and following the target. One problem with this explanation is the obvious fact that masking effects would be expected to affect the CM conditions as well. In addition, considering the VM conditions only, it is easy to disconfirm this hypothesis by considering whether characters or masks precede, follow, or are adjacent to a particular target. The results of these considerations are described in Appendix C. They may be summarized simply here: Lateral and temporal masking are not responsible for the performance differences in the VM conditions.

Second, the CM conditions may not be different in kind from the VM conditions, but a ceiling effect may prevent the effects of load from showing. That is, even in the most difficult CM conditions, performance may be as good as it can possibly be. The argument that a ceiling effect prevents differences between the CM conditions from appearing begs the question at best and is nonsensical at worst, since (a) the CM curves lie atop one another over the entire range of the psychometric functions examined and (b) there appears to be no mechanism by which all of these conditions could show such a radical improvement that they are easier than the easiest VM condition ( $F = 1, M = 1$ ).

Third, the same type of search might be occurring in the two conditions, but the presence of a categorical difference between targets and distractors in the CM conditions (i.e., letters vs. digits) might allow the subject to reduce the scope (but not the nature) of the search; for example, the subject may search for the category "letters" rather than for the particular letters presented prior to the current trial. This argument, that the presence of categories enables search to proceed more efficiently, is a highly plausible hypothesis that will be the subject of a number of experimental investigations in Part II. For now, consider only the simple hypothesis that the subject ignores the individual characters presented at the start of each CM trial and searches instead for any instance of

the target category (say, letters). Such a model may indeed be appropriate in some search studies (see Part II, Experiment 3 for an example), but it cannot be the entire story. In the current study, this category hypothesis can explain the failure to find effects of memory-set size (since the memory set on each trial is ignored), but it cannot explain the failure to find effects of frame size. These issues will be explored at length in Part II. For the time being, we shall accept the hypothesis that different detection processes, automatic and controlled, are being utilized in the CM and VM conditions. We shall defer discussion of the characteristics of automatic detection and controlled search until after presentation of Experiment 2 (the reaction time study).

### 3. Summary

The findings may therefore be summarized as follows. The VM conditions demonstrated enormous effects of load, with both memory-set size and frame size having considerable influence on the level of accuracy. Performance was better in all CM conditions than in the easiest VM condition. Furthermore, the various CM conditions differed little from each other—there was virtually no effect of load. We suggest that different processes, called automatic detection and controlled search, were being utilized in the CM and VM conditions, respectively. In addition, although the present data are only suggestive, we shall see later that the development of the ability to utilize automatic detection depends on the use of a training procedure in which stimuli can be consistently mapped to responses.

### B. A Selective Review of Related Research Results

How do the results of Experiment 1 compare with those already in the literature? The closest study to the present one is that of Sperling et al. (1971). That study, on which the present paradigm was based in part, explored a number of conditions similar

to those we have described as consistent-mapping conditions. The subjects all searched for the presence of a digit in a long sequence of frames of letters, and the frame time and frame size were systematically varied across blocks. Every trial did contain a target, and the subject tried to say both which digit occurred and in what display position the digit appeared. Since 2 and 4 were among the frame sizes Sperling et al. examined, we can compare their results with ours for the same frame size.

The results for two of their subjects (with  $M = 1$ ) were in reasonable agreement with ours in that errors first appeared when frame times dropped to about 40–60 msec. However, these results differed from ours in that detection probability at equivalent frame times was, in general, poorer for larger frame sizes. It is possible that these performance differences with frame size were due to changes in display configurations as frame size increased, since the unused display positions were not filled with masks, as they were in our study. Thus, larger displays covered larger areas, so that acuity could have been lower, and in larger displays, targets tend to be surrounded by more characters, so that lateral masking could have reduced acuity.

An alternative explanation of the frame-size results reported by Sperling et al. is that their subjects had not been trained to reach asymptotic performance—they were given only about 60 trials per condition. (As will be shown in Part II, practice effects are very large, and the development of automatic detection to asymptote can take a great number of trials.) This practice interpretation is supported by the fact that one of their subjects who was given 1,300 additional trials was still improving at the end of the series. On these 1,300 trials, frame size was 9 and frame time was 60 msec; performance in an  $M = 1$  condition was compared to performance in an  $M = 10$  condition (i.e., known vs. unknown digit). During these trials, detection was higher when  $M = 1$  than when  $M = 10$  (.46 vs .38). However, in an additional 1,000 practice trials, there were indications that the difference between the two conditions was decreasing. Thus, with increasing practice,

the results moved into closer agreement with ours, since we found little effect of  $M$  or  $F$ .

Sperling and his colleagues concluded that search was occurring in parallel through the characters in each frame, but with search efficiency dropping as display size increased. Our present results certainly are consistent with the hypothesis that automatic detection occurs in parallel. However, our evidence indicates that detection efficiency drops very little with increases in frame size or memory-set size (at least up to  $M = 4$  and  $F = 4$ ). It is quite possible that performance in Sperling et al.'s study would not have decreased with increasing frame size had the subjects been better practiced and had acuity and lateral masking been made comparable across conditions (although there must surely be some absolute limits on how large  $M$  and  $F$  can become without a drop in the efficiency of automatic detection).

A second set of studies related to Experiment 1 deal with tachistoscopic detection (e.g., to name just a few, Eriksen & Spencer, 1969; Estes, 1972; Estes & Taylor, 1964, 1966; Shiffrin & Gardner, 1972). In tachistoscopic detection studies, only one frame is presented on each trial, sometimes followed or preceded by masks. In the studies in which accuracy is a measure, the subjects are induced to make errors through a reduction in the presentation time (or intensity) to the point where individual characters in the display are incompletely processed. In this case, letters presented are represented in the subject's short-term memory by a set of features that are often insufficient to discriminate among several possible inputs.

Shiffrin and Geisler (1973) built a model for this situation, based on earlier work by Gardner (1973), Shiffrin and Gardner (1972), and Estes (1972). Suppose the subject must decide whether a display contains a "B" or an "F," and assume that the subject has some chance of correctly perceiving any input as a whole letter. Further assume that the subject first searches serially through all locations that are processed to the letter level. If the target is in one of these locations, then a correct response occurs. If the target is not found among the completely processed

letters, then a serial search is made through the locations containing incomplete aggregates of features, starting with the locations having the greatest density of features and proceeding in order of decreasing density. Eventually, an aggregate of features is found that is a subset of the features of one of the target letters, search is terminated, and an appropriate response is made.

In light of the evidence in the present paper, we would like to propose that a variant of this model deserves consideration for CM situations like those presented by Estes (1972) that were modeled by Shiffrin and Geisler (1973). Namely, we suggest that for CM studies, automatic detection (i.e., parallel processing independent of load) might be utilized whenever the target is processed to the letter level. When the target is processed incompletely, however, we see no objection to the search strategy suggested by Shiffrin and Geisler (1973). In fact, it is shown in their paper that due to the confusions caused by incomplete processing their model predicts performance to be highly dependent upon memory-set size and frame size. Such effects have been found in numerous threshold detection studies.

We are arguing that performance decrements in threshold detection studies should be expected when  $M$  or  $F$  is increased, due to confusions arising from incomplete processing. Why, then, are not similar effects of load seen in our CM conditions when frame time drops to 40 msec and letters are incompletely processed? Most likely, the multiple-frame procedure and rapid rates we utilized prevented subjects from going through the time-consuming process of matching sets of partial features in the display to memory-set items. They could only do so if the presentation rates were slowed to those seen in the slowest VM conditions (but then, of course, processing of all letters would be complete). Thus, in our CM conditions, at rapid presentation rates, those targets that do not manage to be completely processed will be ignored by the subject, thereby eliminating effects based on confusions. We do find false alarms in our data, particularly when  $f = 40$  msec, but we suggest that these are most likely caused

by pure guesses rather than by confusions (see Footnote 3).

In summary, we propose that the single-frame procedure of most tachistoscopic detection studies allows the subject to base his decisions upon both complete character encodings and partial character encodings: The incomplete features are examined in a slow, difficult, feature-matching process that is carried out after presentation but before memory fades. Such feature matching improves performance over the level that would obtain were only the complete encodings used for a decision. In addition, analysis of incomplete encodings makes performance dependent on memory-set size and display size. However, the rapid presentation rates in our multiple-frame procedure prevent the subject from utilizing feature matching, and hence, the effects of memory-set size or frame size are reduced or eliminated. By this reasoning, we can eliminate the apparent inconsistency between *single-frame* CM threshold studies that show strong effects of load and our *multiple-frame* CM threshold studies that show little effect of load.

Many search studies related to Experiment 1 utilize a technique in which extended visual search is required. In one paradigm, visual characters appear on a succession of cards, and the cards must be sorted as quickly as possible on the basis of the presence or absence of a member of the memory set (e.g., see Neisser, 1974; Rabbitt, 1964, 1967). In another paradigm, the subjects search as quickly as possible through a column of lines of characters, attempting to locate the presence of a member of the memory set (e.g., see Neisser, 1967). Neisser's studies always utilized a CM design, since the memory-set items were never distractors (and vice versa). The search rate in each study was estimated from the time to locate a given target divided by the number of distractors preceding the target (in reading order). A representative sample of these results is discussed next.

Search for a single, novel, or at least relatively unpracticed, memory-set item embedded in a background of similar characters is quite difficult and slow, requiring perhaps

100 msec per item or more (Neisser, 1963; also, similar results were found by Kristofferson, Groen, & Kristofferson, 1973, in a VM paradigm). The slow search rate for novel, unpracticed memory-set items, or for items in VM situations, probably reflects the use of what we have termed controlled search.

On the other hand, with extended CM training, search becomes markedly faster, reaching rates of at least 60 characters per second (Neisser, 1963). This improvement in search rate with practice in CM situations probably reflects the learning process by which automatic detection is developed, and Neisser's practiced subjects are probably utilizing what we have termed automatic detection. The rate of search for practiced subjects found by Neisser is somewhat slower than that found by Sperling et al. (1971) and by us in the CM conditions of Experiment 1. However, Neisser's search studies required eye movements that may have slowed the search considerably. Eye movements are not needed in the multiple-frame procedure.

In several of Neisser's (1963) studies, and in the studies by Rabbitt (1964, 1967), the number of characters per line was varied (see also Teichner & Krebs, 1974). Care must be taken, however, in equating this variable with the frame-size variable in Experiment 1. Our study filled in all non-utilized display positions with masks; the physical size and arrangement were the same in all frame-size conditions; and the characters were randomly placed within each frame. However, all of these factors differed across frame sizes in the previous studies. Furthermore, eye movements may have limited the search rate in the Neisser paradigm, and motor movements and eye movements may have limited the search rate in the Rabbitt paradigm. In these paradigms, variations in frame size affect the number of characters available per fixation and, hence, affect performance. Thus, although Neisser found that search was somewhat more efficient with 6 than with 2 characters per line (whereas we found little effect of frame size), it is difficult to know which factors are responsible for this effect.

Although memory-set size greatly affects the search rate in VM situations and even for novel or unpracticed targets in CM situations, the effects of memory-set size after extended CM practice remain in some dispute. Neisser, Novick, and Lazar (1963) showed that extended practice caused the search rate for a large memory set to become the same as that for the "slowest" single member of that set. Yonas and Pittenger (1973) and Neisser (1974) replicated this finding, but they also showed that error rates were higher for the slowest member of the large memory set when it was part of a large memory set than when it was the only member of the memory set. Finally, Kristofferson (1972c) showed a convergence of performance with practice, but smaller set sizes still retained some superiority. (Also, see Rabbitt, 1964).

The finding in CM conditions that search rate for a large memory set becomes equal to that for the slowest single member of that set may be consistent with our finding that memory-set size has little effect on the accuracy of automatic detection. If the stimuli we utilized in the memory set did not vary much in their detectability, then the accuracy in our  $M = 4$  condition might indeed be governed by the search rate for the slowest target (which would be about the same rate for any target). This hypothesis cannot be tested directly for our data, because we did not collect the data in such a way that it could be broken down for particular memory-set items.

In summary, then, the visual search paradigms we have discussed give results reasonably consistent with the results from our CM and VM conditions, despite large variations in the paradigms. Conversely, our results and theorizing allow us to organize and understand the various findings in the literature.

#### C. A Single-Frame Attention and Search Task

The goal of Experiment 2 was the development of a direct link between accuracy measures and reaction time measures and between attention and search tasks. To this end,

the paradigm of Experiment 1 was modified slightly so that only a single frame contained characters on each trial. The subject was expected to make very few or no errors, and his response time was the measure of interest. We anticipated that the data would suggest a model for the controlled-search process in the VM conditions. We also hoped that the search process fitting the data of Experiment 2 could be used to fit the results of Experiment 1.

Experiment 2 was also designed with the hope that the results would clear up a number of perplexing findings in previous search studies utilizing reaction time measures. In particular, we intended that the results should delineate the factors governing the shape of the frame-size and memory-size functions. As discussed in the introduction, these functions are often linear with slopes of about 40 msec per item, but sometimes they are curvilinear with a lower slope or even flat. In order to achieve the various goals, a paradigm was chosen that manipulated the same independent variables as did Experiment 1.

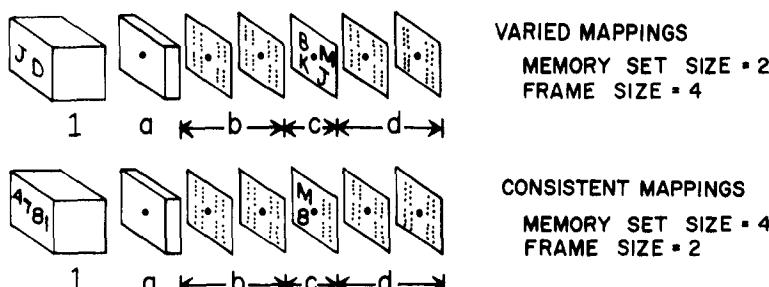
### 1. Method

All procedural details were the same as those in Experiment 1 except for those given below. The basic procedure is depicted in Figure 4. Five frames were presented on each trial, preceded by the presentation of a memory set. The first two frames and the last two frames contained masks only, four masks in each case. The middle frame contained some combination of targets, distractors, and masks. Any noncharacter positions in the middle

frame were filled by masks so that there were always four elements in each frame. The subject was given instructions to maintain high accuracy but to give one of two responses as quickly as possible, indicating whether or not any item from the memory set appeared in the frame. Memory-set sizes were 1, 2, and 4; frame sizes were 1, 2, and 4; and the mapping was VM or CM, as in Experiment 1 (although  $M = 2$  was not examined in the CM condition). All conditions except target present or absent were changed between blocks. Practice blocks and trials were used prior to the experimental blocks, and blocks were presented in random order. There were 120 observations for each subject for each condition (hence, there were 480 trials in each point graphed in Figures 5 and 6). Incorrect responses were signaled to the subject by a tone. Only the mean and variance of the reaction times for correct responses were collected in each block and condition.

The frame time in all conditions was 160 msec. This frame time is much lower than that needed in many of the conditions of Experiment 1, but performance was expected to be accurate nonetheless because the target frame was not followed by any additional frames requiring processing. To maintain accuracy during his search, the subject needed only to retain in short-term memory the memory set and up to four display items. We had no doubt that the characters could be perceived and retained at a frame time of 160 msec because we carried out an experiment in which frame time was varied and the subject attempted to report all the characters from a single frame of four characters (preceded and followed by masks). Errors in report were not made until the frame time dropped to 80 msec or less. Details are given in Appendix D.

The same four subjects from Experiment 1 were used. Furthermore, these subjects were kept in comparable conditions. Thus, a subject who in Experiment 1 searched for consonants in digits (in the CM condition) did so again in Experiment 2, and the same was true for the other conditions and sub-



*Figure 4.* Two examples of a positive trial in the single-frame search paradigm of Experiment 2: varied mapping with memory set = J, D) and consistent mapping with memory set = (4, 7, 8, 1). (1: presentation of memory set; a: fixation dot goes on for .5 sec when subject starts trial; b: two frames of masks; c: target frame; d: two postmask frames. Frame time = 160 msec for each of the five frames.)

jects. Other procedural details may be found in Appendix E.

## 2. Results and Discussion: Mean Reaction Time

The mean reaction times for correct responses in all conditions, averaged over subjects, are graphed as a function of memory-set size in Figure 5 and as a function of frame size in Figure 6. The same data are simply replotted in the two figures for convenience. The error rates are shown in Figure 5 as bars below the horizontal axis.

For the CM conditions, there was almost no change in mean reaction time as a function of either memory-set size or frame size. This result parallels the CM findings of Experiment 1 and suggests the same conclusion: In CM conditions, the subject utilizes automatic detection that is unaffected by load.

The VM conditions show large effects of both memory-set size and frame size. Roughly

speaking, the functions are linear with slopes that are about twice as large for the negative as for the positive trials.

## 3. A Single-Frame Task Using Letters Only

Before discussing these results further, we shall present data from a closely related study by Briggs and Johnsen (1973). Their VM conditions utilized a total of 10 letters from which the memory set (of size 1, 2, or 4) was randomly chosen prior to each block of 48 trials. Thus, although the memory sets and distractor sets did not vary and intermix on every trial, as in our study, they did shift every 48 trials. On each trial, a random choice from frame sizes of 1, 2, and 4 was made. There were two blocks at each memory-set size per session, and there were four sessions altogether.

Their CM conditions differed somewhat from ours in that the target and distractor sets both consisted of letters and the sub-

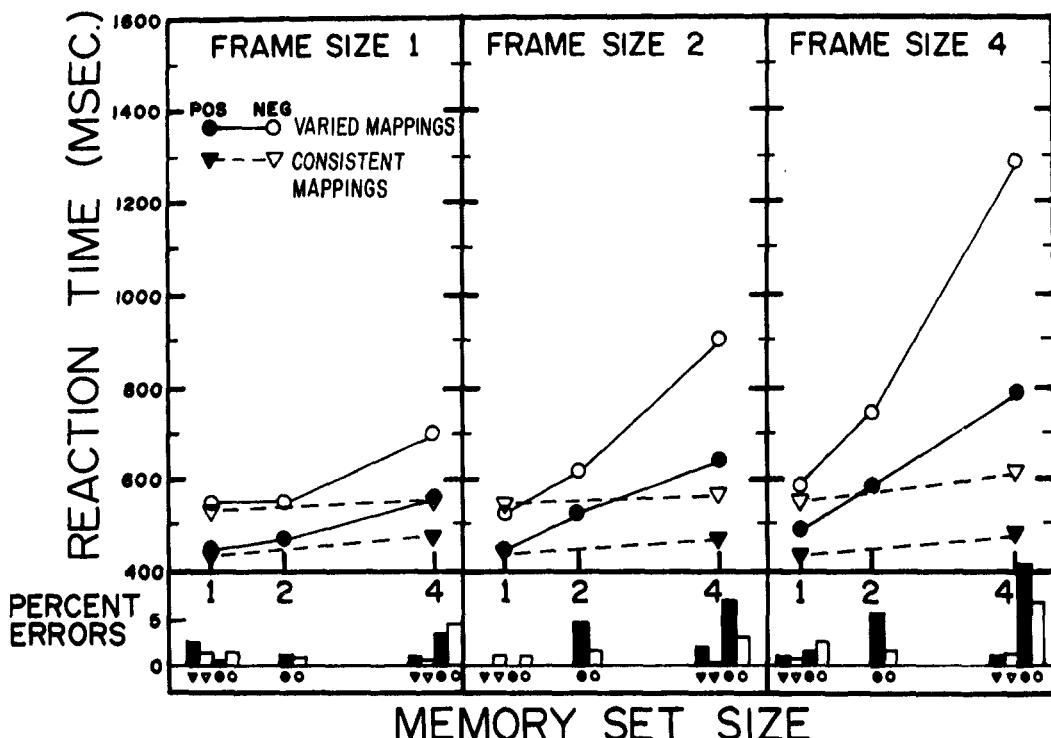


Figure 5. Data from Experiment 2: mean reaction times for correct responses, and percentages of error, as a function of memory-set size, for all conditions.

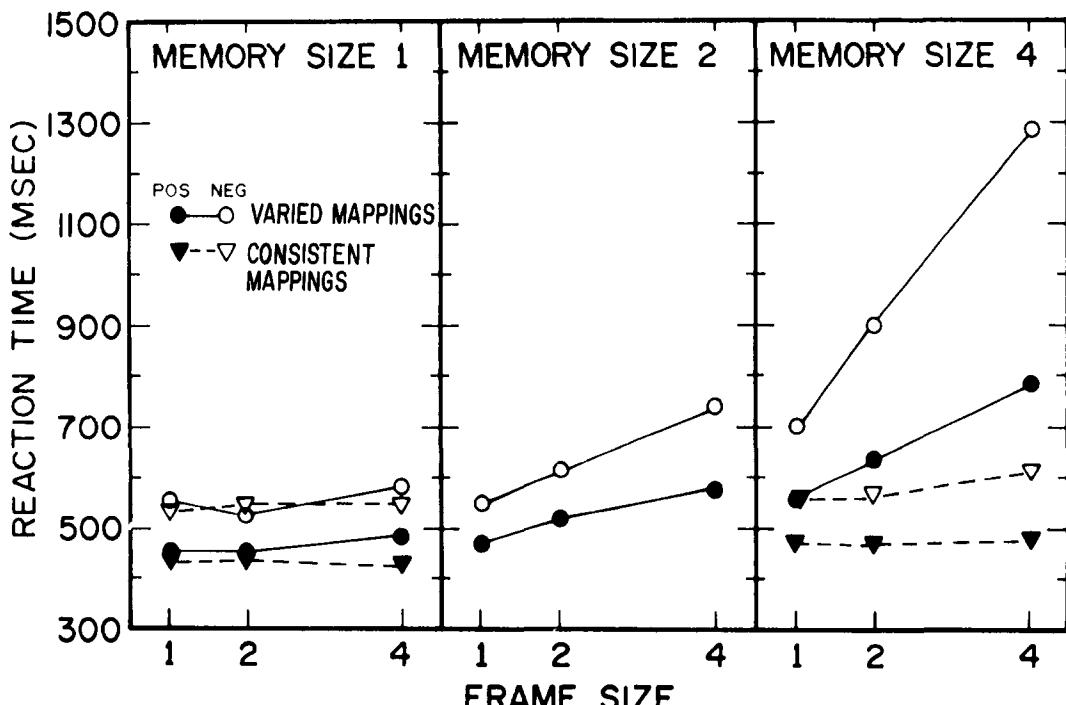


Figure 6. Data from Experiment 2: mean reaction times for correct responses, as a function of frame size, for all conditions. (The data are regraphed from Figure 5 for convenience.)

jects were given considerably less practice: Each subject was trained for four sessions with two blocks of 48 trials per session at each memory-set size. Actually, two groups of CM subjects were run, but at this point we shall discuss only the condition that was closest to our CM conditions. There were seven fixed letters from which all memory sets were chosen and eight fixed letters from which all distractors were chosen. Prior to each block, an appropriate number of letters was chosen from the population of seven, and this set was the memory set for the next block of trials. As in the VM condition, memory-set size was either 1, 2, or 4, and frame size was either 1, 2, or 4.

Briggs and Johnsen graphed their data in an inconvenient fashion from our present viewpoint, so we have regraphed their data in the same fashion as our Figures 5 and 6. The results are shown in Figures 7 and 8.

The results of the VM conditions of Briggs and Johnsen are remarkably similar in pattern to those we have presented in the pre-

vious figures. Compared with our subjects and results, their negative display functions are perhaps more linear, their reaction times are slower overall, their set-size functions are steeper, and their error rates are somewhat lower. Nevertheless, in accord with our data, Briggs and Johnsen found roughly linear functions and found the negative slopes to be about twice the positive slopes.

The CM results of Briggs and Johnsen are uniformly faster than the VM results (whereas our  $M = 1, F = 1$  data are about the same for VM and CM conditions), but too much weight should not be attached to this finding, since their VM and CM results came from different subjects. More importantly, their CM functions show a definite dependence on  $M$  and  $F$ , though with lower slopes than the VM functions (whereas our CM functions are essentially flat).

Probably because their CM and VM functions did not appear qualitatively discrepant, Briggs and Johnsen attempted to construct a single-process model to account for both

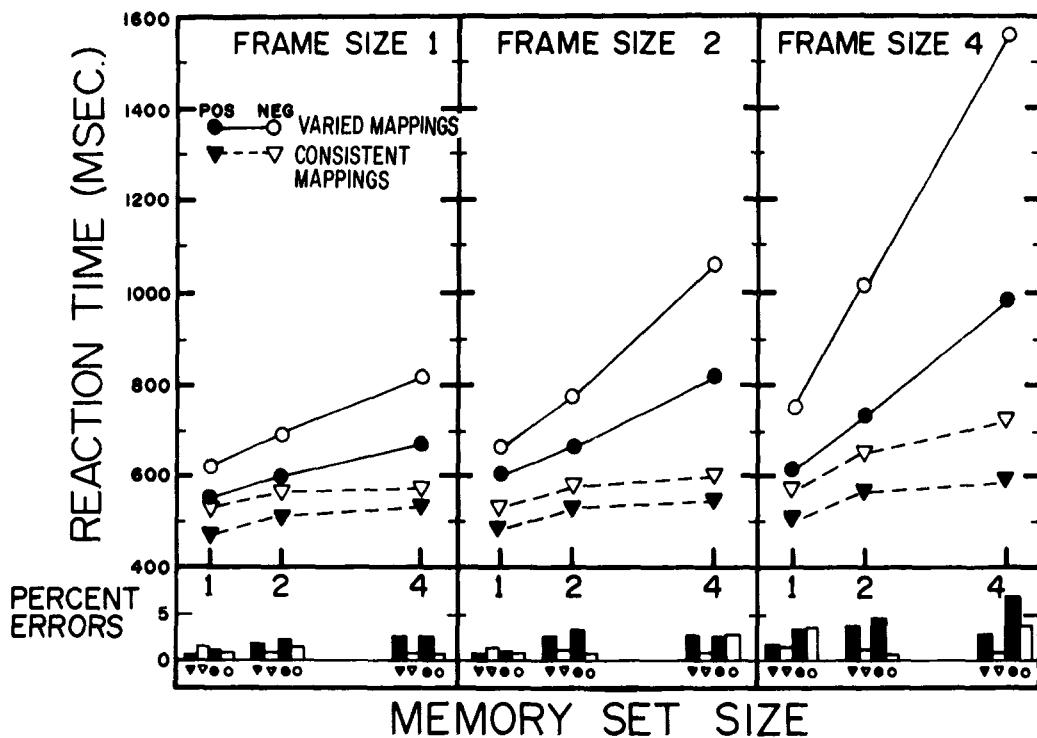


Figure 7. Data from a study by Briggs and Johnsen (1973) that mimics Experiment 2: mean reaction times for correct responses, and percentages of error, as a function of memory-set size, for all conditions.

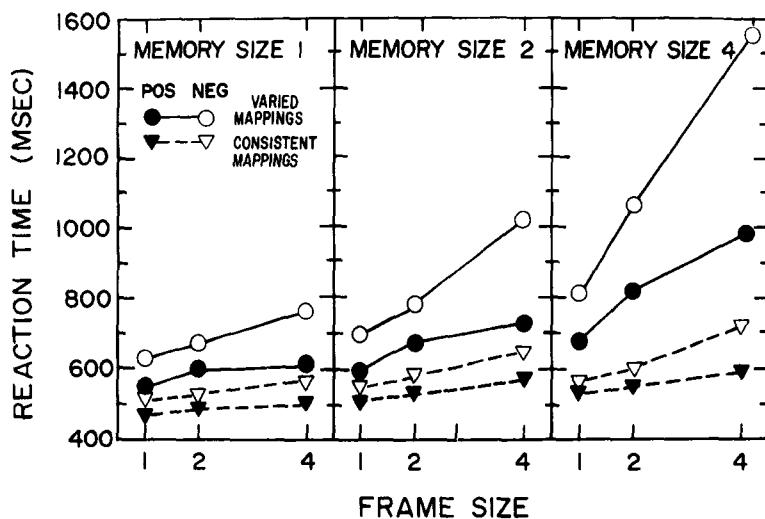


Figure 8. Data from Briggs and Johnsen (1973): mean reaction times for correct responses, as a function of frame size, for all conditions. (The data are regraphed from Figure 7 for convenience.)

sets of data. They proposed a model in which search was related to the informational value of the stimulus ensembles and in which a variable number of rechecking operations occurred in different conditions. In light of our data, however, it seems likely that Briggs and Johnsen's CM functions showed increases with load because their subjects had not been given sufficient practice. After all, it might be expected that automatic detection would take much longer to develop for two subsets of letters not preexperimentally categorized than for digits and consonants already known as categories. If this explanation is correct, then our two-process model might be quite appropriate for the data from their experiment also.

#### 4. Results and Discussion: Variance of the Reaction Time

The variance of the reaction times for the conditions in Experiment 2 are presented in Figure 9. Each point is the average of the variances calculated for each of the four subjects.

For the CM conditions, the variances are relatively unaffected by load, and the variances for negative responses are consistently slightly higher than those for positive responses. This result parallels those for the means and suggests again that the subject utilizes an automatic-detection mechanism in these conditions that is affected very little by load.

The variances for the VM conditions show a very large effect of load but with one very important difference from the means. Namely, when the load becomes large, the positive variance begins to exceed considerably the negative variance, a result opposite to that for the means. It will be shown shortly that this result is to be expected from a search process that terminates when a match is found.

#### D. Quantitative Models of Controlled Search

We wish to demonstrate that the same processes underlie performance in Experiments 1 and 2—that is, that the same search and detection mechanisms are utilized in

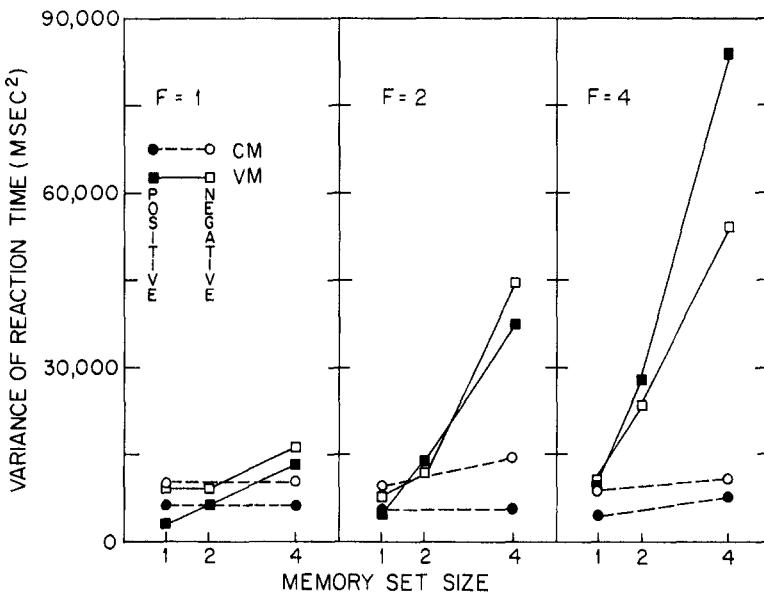


Figure 9. Data from Experiment 2: variances of reaction times for correct responses, as a function of memory-set size, for all conditions. (CM = consistent mapping; VM = varied mapping; F = frame size.)

accuracy and reaction time tasks and in attention and search tasks. In order to demonstrate these facts, we shall fit a quantitative model to the results of both experiments. Because the CM results in both studies show essentially no variations with load, it is uninteresting to model these results, though it seems reasonable that the same detection mechanism is operating in both cases. Thus, the models discussed in this section will be fit to the VM data only.

The organization of this section is as follows: The assumptions of the best-fitting model are described first, and the fit to the data from the two studies is presented. Then an extensive description of alternative models, of the fitting procedures, and of the implications of the modeling are presented.

### *1. A Serial, Terminating Model for Controlled Search*

In either Experiment 1 or 2 the subject is assumed to search each frame in the following fashion:

1. Comparisons are made individually and serially of each member of the memory set against each item in the frame.
2. The order of comparisons is as follows: A memory-set item is chosen and compared to each frame item in turn; then a new memory-set item is chosen and the search continues as before. Each comparison and each choice of a new memory-set item requires some time.
3. Search terminates when and if a match is found.
4. In a single-frame task, the subject initiates a positive response when a match is found or a negative response after all comparisons are completed without a match. The reaction time is the sum of a base time plus the time for each of the comparisons, plus the time for each choice of a new memory-set item.
5. In a multiple-frame task, a target is found if a match occurs in the target frame before the next frame begins. However, some time is needed to switch from frame to frame so that the available time to search each

frame is a constant amount less than the frame time.

The quantitative assumptions are as follows:

1. The time to make each comparison is distributed with a mean of  $c + d + \delta$  msec and a variance of  $\sigma_c^2$  msec<sup>2</sup>.
2. The time to choose each new memory-set item is distributed with a mean of  $\mu$  msec and a variance of  $\sigma_\mu^2$  msec<sup>2</sup>.

To predict the reaction time data of Experiment 2, it is assumed that the positive base time is distributed with a mean of  $B_P$  and a variance of  $\sigma_P^2$  and that the negative base time is distributed with a mean of  $B_N$  and a variance of  $\sigma_N^2$ .

To predict the accuracy data of Experiment 1, it is assumed that the comparison-time and switching-time distributions have the form of a generalized gamma. It is assumed that the time to switch from frame to frame is a constant,  $B_A$ .

The fit to the mean reaction time data of Experiment 2 is shown in Figure 10 for the following parameter values:  $B_P = 347$ ,  $B_N = 397$ ,  $c + d + \delta = 42$ ,  $\mu = 40$ . The fit to the mean reaction time data from Briggs and Johnsen (1973) is shown in Figure 11 for the following parameter values:  $B_P = 468$ ,  $B_N = 522$ ,  $c + d + \delta = 57$ ,  $\mu = 27$ .

The fit to the variance data from Experiment 2 is shown in Figure 12. The values of  $c + d + \delta$  and  $\mu$  are the same as those fit to the means; the values of the variance parameters are:  $\sigma_c^2 = 2,870$ ,  $\sigma_\mu^2 = 2,760$ ,  $\sigma_P^2 = 0$ ,  $\sigma_N^2 = 1,675$ .

The fit to the accuracy data of Experiment 1 is shown in Figure 13. The parameters are the same as those used to fit the reaction time data, with the addition of  $B_A = 45$ . In Figure 14 the fit is shown when  $\sigma_c^2 = 4,200$  and  $\sigma_\mu^2 = 11,428$ , but the other parameters are the same as in the preceding figure (Figure 13).

A perusal of these figures make it clear that this serial, terminating model captures most of the features of the data and supports our contention that the same controlled search process underlies performance in the multiple-

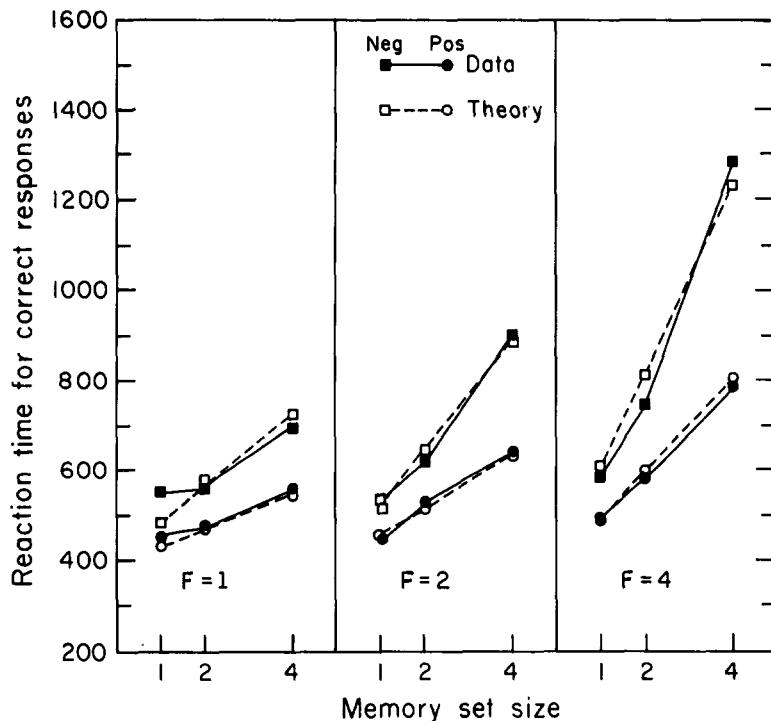


Figure 10. The best fitting predictions of Model 1a for the mean reaction times from the varied-mapping conditions of Experiment 2. ( $F$  = frame size.)

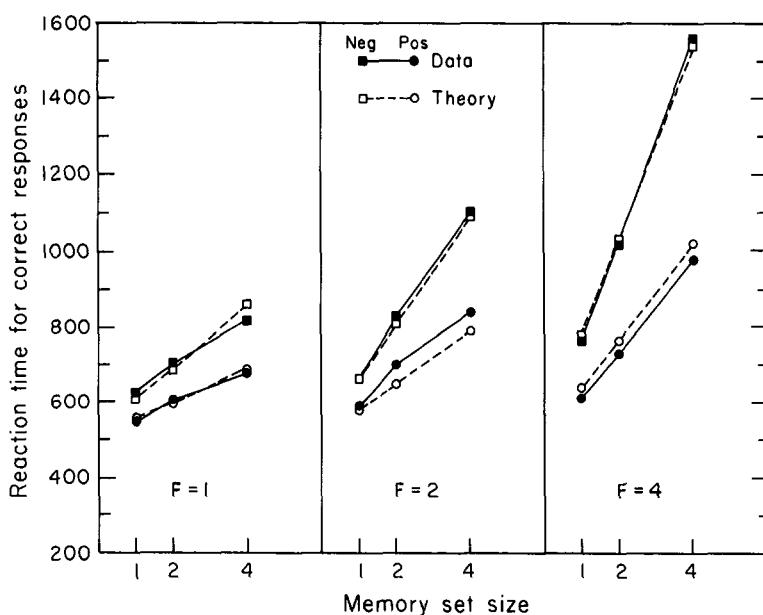


Figure 11. The best fitting predictions of Model 1a for the mean reaction times from the varied-mapping conditions from Briggs and Johnsen (1973). ( $F$  = frame size.)

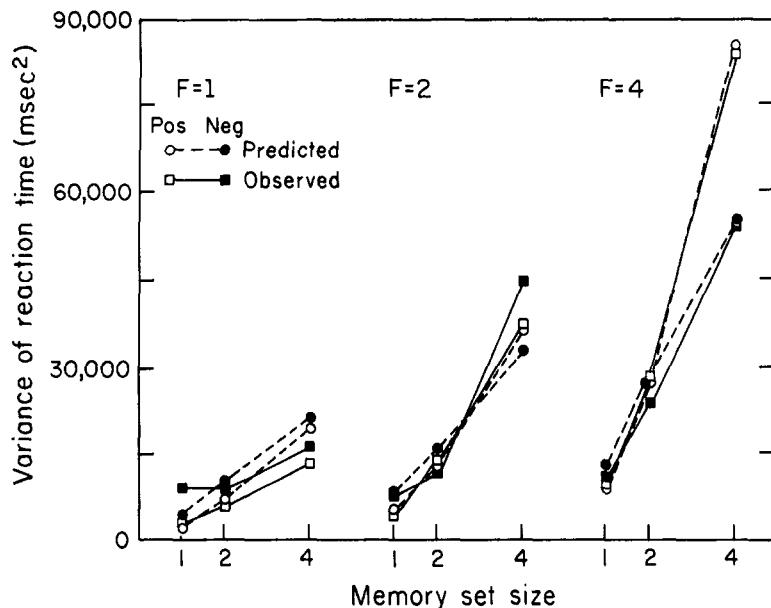


Figure 12. The best fitting predictions of Model 1a for the variance of the reaction times from the varied-mapping conditions from Experiment 2. ( $F$  = frame size.)

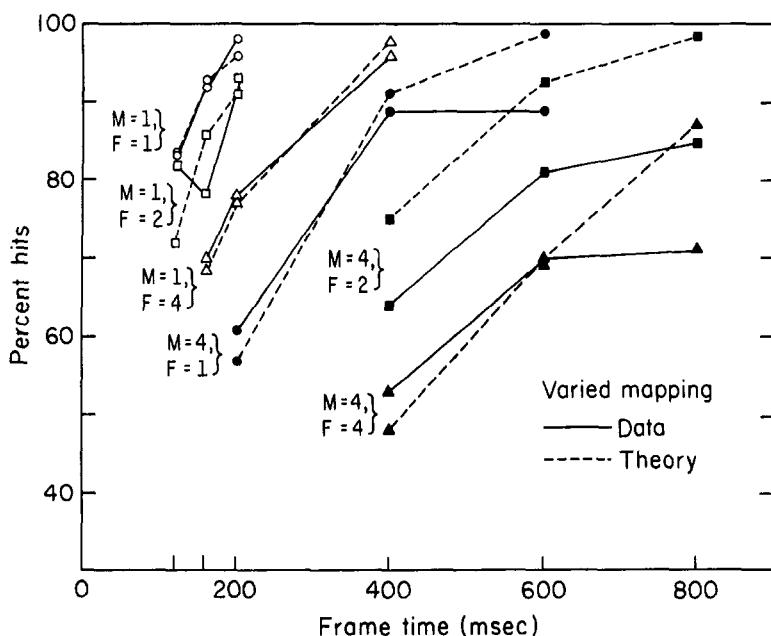


Figure 13. The predictions of Model 1a for the varied-mapping accuracy data from Experiment 1. The parameters of the model were chosen from the fit to the Experiment 2 data (except for a single parameter representing frame switching time, which was constant across conditions). ( $M$  = memory-set size;  $F$  = frame size.)

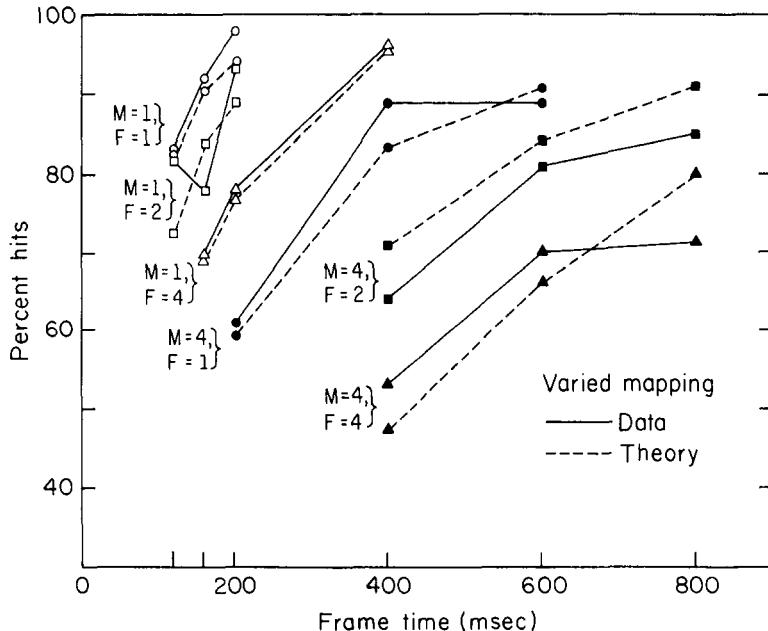


Figure 14. The predictions of Model 1a for the varied-mapping accuracy data from Experiment 1. The parameters of the model were chosen from the fit to the Experiment 2 data, except for a parameter representing frame switching time, and two parameters representing the variances of the two underlying distributions, which were chosen to give a best fit to the data. ( $M$  = memory-set size;  $F$  = frame size.)

frame accuracy task and the single-frame reaction time task—that is, that the same search process underlies divided attention limitations and search limitations.

In the following sections, we shall describe the models, some alternatives, and the important aspects of the data and the predictions.

## 2. Models for VM Reaction Time Data From the Single-Frame Task

It is our general contention that in the situations causing the subject to adopt controlled search (such as those varying the mapping of stimuli to responses), a variety of search schemes are available. In a few situations, the subject is able to choose a search scheme with relative freedom, while in most situations, the paradigmatic constraints lead almost all subjects to adopt similar search strategies. Thus, differences in mode of search would normally be expected between experi-

ments, but reasonable consistency across subjects would be expected within experiments. Our subjects seem fairly consistent (e.g., see Section III.E. of the present paper), and our data agree well with those of Briggs and Johnsen, so it does not seem unreasonable to assume that situational constraints lead subjects to adopt similar search processes.

The most striking features of the VM data in Figures 5, 6, 7, and 8 are (a) the reasonable linearity of frame-size and memory-set-size functions, (b) negative functions with slopes about twice those of positive functions, and (c) monotonically increasing response times as a function of load. By following the reasoning of Sternberg (1966; 1969a, 1969b; 1975), the following implications may be drawn from these data.

First, the linearity of the set-size functions suggests that subjects are using a serial search process in which comparisons of memory-set items of frame items are made sequentially, with each comparison beginning

when the preceding comparison ends. Alternative models can, of course, predict linear functions. For example, Atkinson, Holmgren, and Juola (1969) suggested a search model in which all items are simultaneously compared but in which each comparison proceeds at a rate inversely proportional to the total number of comparisons. Townsend (1971, 1972) extended these arguments and demonstrated quantitatively that certain types of parallel search models may predict linear set-size functions. Thus, we shall restrict ourselves to the weaker conclusions of Sternberg: first, that a parallel search with a per item search rate independent of load cannot predict the observed results, and second, that a serial item-by-item search does provide a simple, parsimonious model.

The second implication of our results derives from the two-to-one ratio of slopes for negative and positive functions. This result is inconsistent with an exhaustive search model in which all possible comparisons are finished before a negative or positive response is made (even if a match is located early in the sequence of comparisons). Rather, the two-to-one slope ratio suggests a serial comparison process which terminates when a match is found, since on the average, if the target is present, a match will be found half way through the sequence of all possible comparisons. Rather strong confirming evidence for a terminating search model can be seen in the variances of the response times for each condition, which are shown in Figure 9. Namely, the positive variance becomes larger than the negative variance when the load is large, a fact predicted by terminating search models (as shown later in this section).

The third implication of the obtained results derives from the finding that mean reaction time is a monotonically increasing function of  $M \times F$ . This fact suggests that the comparison process involves comparing each possible memory-set item against each frame item. To explore this point further, we replotted the reaction times as a function of the total number of comparisons needed for a complete search, namely the product of memory-set size and frame size ( $M \times F$ ).

A straight line was then fitted to the positive and negative functions for both experiments. The results are shown in Figures 15 and 16 for our data and for Briggs and Johnsen's data, respectively. There are a number of cases in these figures where the total number of comparisons is the same for different combinations of  $M$  and  $F$ . In each such case, the values of  $M$  and  $F$ , in that order, are indicated adjacent to the data point.

Although both the overall dependence of reaction time on total number of comparisons and the two-to-one slope ratios for negative and positive trials are well supported by these figures, there are important deviations from the best fitting straight lines. In particular, in every instance in which there are several points representing the same product,  $M \times F$ , the reaction times are slower when the memory-set size is larger than the frame size, and this effect is quite substantial. Obviously, then, the search process does not depend solely on the total number of comparisons to be made. The data show that there is additional time used during the search for every additional memory-set item (which could reflect switching time between memory-set items).

*Models for the mean reaction times.* At this point we decided to explore systematically and quantitatively a class of possible serial search models. The general assumptions are as follows:

1. There is one base time, with mean  $B_P$ , for positive trials, and another base time, with mean  $B_N$ , for negative trials. These represent motor response times and any other components of the response process that do not vary across conditions and, in particular, that do not vary with  $M$  or  $F$ .

2. The basic unit of the search process is a comparison of one member of the memory set against one member of the frame set. The time to make any one such comparison has a distribution with a mean of  $c$  msec.

3. The comparisons are made sequentially (i.e., they do not overlap in time).

4. At various times during the search, depending on the model, a decision is made concerning whether a match has been found be-

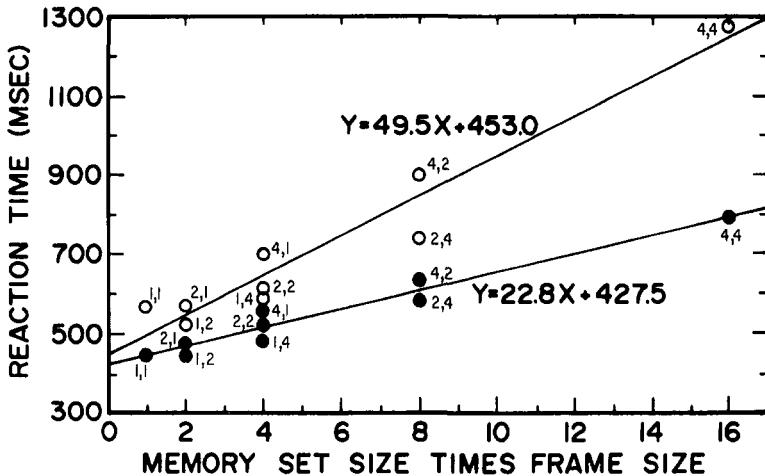


Figure 15. Data from Experiment 2: mean reaction time as a function of memory-set size times frame size, for negative trials (open circles) and positive trials (closed circles). The first number adjacent to each point designates the memory-set size; the second designates the frame size. The best fitting lines are depicted.

tween a memory-set item and a frame item. The time to make any one such matching decision has a distribution with a mean of  $d$  msec.

5. During the course of the search, there are occasions when it is necessary to switch

from a comparison involving one memory-set item to a comparison involving another, or from a comparison involving one frame item to a comparison involving another. The time to switch from one frame item to the next has a distribution with a mean of  $\delta$  msec, and the

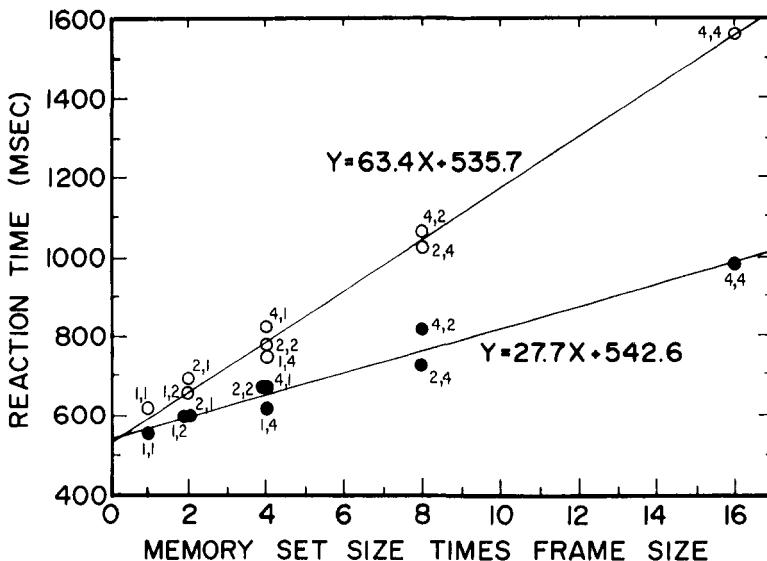


Figure 16. Data from Briggs and Johnsen (1973): reaction time as a function of memory-set size times frame size for negative trials (open circles) and positive trials (closed circles). The first number adjacent to each point designates the memory-set size; the second designates the frame size. The best fitting lines are depicted.

time to switch from one memory-set item to the next has a distribution with a mean of  $\mu$  msec.

Within this general framework, a number of models may be distinguished, depending on the order in which comparisons are carried out and the times at which matching decisions are made. In none of those models, however, is the possibility considered that a single matching decision is made at the conclusion of all comparisons. Such an exhaustive model has been used successfully by Sternberg (1966) and others in simpler situations, but it requires negative and positive functions to have the same slopes. Since the present data have slope ratios close to 2:1, models of this sort may be ruled out in advance. The models to be considered, then, are as follows.

Model 1: The subject first chooses a memory-set item and then compares it to each frame item in turn (in some order). Then the subject chooses another memory-set item and continues. There are two versions of this model.

Model 1a: A matching decision is made after every comparison.

Model 1b: A matching decision is made each time comparisons have cycled through the frame.

Model 2: The subject first chooses a frame item and then compares it against each memory-set item in turn. The subject then chooses another frame item and continues. There are two versions of this model.

Model 2a: A matching decision is made after every comparison.

Model 2b: A matching decision is made each time comparisons have cycled completely through the memory set.

The quantitative predictions for each of these models are derived in Appendix F. Parameters were chosen via a computerized grid search so as to minimize the sum of squared deviations of the data from the predictions. Separate sets of parameters were fit to our experiment and to Briggs and Johnsen's experiment. The parameter estimates for each model for each experiment are shown in Table 1. Also shown in each

case is the square root of the average squared difference between observed and predicted reaction times. Note that in each model, certain combinations of parameters are not separable, and in these cases only their sum is estimated.

Table 1 shows that the best fitting model is Model 1a for both our data and Briggs and Johnsen's data. The predictions of Model 1a have already been presented in Figures 10 and 11 for the two sets of data. This model provides a reasonably good description of the data. It assumes that extra time is utilized for switching from one memory-set item to another, and therefore, it predicts the finding that reaction time is slower when  $M$  is larger, for the cases in which  $M \times F$  is constant.

There are, of course, some mispredictions seen in Figure 10. The major discrepancy occurs at the  $M = 1, F = 1$  negative point; in fact, about one third of the total sum of squared deviations for Model 1a arises from this point alone. Partly for the same reason, the predictions for the  $M = 1$  positive and negative functions of frame size are not as accurate as would be desirable. That is, as seen in Figure 6, the slope of the  $M = 1$  function is quite small, a finding not predicted well by Model 1a. In order to find a model that could reduce the discrepancies, we considered a model in which the search order was entirely random through all possible comparisons. This model we regard as implausible on several grounds, but it does provide a somewhat better fit to the data than even Model 1a. The predictions of this model and a full discussion of it are included in Appendix F.

In the remaining sections of this paper we shall assume that Model 1a gives an accurate description of controlled search in our experimental situation. We are assuming, therefore, that controlled search in our paradigm is a serial, terminating, search process in which the characters in each frame are searched for each memory-set item in turn, and in which each switch to a new memory-set item requires time beyond that used in the basic comparison process.

*Models for the variance of the reaction times.* If a serial comparison process deter-

Table 1

*Models for Mean Reaction Time: Parameter Estimates and Deviations*

Model and study	Parameters				Standard error of prediction
Model 1a	$B_N$	$B_P$	$c + d + \delta$	$\mu$	
SS	397	347	42	40	28.6
BJ	522	468	57	27	25.7
Model 1b	$B_N$	$B_P$	$c + \delta$	$\mu + d$	
SS	397	319	39	47	35.7
BJ	522	430	51	41	40.6
Model 2a	$B_N$	$B_P$	$c + d + \mu$	$\delta$	
SS	444	386	50	1	41.0
BJ	556	495	62	1	33.4
Model 2b	$B_N$	$B_P$	$c + \mu$	$d + \delta$	
SS	461	364	47	1	41.8
BJ	572	466	59	1	35.7

Note. SS (Schneider & Shiffrin), refers to the data from the Experiment 2 varied-mapping conditions.

BJ refers to the data from Briggs and Johnsen (1973). Standard error =  $\sqrt{\sum_{i=1}^n (O_i - P_i)^2 / n}$ .

mines the reaction time, and if the comparisons are identically distributed, independent, random variables, then the variances must be a linear function of the number of comparisons. This holds because the variance of a sum or difference of independent random variables is the sum of the individual variances. (The assumption of independence is thus needed to predict linearity for the variances, but it is not needed to predict linearity for the means.)

On a negative trial in the VM conditions, many models assume that the number of comparisons needed is always  $M$  times  $F$ ; this assumption is common to all the models of the preceding section. If the reaction time consists only of a base time and of the various comparisons (which is not true of the preceding models, since they include switches, but which is true of the Sternberg, 1966, model, for example), then the variances of negative responses will be a linear function of  $M$ , of  $F$ , or of  $M$  times  $F$ . Even when switches are included in the models, linear functions are expected when either  $M$  or  $F$  is held constant and the other varies. The data for the negative VM conditions in Figure 9 are not com-

pletely linear, but they fail badly only when the frame size varies for  $M = 4$ . On the whole, despite a certain amount of noise in the data, the variances are approximately a linear function of the load.

The predictions for the variances for positive trials depend on the termination rule assumed by the model. If a model assumes exhaustive search, then all comparisons are completed whether or not a match has been found. In this case, the distribution of variances for positive trials should have exactly the same form as that for negative trials, differing only, perhaps, in the base time. It is clear from Figure 9 that the negative and positive variances for the VM conditions do not simply differ by a constant.

On the other hand, models assuming that search terminates when a match is found predict a quite complicated relationship between load and variance. In addition to a variance component due to the variability associated with the individual comparisons, a terminating model predicts a variance component associated with the randomly distributed stopping point of the search. In fact, if the load is large enough, this second vari-

ance component becomes so large that the positive variance is predicted to become larger than the negative variance, a result that is found in our data.

Let us now fit the variance data with Model 1a. Suppose we denote the variance of the positive base time as  $\sigma_p^2$  and the variance of the negative base time as  $\sigma_n^2$ . Denote the variance of the combined comparison, decision, and visual switching time as  $\sigma_e^2$  and the variance of the switching time from one memory item to the next as  $\sigma_\mu^2$ .

The predictions for the variances, assuming Model 1a to hold and assuming independence of the different reaction time components, are derived in Appendix G. It is shown there that the variance of the negative reaction time is a function of  $\sigma_e^2$ ,  $\sigma_\mu^2$ , and  $\sigma_n^2$ , while the variance of the positive reaction time is a function of  $\sigma_e^2$ ,  $\sigma_\mu^2$ ,  $\sigma_p^2$ ,  $(c + d + \delta)$ , and  $\mu$ . We therefore chose the values for  $(c + d + \delta)$  and  $\mu$  that were estimated from the data for the mean reaction time (42 and 40, respectively—see Table 1). The remaining parameters were then estimated so as to give a best fit to the variance data (as was done for the means). The best estimates are  $\sigma_e^2 = 2,870$ ,  $\hat{\sigma}_\mu^2 = 2,760$ ,  $\hat{\sigma}_p^2 = 0$ ,  $\hat{\sigma}_n^2 = 1,675$ . The predictions have already been given in Figure 12.

Figure 12 makes it clear that Model 1a provides a reasonably good fit to the variances as well as the means of the reaction times from the VM conditions of Experiment 2. Note that the variances for positive responses are predicted to be higher than those for negative responses at the high values of  $M$  and  $F$ , which is in close agreement with the data. These predictions for the variances lend further support to our model for controlled search.

### *3. Evaluation of the Model for Reaction Times*

In summary, then, we are proposing that controlled search in the paradigm of Experiment 2 is a serial, self-terminating scanning process in which all frame items are compared to each memory-set item before the

next memory-set item is compared in turn, and in which each recruitment of the next memory-set item takes a certain amount of time. This model does an adequate job of fitting the means and the variances of the reaction time data.

Is such a serial, terminating model reasonable? Serial search models are quite prevalent in the literature (see Sternberg, 1975, for a recent review), but almost all of the extant models assume exhaustive search, because the data usually are in the form of parallel negative and positive set-size functions. Exhaustive models predict parallel functions because only one matching decision is made at the conclusion of all comparisons. Thus, the presence or absence of a target does not affect the search, but only (perhaps) the base time. The great majority of such exhaustive-search findings occur in paradigms in which either memory-set size or frame size is equal to 1.

While our data for  $M > 1$ ,  $F > 1$  clearly show the negative slopes to be greater than the positive slopes, the facts are less clear for  $M = 1$  and  $F = 1$ . For  $M = 1$ , the slopes are close to zero, a finding to be discussed further in Section III.F. and Appendix J but one which makes it impossible to discuss ratios of slopes. For  $F = 1$ , the ratio of negative to positive slopes, averaged across the four subjects, is 1.43 (and ranges from 1.19 to 1.65). It may be that the subjects are mixing strategies when  $F = 1$ , sometimes using exhaustive search and sometimes using terminating search. A mixture of this sort would reduce the slope ratio.

Sternberg (1966) has suggested that exhaustive search is utilized, because the time for matching decisions (our parameter  $d$ ) is substantial relative to the comparison time (our parameter  $c$ ). If so, the subject will gain considerable time on negative trials by waiting until all comparisons are completed before making a single decision, rather than making a decision after every comparison. The gain on negative trials makes up for any loss in time that may occur on positive trials using the same exhaustive strategy. However, this argument applies equally well to Experiment 2 and to Briggs and Johnsen's study;

why then do these data indicate terminating search?

One possible reason for the adoption of a terminating strategy may be found in the very long time needed to complete search in difficult conditions such as  $M = 4$ ,  $F = 4$ . Since reaction time may average 1.5 sec or more in these cases, the subjects (who are given instructions to respond as quickly as possible) may be unwilling to wait this long on every trial before responding. By using a terminating strategy, the subject will produce reaction times on at least some (positive) trials that are in the subjectively fast region around .5 sec. Such a set to terminate in the difficult conditions may carry over to the other, easier, conditions. By using termination, the subjects may well produce slower average response times, but they may not realize this fact.

The same considerations are accentuated during training in the multiple-frame paradigm, for reaction times would often be very long if the subject waited till the end of a trial to respond. The consequent history of reinforcement for termination may lead the subject to continue using a terminating search strategy during a subsequent series in the single-frame paradigm.

One implication of terminating-search strategies is that the slope of the reaction time functions, per comparison, may reflect both comparison and decision times; in exhaustive search, the slope would reflect comparison time only. Thus, a greater slope might be expected in conditions producing terminating search. The usual estimates of  $c$  are about 38 msec, (e.g., Sternberg, 1966), whereas the estimates of  $c + d$  are 42 msec for our study and 57 msec for Briggs and Johnsen's study (see Table 1).

A few previous studies have varied both memory-set size and frame size. Sternberg (Note 2) and Nickerson (1966) carried out similar studies, and their results differed somewhat from those we have presented. Sternberg (Note 2) suggested that a variant of the model we have designated Model 2b best fit his and Nickerson's results. It is not at all obvious that one of the other search models would not fit these previous data

(especially Nickerson's) as well as or better than Model 2b. We will not fit our models to these data, however, because the error rates of their subjects were allowed to become much too large in the difficult conditions. For example, Nickerson reported an error rate of 37% when  $M = 4$  and  $F = 4$ . The strategy giving rise to these errors probably cannot be used when subjects are forced to maintain a low error rate. Furthermore, speed-accuracy tradeoffs could have been affecting the reaction times for the correct responses.

In our own study, we found that several training sessions and many repetitions of the accuracy instructions were required before the subjects reduced their error rates to acceptable levels when  $F = 4$  and  $M = 4$ . It is very likely that subjects given instructions to speed their responses have considerable reluctance to wait as long as 1.5 sec before responding. If error rates are high, therefore, we cannot expect our model to apply in an unmodified form.

In general, we do not assume complete uniformity of search mechanisms across experiments. In fact, we wish to argue that controlled search is modifiable by the subject and does alter with changes in paradigms. It seems clear from prior studies that exhaustive search is most often utilized when  $F = 1$  or  $M = 1$ . When  $F \geq 2$  and  $M \geq 2$ , the type of search adopted may depend on the relative importance given to responding quickly or accurately. Of course, other paradigmatic differences between the various experiments that have been discussed could also be responsible for the apparently differing search schemes adopted. Some controlled strategies that arise in different paradigms will be discussed in the literature review section (III.G.).

#### *4. Models for VM Accuracy Data From the Multiple-Frame Task*

We shall now attempt to predict the VM accuracy results of Experiment 1 using the search model and parameters that were fit to the VM reaction time results of Experiment 2. Consider a model for Experiment 1 with the following assumptions:

1. Search is a serial process, obeying the same rules as the search used in the VM conditions of Experiment 2.

2. The search time available for each frame,  $e$ , is equal to the frame time,  $f$ , minus some general adjustment time,  $B_A$ . At the end of each frame, a new search begins through the next frame.

3. An error (a miss) is made if the target is not located during the available search time.

It would seem as though the model could now be applied directly to fit the Experiment 1 results, but unfortunately we did not collect the entire distributions of reaction times in each condition, only the means and the variances. Hence, the shape of the underlying comparison-time and switching-time distributions cannot be culled from the data but must be assumed. Note that the accuracy predictions (unlike the predictions for the means and variances of the reaction times) depend on the exact shapes, not just on the means and variances of the underlying distributions. To see why this is true, note that according to the model, misses in the accuracy task are caused largely by the occurrence of one or more exceptionally long comparison times during the search of the target frame.

For a variety of reasons, we therefore assumed that the comparison times and switching times were both described by gamma distributions. We assumed further that the means and variances of these gamma distributions were equal to the estimates already derived from the reaction time data. This assumption completely specified the distributions, with no parameters left to be estimated. The mathematical derivations and details are given in Appendix H.

We first assumed that the parameter  $B_A$  was zero, that is, that the available search time was equal to the frame time. No free parameters remained in the model, and the predictions were derived for the data of Experiment 1. These predictions were not very close to the data, primarily because performance was predicted to be much too high for the  $M = 1$  conditions. We therefore let

the parameter  $B_A$  vary and found that a best fit of the data occurred (according to a minimum chi-square criterion) when  $B_A = 45$ .

The resulting predictions for the accuracy data have already been presented in Figure 13. One additional version of the model was also fitted to the data: The variances of the underlying comparison-time and switching-time distributions were chosen so as to best fit the data (rather than chosen to match the values estimated from the reaction time data). The estimated parameter values were  $\bar{B}_A = 45$  msec;  $\sigma_o^2 = 4,200$  msec $^2$ ;  $\sigma_\mu^2 = 11,428$  msec $^2$ . The predictions of this version of the model have been presented in Figure 14.

The predictions in Figures 13 and 14 certainly lend support to the hypothesis that the same search mechanism underlies performance in the accuracy task of Experiment 1 and the reaction time task of Experiment 2. These findings therefore suggest that the limitations on the rate of short-term search are the cause of divided-attention deficits.

It is somewhat difficult to evaluate the predictions of the model for the accuracy data. Considering the different nature of the two studies and the fact that the predictions (in Figure 13) were derived in an essentially parameter-free fashion, the model seems to be capturing the main characteristics of the data. On the other hand, there are some marked quantitative deviations. Let us consider these in turn.

First, the data show almost no performance increment when frame time is raised from 600 to 800 msec (in the  $M = 4$ ,  $F = 2$  and the  $M = 4$ ,  $F = 4$  conditions). This finding would be difficult to predict for any model. In fact, we informally ran several sessions in the  $M = 4$ ,  $F = 4$  condition with frame time increased to 2,000 msec, without substantially improving performance. We suspect, therefore, that the subject occasionally forgets one (or more) of the members of the memory set when  $M = 4$ . Probably this forgetting occurs at the start of the series of frames (since rehearsal during the frame presentations would probably prevent forgetting later). The forgetting hypothesis could explain this apparently anomalous result, and it coincides

with subject reports. In retrospect, we wish we had collected posttrial recalls of the memory set, a procedure which might either have verified our hypothesis or forced the subjects to better memorize the memory set. The effect of forgetting is, of course, to flatten the observed psychometric function (of frame time) and thereby cause the model to require very high variance estimates for the underlying distributions (since high variances tend to flatten the predicted functions). Thus, the forgetting hypothesis may also explain why the fit in Figure 14 is better than that in Figure 13.

Note finally that forgetting is not the only hypothesis that could explain the lowered performance at high loads and long frame times. Fatigue is another possibility, since our subjects reported that the high-load conditions were extremely taxing. Each trial at an 800-msec frame time took 16 sec to complete, and it may have been too much to ask of the subjects that they maintain high efficiency during the 320 comparisons that may have been required on a 16-sec trial. In any event, the effects of fatigue could be expected to manifest themselves in a manner similar to that expected for forgetting, so that the arguments of the preceding paragraph would apply.

Second, the observed psychometric functions for the  $M = 4, F = 2$  and the  $M = 4, F = 4$  conditions are closer together than those predicted, even when the variances are estimated separately. At least part of this discrepancy may be due to forgetting, as discussed in the preceding paragraph.

In light of the quantitative deviations observed, it cannot be argued that our model is an accurate and complete description of the processing occurring in the multiple-frame task. Nevertheless, the results are highly encouraging. Further research to clear up some of these problems is now under way, but in advance of such research we feel that the present results still provide good support for the hypothesis that a common, serial, terminating search mechanism underlies performance in the VM conditions of Experiments 1 and 2.<sup>4</sup>

To summarize, the quantitative models

that have been applied to the VM conditions suggest a number of conclusions. First, the search strategy in the reaction time study is serial and terminating. The comparisons cycle through the set of frame items for a given memory item before switching to the next memory item. Each switch requires a certain amount of time apart from the basic comparison and decision time. Second, the results suggest that the search strategy used in the reaction time study is also used in the accuracy study. Thus, the limitations on

---

<sup>4</sup> It should be noted that the gamma distributions that we have estimated for the comparison time and the switching time have a somewhat unusual shape. In particular, the highest point in the density is at  $t = 0$ , and the density decreases monotonically as  $t$  increases. Although such a shape appears at first glance to be an unusual one for a distribution for a single comparison time or a single switching time, the shape is forced upon us by the fact that the variance is so large relative to the mean. In the fit to the means and variances of the reaction time data of Experiment 3 (sections III.D and III.E.), the variance estimates were larger than the estimates of the mean squared. If a unimodal reaction time distribution has a mode near the mean, then the tail above the mode cannot be far extended relative to the tail below the mode (or else the mean would not be close to the mode). Under these conditions, where the distribution is relatively symmetrical, the variance is bounded above. For example, the variance of a symmetric distribution with mean  $x$  is maximized if the density is  $\frac{1}{2}$  at  $t = 0$  and  $\frac{1}{2}$  at  $t = 2x$ . Such a distribution has a variance of  $x^2$ . Thus, for the variance to be larger than the mean squared, the mode must be smaller than the mean, perhaps considerably so. The estimated gamma distributions for our model have this property, as must any reasonable distribution with a variance larger than the mean squared. Of course, we do not necessarily believe that the density maximum is at  $t = 0$  but only that the mode must be at a small value of  $t$ .

Sternberg (Note 3) estimated the shape of the underlying comparison-time distribution from the first four moments of the observed reaction time distributions in one of his memory-search studies. The estimated shape was a beta, rather than a gamma, but the two distributions are actually very similar: Both densities descend sharply and monotonically from an infinite peak when  $t$  is very small, and both have long tails extending toward large values of  $t$ . Whether such a shape is reasonable for the comparison-time distribution is another matter, but we see no a priori reason for ruling out such a distribution at the present time.

search rate may be inferred to be the cause of missed targets in accuracy experiments. In slightly different words, attentional deficits, such as those seen in Experiment 1, are due to the limitations upon the short-term search process. This model, and applications to studies of attention, will be elaborated in later sections of this paper and in Part II.

### E. Individual Differences in Experiments 1 and 2

The models in the preceding section have all been applied to the data averaged across the four subjects. Such a procedure is never really justifiable and certainly cannot be tolerated unless the pattern of results is much the same for each of the subjects. In this section, therefore, we show that there are considerable differences among subjects, but largely in performance level rather than pattern, so that it is reasonable to suppose that the subjects are all using the same strategies.

Table 2 gives the individual best-fit slopes for each subject for each condition, corresponding to the average VM functions depicted in Figures 5 and 6. For the memory-set-size functions, the slope is given for frame size varying. For the frame-size functions, the slope is given for memory-set size varying. Subject P.D. has much smaller slopes than the others and Subject W.S. has slightly larger slopes. All subjects, however, have

negative slopes roughly twice the positive slopes. Thus, it seems likely that these subjects are utilizing similar search strategies but with differing comparison and switching times.

The approximation to 2:1 slope ratios for each subject holds up best for the largest frame sizes and memory-set sizes. For  $M = 1$  or for  $F = 1$  (especially  $M = 1$ ) the results are much more variable, and the slope ratios deviate considerably from 2:1. It is natural to suspect that when  $M = 1$  or  $F = 1$ , the subjects tend to switch to an alternative search strategy, exhaustive when  $F = 1$  or even parallel when  $M = 1$ , on at least some large proportion of the trials (see Appendix J).

Next, Model 1a was fit to the VM data of each of the four subjects separately. The estimated parameters and the square root of the average squared deviations are given for each subject in Table 3. As might have been expected in view of Table 2, Subject P.D. shows a much smaller estimate for  $c + d + \delta$  than do the others.

Finally, the individual subject data from the VM conditions of Experiment 1 are summarized in Table 4. For each combination of  $M$  and  $F$ , the table gives the hit probability averaged across the three frame times for that combination. Although this description of the data does not depict the shape of the psychometric functions, in fact each subject's data corresponded closely to

Table 2  
*Slopes of the Memory-Set-Size and Frame-Size Functions for Individual Subjects in the Varied-Mapping Conditions of Experiment 2*

Subject	Memory-set size			Frame size		
	1	2	4	1	2	4
<b>Positive trials</b>						
C.B.	-4.8	42.6	81.8	16.2	54.7	101.0
T.D.	38.0	54.9	94.3	45.0	76.2	104.1
W.S.	12.3	34.1	75.7	54.8	72.3	117.4
P.D.	14.1	21.8	48.0	34.4	55.6	71.0
<b>Negative trials</b>						
C.B.	20.1	77.2	214.2	23.2	127.5	226.8
T.D.	26.9	87.4	242.8	65.0	134.7	282.1
W.S.	4.4	64.1	236.5	65.4	145.4	301.5
P.D.	6.1	29.3	83.8	56.8	95.3	136.6

Table 3

*Parameter Estimates and Deviations for Model 1a for Individual Subjects in the Varied-Mapping Conditions of Experiment 2*

Subject	$B_N$	$B_P$	$c + d + \delta$	$\mu$	Standard error of prediction
C.B.	490	447	46	20	32.6
T.D.	343	292	55	30	29.5
W.S.	349	282	48	57	45.5
P.D.	404	366	20	52	34.4

the shapes of the curves in Figure 3, with only the overall level of accuracy varying across subjects. Note that Subject P.D., who had the smallest slopes in Experiment 2, had the highest accuracy level in Experiment 1. Also, Subject W.S., who had the largest slopes in Experiment 2, had the lowest accuracy in Experiment 1. Thus, the pattern of individual differences is consistent across the two experiments.

In summary then, we feel it is reasonable to assume that the subjects were utilizing similar controlled-search strategies, though with rates that varied from one subject to the next. While this conclusion does not justify the fitting of the models to the average data, it does make that procedure somewhat less reprehensible. We have not fit the individual data, primarily because we were worried that too few observations were collected for any one subject in Experiment 1 to support such an undertaking.

#### F. Automatic Detection versus Controlled Search: Evidence From Experiments 1 and 2

In this section we shall begin to discuss the nature of automatic search, and the differences between the two search modes, using comparative evidence from Experiments 1 and 2. One main result from both studies showed that neither reaction time nor accuracy varied as a function of load in the CM conditions. It is natural to interpret this finding as evidence for a parallel and independent search through the visual (and memory) characters for a given frame. An additional question arises for Experiment 1, however: Does automatic search occur in parallel for the visual characters in successive frames?

The answer to this basic question lies in the resolution of an apparently discrepant finding between Experiments 1 and 2. In

Table 4

*Mean Hit Probabilities for Each Subject for Each Varied-Mapping Condition in Experiment 1 (Averaged Across Frame Times)*

Subject	Condition						Row mean
	$M = 1$	$M = 1$	$M = 1$	$M = 4$	$M = 4$	$M = 4$	
	$F = 1$	$F = 2$	$F = 4$	$F = 1$	$F = 2$	$F = 4$	
	$f = 160$	$f = 160$	$f = 253$	$f = 400$	$f = 600$	$f = 600$	
C.B.	92	81	80	76	76	62	78
T.D.	87	74	81	81	76	68	78
W.S.	90	90	76	69	66	51	74
P.D.	99	91	89	92	89	78	90
Column mean	92	84	82	80	77	65	80

Note.  $M$  = memory-set size;  $F$  = frame size; and  $f$  = mean frame time (in msec).

Experiment 1, the easiest VM condition resulted in much lower performance (accuracy) than any of the CM conditions. In Experiment 2, however, the means and variances of reaction times for the easiest VM conditions were equal to or lower than those for the CM conditions.

It may be that subjects adopt a parallel search strategy when  $M = 1$  in the VM conditions of Experiment 2. This hypothesis could explain why the reaction times are nearly equal for the CM conditions and the VM conditions when  $M = 1$ . However, this hypothesis only accentuates the discrepancy between Experiments 1 and 2, since an extrapolation of the data when  $M > 1$  suggests that the reaction times for the  $M = 1, F = 1$ , VM condition would have been even faster than those observed if the subject had utilized serial rather than parallel search. This reasoning is confirmed by the predictions of the model shown in Figure 10. It may be seen there that the serial search model predicts reaction times even faster than those observed when  $M = 1$  and  $F = 1$ . In short, we are asking why the comparison between the easiest VM conditions and the CM conditions goes in opposite directions for the two studies.

There are many possible answers to this question (some quite uninteresting—for example, the base reaction times,  $B_P$  and  $B_N$ , might be supposed to differ in the CM and VM conditions). The simplest, most elegant, and most plausible explanation is based on the presumption that automatic detection operates in parallel and independently across frames as well as within frames. If so, the time for an automatic response to occur would be independent of the frame time.

To make this argument clear, consider an extreme example. Suppose that automatic responses were made with high accuracy but required 10 minutes to be emitted. Then the response time in the CM conditions of Experiment 2 would average 10 minutes, much longer than the 1-sec responses for the VM conditions. But what would happen in the multiple-frame task? If an automatic-detection response occurs in parallel across frames, then whenever a target appears, it will at once initiate an accurate response system

(however long the response system takes to run to completion). Thus, all 20 frames could be presented very quickly, within several hundred milliseconds, say, and an accurate response would be given 10 minutes after the occurrence of the target frame. This example makes it clear that the only limitation on frame time in the CM conditions, if automatic detection occurs in parallel across frames, would be the physical limits of the system to encode briefly presented material. When the frame time drops to the point where the input stimuli are incompletely or inaccurately processed, then the limits will have been reached.

To summarize and conclude, we feel that the CM conditions in Experiment 1 produce higher performance than even the easiest VM conditions because the automatic detection used in the CM conditions is parallel within and between frames. Thus, the frame time limits automatic-detection performance only because of peripheral physical factors like lateral and temporal masking. On the other hand, the response-time data of Experiment 2 suggest that the automatic-detection system may actually be somewhat slower than a very easy controlled search (i.e.,  $M = 1, F = 1$ ) under the conditions of our study.

We now turn to another very basic question. Why are those frame locations that are filled with masks not included in VM search? Clearly, such locations do not require appreciable processing time, since the number of masks increases as the load and the mean reaction time decrease. At first glance, this question appears foolish; it appears that one might as well ask why the screen, the fixation dot, or the response box do not require processing. In fact, these questions also must be answered.

Historically, these questions have been treated seriously by many authors. Most have argued that stimuli may be selected "preattentively" on the basis of simple physical features like color, size, or location (e.g., see Broadbent, 1971; Neisser, 1967). Thus, location would provide a basis for excluding from the search all features of the visual field except the four frame elements. The

masks would be excluded from search on the basis of a gross difference in shape or form. (See Part II for a fuller discussion of these issues.)

On the basis of our findings that training mode (VM or CM) determines the ability to use automatic detection, we would like to propose an alternative hypothesis. Suppose that an automatic process develops that directs controlled search to consistently relevant locations and consistently relevant types of stimuli (characters). In our studies, masks are always irrelevant, while characters are always relevant in the sense that they always contain the target if one is present. Such a hypothesis suggests that a change in our procedure might reduce the ability to exclude masks from the controlled-search process. For example, suppose that trials with masks making up the memory set alternate with trials using characters to make up the memory set. Then masks and characters would be alternately relevant and an automatic process presumably could not develop. In such a case, it might be necessary on every trial to include both mask and character locations in the search. Results are not presently available to confirm the validity of this hypothesis.

Let us finally turn to a consideration of the relation of the flat frame-size function when  $M = 1$  in the VM conditions to the flat set-size functions in all the CM conditions. At several previous points we have suggested that the flat,  $M = 1$ , VM function indicates some type of parallel search. Can this parallel search be equated with the one seen in the CM conditions? There are several reasons why we feel that these processes differ. First, in the multiple-frame task, performance in the  $M = 1$ , VM conditions differs with frame size and also is worse than that in the CM conditions. Second, the constancy of the  $M = 1$ , VM conditions occurs only at the cost of additional errors at the higher frame sizes. This fact, and a more extensive discussion, are provided in Appendix J. Finally, subjects report considerably greater attention demands (i.e., subjective task difficulty and effort) for the  $M = 1$ , VM conditions than for any of the

CM conditions. For all of these reasons we prefer the hypothesis that controlled search requiring attention occurs in the VM conditions even when  $M = 1$ , and even though the search when  $M = 1$  might be parallel in nature.

We propose that the parallel process that occurs in the CM conditions, termed *automatic detection*, differs qualitatively from those search mechanisms that occur in any of the VM conditions. The remaining studies presented in this paper, and those presented in Part II further demonstrate the validity of this contention.

#### G. A Selective Review of Reaction Time Studies of Search

In this section, we will briefly review studies that use reaction time to demonstrate automatic detection and controlled search. The role of consistent training will be carefully considered. We will show how the contrast between the two processing modes helps us to understand a number of puzzling results in the literature.

We propose that subjects can search in one of two qualitatively different modes: Automatic detection develops if the memory-set items are consistently mapped to positive responses across trials; controlled search is utilized when the mapping is inconsistent across trials. We suppose automatic detection reflects the occurrence of an attention-directing response that results from the presentation of a consistently mapped target. Attention will be directed automatically to such a target so that no serial search through the memory-set items or display items is needed. In such a case, flat reaction time functions are expected. On the other hand, when no item produces an attention-directing response, then each input must be compared in turn to each memory-set item; in such a case, linear set-size functions are expected (with slopes of approximately 30–80 msec per comparison, depending on the study). In many studies, consistent-mapping procedures are utilized, but the amount of training may be too small for automatic detection to develop fully. In such studies, a mixture

of the two processing modes is to be expected, and the set-size functions may take on a variety of forms intermediate between the two extremes. Let us consider some of these possible intermediate forms and their causes.

One possible set-size relationship would arise if automatic detection were utilized on some proportion of the trials or for some subset of the memory set. In this case, one would expect linear results with reduced slopes. Another possible outcome is based on the assumption that automatic detection will gradually increase in speed until it is as fast as even the fastest controlled search ( $M = 1, F = 1$ ). Before this point is reached, the subject may carry out both processes in parallel, responding in accord with the process that ends first. In such a case, one might expect a latency function which would be linear and have the usual slope for small memory-set sizes or frame sizes but which would tend to become flatter as the memory-set sizes or frame sizes increased (since automatic detection is often completed first under conditions in which controlled search, on the average, proceeds slowly). A curvilinear, or two-limbed, form for the set-size function can also be predicted at intermediate levels of practice if conditions are blocked or if the visual and memory-set sizes are known in advance of a given trial. Under such conditions, the subject might decide in advance of the trial to adopt a controlled search when the load is small but to utilize only automatic detection when the load is large. Thus, at intermediate levels of practice, we might expect either linear functions with reduced slope, two-limbed functions (a limb for small set sizes with a normal slope, followed by a limb at large set sizes with a reduced slope), or curvilinear functions (negatively accelerated).

Let us now turn to the data. Consider first situations in which  $F = 1$  and  $M$  varies, situations usually termed memory search. Sternberg (1969) reported linear functions, parallel for negative and positive tests, with slopes of 38 msec/item, results that have been replicated many times since for VM conditions. (See Sternberg, 1975, for a recent re-

view; Kristofferson, 1972a, in particular, showed that even enormous amounts of VM training did not change the basic linearity of the set-size function.) However, almost identical results were found in a second experiment run under CM conditions (Sternberg, 1966). It seems likely that indications of automatic detection did not appear in Sternberg's CM conditions for three reasons. First, very low levels of practice were used. Each subject was given, for each memory-set size, 60 practice trials and 120 test trials. These trials for each memory-set size were run in a single block. Second, from block to block (as memory-set size changed), the mapping of stimuli to responses was altered. The positive items on one block became negative on the next (and vice versa to a considerable degree). Third, Sternberg stressed accuracy in his studies; this stress on accuracy may have inhibited subjects from relying on an automatic-detection mechanism that would not be fully developed at low practice levels and, hence, that would be error prone. Any or all of these factors could have worked against the development of an appreciable degree of automatic detection.

In fact, in almost all other memory-scanning studies using CM designs, reaction time is not a linear function of memory-set size. Typically, in CM designs, reaction time is close to a linear function of  $\log_2 M$  (e.g., Briggs, 1974; Kristofferson, 1972a; Ross, 1970, Simpson, 1972; Swanson & Briggs, 1969). Simpson (1972) reviewed some of the results and contributed a number of new findings. His first study used a CM design with letter stimuli. Each condition utilized 160 trials on each of 3 days. These results showed reaction times to be a linear function of  $\log_2 M$ . On each day, reaction time decreased but the functions remained parallel. This result is very similar to that of Kristofferson (1972b), who gave her subjects much greater amounts of practice. Simpson's second experiment used a CM design with digits as targets and letters as distractors; 192 trials were given per condition. Again the results showed reaction time to be a linear function of  $\log_2 M$ . The third experiment used all letters in a CM task and

showed that on Trials 65-128, the reaction time was a linear function of  $M$  but that on trials 129-192, the reaction time was linear with  $\log_2 M$ . In all three studies, the positive and negative functions were parallel. These results are very similar to the CM functions of Briggs and Johnsen (1973), who also found that reaction time during several sessions of CM practice moved from a linear function of  $M$  to a linear function of  $\log_2 M$ .

These curvilinear set-size functions (except for the study by Kristofferson, 1972b, to be discussed below) are very likely the result of intermediate levels of practice causing a mixture of automatic detection and controlled search.

Ellis and Chase (1971) presented an excellent demonstration of the mixing of automatic detection and controlled search. Their subjects performed a memory-search task ( $F = 1$ ) under VM conditions. The novel feature of this study was the manipulation of the size of the distractors: A given group of subjects had a mixture of same-size distractors and smaller-size distractors. For different subjects, the smaller size was five sixths, two thirds, or one half the size of the large-sized distractors (the targets were always large size). Hence, there was a consistent mapping of small distractors to negative responses. Thus we might expect linear set-size functions for same-size distractors, and set-size functions with lowered slopes for smaller-sized distractors. These results would support a model in which automatic detection on the basis of size is operating simultaneously with serial search of the memory set. The two processes occur in parallel, and the response is based on the first process to finish.

Assuming that smaller size-differences require longer automatic detections, then one would expect the set-size functions for small-sized distractors to mimic the function for same-sized distractors at small set sizes (when serial search is almost always faster), but to fall below the function for same-sized distractors at the larger set sizes (when the automatic detection would tend to finish first). Furthermore, the set size at which the small-size distractor function should

begin to fall below the same-size distractor function should be smaller as the size discrepancy becomes larger.

All of these predictions were confirmed in the data. It is possible, of course, that if training in this task were continued long enough, the detection process based on size would become so efficient that it would be faster than even the easiest serial search. In such an event, all the set-size functions with small-sized distractors would become flat.

The possibility that automatic detection and controlled search can sometimes operate in parallel provides one explanation for the finding of Kristofferson (1972b). She used a CM, memory-search ( $F = 1$ ) design very similar to our CM conditions in Experiment 2, except that the stimuli for some subjects were all digits and for other subjects all faces. There were 10 stimuli, four of which composed the memory sets, and six of which were always distractors. The subjects were given large amounts of practice, and the set-size functions remained curvilinear without tending to flatten from Days 7-36 (although overall performance level continued to improve).

We ask why these set-size functions remained curvilinear and did not approximate zero slope (as did ours in Experiment 2). It is doubtful that the reason lies in the fact that we used digits and letters for our sets, whereas Kristofferson used either all digits or all faces. We reject this possibility because in Experiment 3 of Part II (to be reported later), no memory-set size differences occurred in a CM procedure using a memory ensemble of four letters and a distractor set of four visually matched letters. (Experiment 3 of Part II used the multiple-frame procedure, but we doubt this fact should change our conclusion.)

Instead, the reason may be based on the use in parallel of serial search and automatic detection. Kristofferson's subjects may have used both and responded when the first was finished (as in the Ellis & Chase, 1971, study). The serial component would tend to beat the automatic component more often at the smaller set sizes, thereby producing

the observed findings. Of course, this explanation requires that the VM functions become speeded with extended practice at about the same rate as the CM functions, but this is just what Kristofferson found (1972a, 1972b). Perhaps the overall improvement in both studies was due to a decrease in motor response times. (Kristofferson offers a somewhat different explanation of her finding: She suggests that a single unitary search is organized in a hierarchical series of feature tests, a model similar to that of Hick, 1952, discussed below.)

Why do the subjects in our studies not show findings similar to those of Kristofferson? Our subjects were run in many conditions with higher loads (up to 16), so the serial component, if it were used, would seldom help facilitate search. Furthermore, our subjects practiced previously in the multiple-frame procedure in which any serial component would tend to be useless at CM frame times. Finally, our subjects had training in VM paradigms, in which effortful attention and controlled search was required, and they tended to relax in the subjectively much easier CM conditions, so they may not have attempted to speed their processing with a controlled serial search. From their point of view, a great deal of effort would be expanded for very little actual gain.

One alternative to the theory that practice leads to flat set-size functions through the development of automatic detection is the theory that flat functions are due to a categorical difference between targets and distractors. Some evidence against this alternative theory can be found in Simpson's (1972) second experiment, in which the use of digits as memory-set items and letters as distractor-set items produced results similar to those obtained when only letters were used for both sets. Much more persuasive evidence will be shown in Experiment 3 of Part II.

Alternative models for the  $\log_2$  functions have been proposed since Hick (1952) first suggested that a discrimination process which proceeds according to a series of binary subdecisions could give rise to such a result. Many other workers have since suggested models of this general nature, includ-

ing those models discussed above. However, we have demonstrated through the studies reported in this paper that linear functions with 40-msec slopes occur when the mappings vary from trial to trial and that flat functions occur when much practice is given with consistent mappings. A strong case has been made (which will be made stronger in later studies) for the presence of two different search modes in these two cases. Thus, although we cannot rule out the possibility that there is a third, hierarchical, search process that occurs at intermediate levels of practice in CM conditions, it seems most parsimonious to assume that the curvilinear functions arising at intermediate practice levels are the results of mixtures of the processes occurring at the two extremes of practice.

Consider next, conditions in which  $M = 1$  and  $F$  varies. Such paradigms are usually termed "visual search." Atkinson, Holmgren, and Juola (1969) used a VM paradigm and found reaction time to be a linear function of  $F$ , with parallel negative and positive functions.

The results from CM, visual-search paradigms are quite different. Egeth, Jonides, and Wall (1972) showed that a "4" as the memory-set item in a distractor set of "C"s gave reaction times independent of frame size. Jonides and Gleitman (1972) extended this work by showing that the category of a target could determine the shape of the reaction time function. An oval shape as the memory-set item in a distractor set of letters gave rise to a flat set-size function for subjects who thought the oval shape was a number (zero) but gave a linear set-size function with a slope of 26 msec for subjects who thought the oval shape was a letter. These results reversed if the distractor stimuli were digits rather than letters. For a given subject, the category to which the oval shape belonged and the type of distractors used were fixed over the experiment. Thus the category of the "O" was induced both by the other possible memory-set items and by verbally stating, prior to a trial using the oval shape, either zero or oh.

The Jonides and Gleitman findings point to the importance of categorical differences in

producing flat functions. All conditions were run with CM procedures on different groups of subjects, for 126 trials each. Thus, subjects had identical amounts of consistent practice with the oval shape in each of the conditions of interest, yet very different slopes resulted in the two cases.

We suggest that an automatic-attention response can be attached not only to the "name" of a stimulus but also to its category if the category is consistently mapped to responses. Furthermore, we propose that it takes less training to develop automatic detection for a single category, already known, than for the individual stimulus names making up the category. With these assumptions, the Jonides and Gleitman results can be accommodated in our framework. When an oval shape is embedded in other memory-set items of the same category as the distractors, then no consistent mapping of the category to responses is available. Since the individual memory-set characters cannot develop automatic detection in just 118 trials, a controlled-search strategy is used, and linear set-size functions with substantial slopes occur. On the other hand, when a consistent categorical difference between memory-set items and distractors is available, then the category feature abstracted from any memory-set item may develop an automatic-attention response; since only this single feature-response relationship need be learned, it develops quickly (within the 118 trials), and flat functions result. The model we have suggested emphasizes the role of consistent practice (whether or not a categorization is available) and is in contrast to the suggestions of Jonides and Gleitman, which relied solely upon categorization (see also Experiment 3 of Part II).

Additional results bearing on practice and categorization effects can be found in the work of Egeth, Atkinson, Gilmore, and Marcus (1973). Again, in their studies,  $M = 1$  and  $F$  varied. In their Experiment 1 they utilized CM conditions: In some, an "A" was the memory-set item among distractors of letters or numbers, and in others, a "4" was the memory-set item among distractors of letters or numbers. All conditions were run in

different groups, and there were 180 training trials in each condition. The findings differed considerably from those of Jonides and Gleitman (1972). The visual set-size functions were all relatively flat with slopes ranging from 5 to 10 msec, but categorization had almost no effect. Distractor sets of numbers resulted in differing slopes than distractor sets of letters, whatever the memory-set item, but the categorical relationship between the memory set and the distractor set made no difference.

This result from Egeth et al. (1973) may have been caused in part by the fact that there was just one target for a given subject. Thus, the subject had a choice of perfectly consistent response rules—letter name or category name. In such a case, automatic detection could conceivably develop faster for the better-learned feature (i.e., letter or digit name) than for the less well-learned feature (i.e., category name). Thus, automatic detection would tend to be based on the name, rather than the category, of the memory-set item. A different result was obtained by Jonides and Gleitman (1972) because the memory-set items varied over trials for a given subject; thus the category name was mapped to a positive response much more often than was any single item.

An alternative or supplementary explanation for the lack of a categorical effect in the Egeth et al. study is based on the physical similarity of the memory-set items to the distractor set. Both "A" and "4" tend to be similar to letter backgrounds (being constructed from straight lines) and dissimilar to digit backgrounds. This average physical difference could have been used as an automatic-detection cue taking priority over either letter or category name. Possibly, the subject might select the subset of straight-line characters via automatic detection and then check them via controlled search. This hypothesis is supported by a visual-search study by Ingling (1972) that is similar to the studies of Neisser discussed earlier. Ingling varied both the categorical relationship and the physical relationship of the memory-set item to the distractors in a CM design. The results showed both effects:

Search was faster when a categorical difference existed and it was also faster when a physical difference existed (the first effect occurred when physical differences were controlled). Note that these findings are perfectly compatible with our approach. Either, or both, physical or category differences could be used as a basis for automatic detection under CM training conditions.

The final reaction time studies to be reviewed at this time are those in which  $M > 1$  and  $F > 1$ . These hybrids of memory search and visual search are closest to Experiment 2 in design. We have already discussed the studies of Briggs and Johnsen (1973), Nickerson (1966), and Sternberg (Note 2), but only the first of these utilized a set of CM conditions. The remaining three hybrid studies of this type are those of Egeth et al. (1972, Experiments 3 and 4; 1973, Experiment 2) and Sternberg and Scarborough (Note 1).

In the Egeth et al. study (1972; Experiments 3 and 4), all 10 digits made up the memory set, and the distractor items were letters. In these studies,  $F$  varied from 1 to 6. Over 4 days of practice, the frame size functions consistently flattened until the slopes were essentially zero. In the later Egeth et al. (1973) study, the memory set again consisted of all the digits, and the distractor items were again letters, but there were three conditions: (a) Frame size ranged from 2 to 8 with frame time equal to 200 msec; (b) frame size ranged from 2 to 8 with frame time equal to 500 msec; and (c) frame size ranged from 2 to 16 with frame time equal to 200 msec. Over 4 days of practice, all slopes again decreased, reaching essentially zero for conditions (a) and (c). With the long frame time (condition b), the slope did not quite reach zero, which led Egeth et al. to suggest that the subjects may have been "rechecking" their answers. Alternatively, the long frame time may have allowed subjects to complete a serial search and may thereby have hindered the development of automatic detection. In any event, the results of condition (c) show that automatic detection can operate efficiently even with simultaneous presentation of as many as 16 items,

and with search being maintained for a memory set consisting of all the digits in a distractor set of letters.

Sternberg and Scarborough (Note 1) used a paradigm reminiscent of our Experiment 1, but with reaction time as a dependent measure. Twenty successive digits were presented, one every 70 msec, in the same spatial location. Thus there were 20 frames with  $F = 1$ , and  $f = 70$  msec. The subject responded upon detecting a member of the memory set. The memory set was constant for a block of trials and consisted of 1, 3, or 5 digits in different blocks of trials. In our terminology, the paradigm utilized consistent mapping within each block but not between blocks. Each block consisted of 90 trials.

The results showed a linear relationship between reaction time and  $M$ , with a slope of 40 msec per item in the memory set. When  $M$  was equal to 5, the subjects could not possibly have searched serially at a rate of 40 msec per item and kept up with the successive frames that arrived every 70 msec. Thus Sternberg and Scarborough suggested that serial search through the memory set occurred for each frame item but that the successive frame items were each compared to the memory set in temporally overlapping searches. That is, multiple simultaneous searches were hypothesized. Our results and theory, however, suggest a possible alternative explanation. Because each block used consistent mapping, the subjects may have been able to use a somewhat error-prone automatic detection to single out (with high probability) the frame containing the target. However, to prevent excessive errors, the subject may then have checked the detected item against the memory set using controlled search. This latter process would give rise to the observed reaction time function with a slope of 40 msec per item. Presumably, if our hypothesis is correct, additional practice in this paradigm under CM conditions (over blocks as well as within blocks) would have eventually resulted in the subject's use of automatic detection only, and hence would have produced flat set-size functions (as our data showed.)

There is an important implication of our work for research into memory and visual

search. If the stimulus-to-response mapping is changed from trial to trial, pure controlled search performance will be observed. If the stimulus-to-response mapping is fixed and the subject is well practiced, automatic-detection performance will be observed. Experiments changing the mapping between blocks will show a mixture of the two processes and thus be difficult to interpret. Therefore, unless the experimenter wishes to investigate the development of automatic detection, the stimulus-to-response relationship should only be manipulated between trials or between subjects.

In summary, this brief review of reaction time experiments suggests the following: Our division of search processes into *controlled search*, operating in VM procedures or at very low levels of practice in CM procedures, and *automatic detection*, learned gradually after practice in CM procedures, can organize a variety of apparently discrepant results into a relatively simple framework. A similar case can be made for experiments in selective attention (that generally use accuracy measures), but any extensive discussion of attentional studies will be deferred until Part II, following the presentation of several studies to explore attentional mechanisms directly.

#### H. An Empirical Organization of Search Results

Quite apart from our theoretical proposal of two distinct detection mechanisms, the results of our experiments provide a means of classifying and organizing a wide variety of search and attention studies. Even were one to take the view that detection is a continuum of processes ranging from slow serial search to fast parallel detection (a view we hope to argue against with subsequent experiments), our studies allow us to predict empirically just what results to expect from a given paradigm.

The dimensions of primary interest are those of *mapping consistency*, *amount of training*, and *load*. With inconsistent mapping, load will have a large effect at all levels of practice. With consistent mapping, load should have a decreasing effect, and perform-

ance should improve, as the amount of training increases. In this classification, categorical differences would primarily affect the relative rate of change that occurs with practice. The validity of such an empirical organization is supported by the reviews in Section III.B. and III.G.

#### IV. Searching for Multiple Targets: Differentiating Automatic Detection from Controlled Search

In the previous sections of this paper we have introduced the basic concepts of controlled and automatic processing modes. The approach was based on the results of Experiments 1 and 2 and on a brief survey of the literature. However, on the basis of our previous studies, one could possibly maintain that only a single search process was operating, one that could be speeded by certain conditions of practice or by categorical differences between targets and distractors. In the following experiments, a stronger case will be established for the existence of two qualitatively distinct search mechanisms.

The vehicle for this demonstration is an experimental paradigm similar to that of Experiment 1. However, in Experiment 1 we utilized either no targets or one target on a given trial. In Experiment 3, to be reported next, we used either zero, one, or two targets per trial. The rationale for this paradigmatic shift lies in the selective-attention literature. Numerous studies have shown that it is more difficult to detect a target presented on one channel when another target is presented simultaneously on another channel (if both channels require responses). This difficulty arises even though the requirement to monitor multiple channels may in itself cause no deficit (in many studies, only the simultaneous presence of two targets causes performance decrements—see Moray, 1975; Sorkin & Pohlman, 1973). Thus, a performance decrement is engendered by multiple simultaneous targets in situations similar to our CM conditions. Therefore, although our CM conditions with single targets show no effect of load, performance drops might appear when multiple targets are presented. Any

such effects could then be compared to those occurring in comparable VM conditions, and differences in the results might serve to distinguish the two types of search.

### A. Experiment 3a: Detection of Multiple Targets When $M = 2$ and $F = 2$

In this study, we used an accuracy paradigm similar to that of Experiment 1, with the exception that zero, one, or two targets could occur on a given trial. The subject responded at the trial's end by indicating the number of targets detected. Two new independent variables were manipulated in this study on those trials containing two targets. First, the temporal relation between the two targets was varied: The number of frames separating the two targets (the *spacing*) was either 0, 1, 2, or 4. Second, the relationship of the two targets to each other was varied: The two targets could be identical to each other (termed II for *identical items*) or could both be different characters from the memory set (termed NI for *nonidentical items*).

#### 1. Method

Experiment 3a utilized a 20-frame procedure as in Experiment 1, except that the subject knew that multiple targets might be present and was required to note these as they occurred and to respond with the number of detected targets at the end of the trial. The subject heard an error tone if the reported number of detected targets did not equal the actual number. In fact, 25% of the trials had no targets, 25% had one target, and 50% had two targets. Memory-set size and frame size were both 2 in all conditions. When two targets were presented, one half of the time they were identical (II), and one half of the time they were different (NI). On  $\frac{1}{4}$  of the trials the two targets were in the same frame (Spacing 0); on  $\frac{1}{4}$  of the trials they were in successive frames (Spacing 1); on  $\frac{1}{4}$  of the trials one frame intervened (Spacing 2); and on  $\frac{1}{4}$  of the trials three frames intervened (Spacing 4). Examples of some of these conditions are given in Figure 17. Regardless of the spacing, two targets were always in different display positions, but subjects were unaware of this constraint. When two targets were presented, the first occurred randomly in any of Frames 4 to 13.

The nature of the mapping (VM or CM) was varied between blocks; all other conditions were

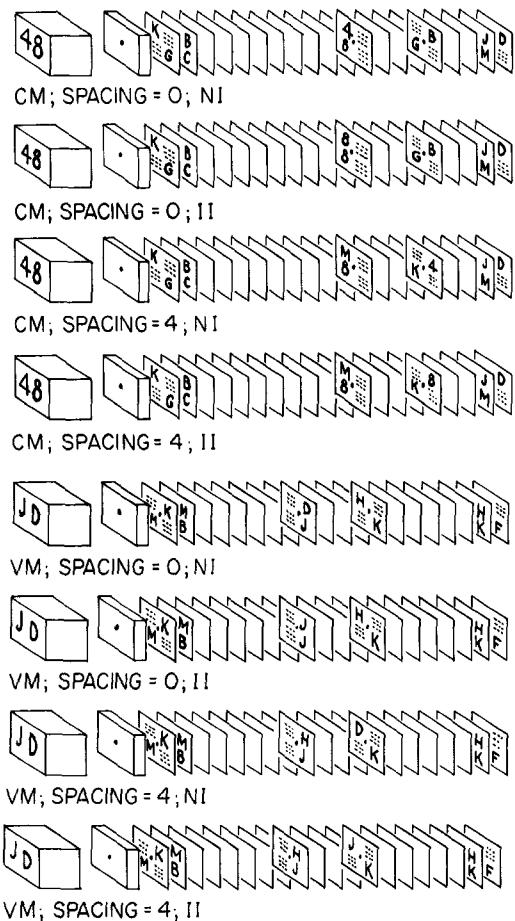


Figure 17. Examples of trials used in Experiment 3a. In each case, a trial is illustrated that has two targets, though other trials sometimes contained just one target, or no target. (CM = consistent mapping; VM = varied mapping; II = identical items; NI = nonidentical items.)

varied within blocks. There were four blocks of 132 test trials in each session; the first block began with 15 practice trials, and the others began with 5 practice trials. A total of 12 VM and 12 CM blocks were run per subject. In VM blocks, frame time was set to 200 msec, and in CM blocks, frame time was set to 80 msec.

#### 2. Results and Discussion

The results averaged across subjects are graphed in Figure 18. The figure gives the estimated probability that the subject cor-

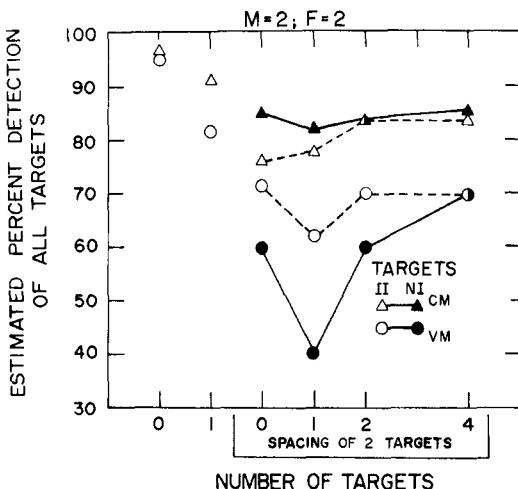


Figure 18. Data from Experiment 3a. Probabilities of correctly detecting one target when one was presented, or two targets when two were presented, as a function of spacing, for each condition. The data shown have been adjusted to remove effects of guessing and false alarms. The observed percentage of "none" responses to no-target trials is also shown. Frame time = 80 msec for the consistent-mapping (CM) conditions and 200 msec for the varied-mapping (VM) conditions. (The dashed or straight lines indicate the same thing as the open or closed points, respectively—namely, use of non-identical or identical items.  $M$  = memory-set size;  $F$  = frame size.)

rectly detects one target when one is presented or both targets when two are presented, for each condition. The left-most data point, however, is not an estimate but is instead the observed percentage of "zero" responses when "zero" targets were present. There are 1,440 observations per point in the conditions in which no target or one target was presented and 432 per point for each of the double-target conditions.

The estimation procedure used to construct this figure (and also all subsequent figures of a similar nature) is given in Appendix K. This estimation procedure is, in effect, a simple correction for guesses and false alarms. It should be noted that the pattern of the results is the same when the observed proportions of "one" responses to single targets and "two" responses to double targets are graphed instead. Thus, the correction procedure ad-

justs the quantitative but not the qualitative features of the results.

Consider first the VM results (the circles). The square of the probability of detecting a single target is about equal to the probability of detecting both targets at a spacing of 4. This finding is consistent with the hypothesis that the two targets are detected independently when the spacing is at least 4.

The VM data of greatest importance occur when two targets are presented. There are two main findings: First, there is the *Spacing 1 decrement*, that is, a considerable drop in performance when two targets are presented in adjacent frames. (In Part II, we shall see that it is the second of the two targets that is missed.) Second, there is a *target-similarity effect*: Performance is considerably lower when the two targets are physically different (NI) than when they are identical (II). It should be noted that the Spacing 1 decrement is quite sizeable and is somewhat larger in the NI than in the II condition. The target-similarity effect is also quite large, with performance for the II double targets averaging 11% higher (absolutely) than that for the NI double targets.

Thus in the VM conditions there is no question concerning the presence of the Spacing 1 effect and the target-similarity effect. The most important outcome of this experiment is that these effects are reversed in the CM conditions. The CM results are the upper curves indicated by triangles in Figure 18. First, the spacing effect is different than in the VM conditions: The greatest decrement in performance caused by the presence of two targets occurs when the spacing is zero rather than one. Second, the target-similarity effect is the reverse of that seen in the VM condition: Performance is lower when the two targets are identical than when they are different.

Thus, there is a clear qualitative change in the pattern of results from the VM conditions to the CM conditions. The spacing effect and the target-similarity effect are reversed in the two cases. We suggest that such an outcome provides evidence that the two search processes differ qualitatively as well as

quantitatively.<sup>5</sup> If the search in the CM conditions were simply a speeded version of the search in the VM conditions, then the pattern of results in the two conditions of Experiment 3 would be similar.

In Part II, considerable additional evidence will be presented attesting to the qualitative differences between the two modes of detection and search. Nevertheless, we feel a reasonably strong case has been made on the basis of the results of Experiments 1, 2, and 3 alone.

### *3. Additional Results and Speculations Concerning Experiment 3a*

In the VM conditions of Experiment 3a, the observed probability of reporting that one or more targets were present when none were in fact presented is .05, so this type of false alarm occurs very seldom. However, when one target was presented, the subjects fairly often said that two were present. This probability is not shown in Figure 18, but it is equal to .17. Why does this type of false alarm occur so frequently? It is possible that the subject was aware that he had difficulty distinguishing the presentation of one target from the presentation of two targets. If so, he might have decided to respond "two" on a proportion of the trials when only one was detected so as to better match the response frequencies with the presentation frequencies (given that at least one target was presented, there was a  $\frac{2}{3}$  probability that two targets were presented).

Let us examine next the possible processes underlying the spacing functions shown in Figure 18. Consider first the VM conditions. The Spacing 1 decrement has several possible interpretations. It may be that the subject spends extra time processing a frame on which a target is found, borrowing the additional time from the following frame. Alternatively, the detection of a target may be associated with a somewhat delayed, time-demanding, startle or registration response, in which the subject notes and counts the occurrence of a target.

The VM target-similarity effect may be

explained in terms of a search order, that is, a preferential order of scanning in favor of the target first located. When the spacing is zero in the II conditions and the first target is located, then the subject will continue his search using the same memory-set item, an item which by definition is appropriate for the second target. In fact, the very first comparison must be a match of the second target, since  $F = 2$  and we are assuming that the search of the frame finishes before a switch is made to the next memory-set item. The only way a second of two simultaneous identical items may be missed, according to this view, is for the subject to begin the search of the following frame just after the first target has been located, and before any further comparisons occur in the target frame.

On the other hand, if two simultaneous targets are not identical, after the first target is located, the subject could need as many as three additional comparisons and one additional switch to a new memory-set item before the next match occurs. These extra search steps increase the likelihood that the search of the next frame will begin before the second target is found, and hence, they decrease the probability that two simultaneous NI targets will both be found. We argue that this factor is large enough to overcome any advantage accruing to the NI condition, because the average number of comparisons until the first target is compared is lower in the NI condition (1.5) than in the II condition (2.0).

When the spacing is greater than zero, the target-similarity effect is likely due to an ordering of the search in subsequent frames so that comparisons will first be made for a memory-set item already located in a previous frame. Call this the search-order strategy.

---

<sup>5</sup> This conclusion is not quite as strongly justified as we would like, because the two conditions were run with different frame times. However, trying to equate the frame times would either have caused the VM performance to drop to chance or the CM performance to rise to perfection. In addition, we see no apparent reason why the change in frame times should have resulted in the differing patterns of results that were observed.

This search-order strategy presumably dissipates by the time Spacing 4 is reached.

Let us now turn to the CM spacing results. It is necessary to explain why performance is largely unaffected except when two identical targets occur simultaneously, in which case performance drops. Suppose that any target for which automatic detection has developed not only elicits an automatic-attention response but also becomes established in short-term memory with a stronger trace than distracting items. Then two such targets, if they are processed sufficiently well that an automatic-attention response occurs to each, will both remain in short-term store at a strength high enough to insure both will be found in a subsequent controlled search.<sup>6</sup>

Such a hypothesis needs to be modified for a multiple-frame task, however. When many frames arrive very rapidly, then the two separate visual codes for the automatic targets will tend to be replaced by the visual codes for subsequent items in those same positions. A more central verbal or visual code that is not position-specific will probably remain in short-term store for automatic targets even during the presentation of subsequent frames. If the two targets are different, then two different central codes will remain; if the two targets are the same, however, only one central code may remain—without position-specific codes, it may be difficult to tell whether the central code represents one or two different stimuli.

In short, we are arguing that in the CM conditions, the subject may occasionally be forced to rely not upon position-specific visual codes, but upon positionless central codes, for his decision regarding the number of targets present. The reliance upon central codes will not hurt performance if these codes differ (the NI condition) but will sometimes hurt performance if the central code is singular (the II condition). The detrimental effect of identical targets would be reduced as the target spacing increases, either because the position-specific codes could both be found, or because the central code would be perceived doubly on the basis of temporal separation.

In the absence of further research, all of these hypotheses are highly speculative. Whether one agrees with these hypotheses or not, however, the importance of the main findings is undiminished. In the VM conditions, the II performance is better than the NI performance, and the greatest performance decrement occurs at Spacing 1. In the CM conditions, the NI performance is better than the II performance, and the greatest performance decrement occurs at Spacing 0. These results suggest that automatic and controlled search are qualitatively differing detection processes.

#### B. *Experiment 3b: Detection of Multiple Targets When M and F are Varied*

The qualitative differences between CM and VM conditions in Experiment 3a point to qualitative differences in automatic and controlled search. As an extension and verification, we carried out similar experiments in which  $M$  and  $F$  were varied. We expected considerable changes in overall level of performance in the VM conditions, but we expected a qualitative pattern similar to the one seen when  $M = 2$  and  $F = 2$ . On the other hand, we expected little change at all in CM performance with variations in load, since Experiment 1 showed, at most, small effects in performance when load was varied in a study using single targets.

#### 1. Method

Experiment 3b was run exactly as Experiment 3a with the following exceptions. There were three basic conditions run in different blocks of trials: (a) a VM condition with  $M = 2$  and  $F = 4$ , (b) a VM condition with  $M = 4$  and  $F = 2$ , and (c) a CM condition with  $M = 2$  and  $F = 4$ . The VM conditions used a frame time of 200 msec and the CM conditions used a frame time of 80 msec.

---

<sup>6</sup>If the number of simultaneous targets became large enough, then this reasoning would of course not apply, because all items would have equal strength and forgetting would be almost as fast as that occurring when no targets are present.

## 2. Results and Discussion

The results of Experiment 3b are shown in Figure 19. Consider first the VM conditions when  $M = 2$  and  $F = 4$  (the circles). In this case, performance is much worse than when  $M = 2$  and  $F = 2$  (Figure 18), but it shows a similar qualitative pattern: Performance is lowest at Spacing 1, and performance in II conditions is superior to that in NI conditions. These results suggest no alteration in our previous hypothesis; as expected, the greater frame size reduced overall performance. Second, consider the VM results when  $M = 4$  and  $F = 2$  (the squares). Again, performance is much reduced compared to that in Figure 18 ( $M = 2, F = 2$ ), and again

the two main qualitative features remain: Performance is lowest at Spacing 1, and performance in II conditions is superior to that in NI conditions. However, one new effect can be seen: At Spacing 4, the II performance remains considerably higher than the NI performance.

It was suggested earlier that the performance superiority shown in II as compared to NI conditions at spacings greater than zero could be the result of a subject's strategy to search first for the memory-set item already found (termed the search-order strategy). We suggested that this tendency dissipated by Spacing 4, when  $M = 2$  and  $F = 2$ . The search-order strategy could be expected to dissipate, since it results in no real gain in

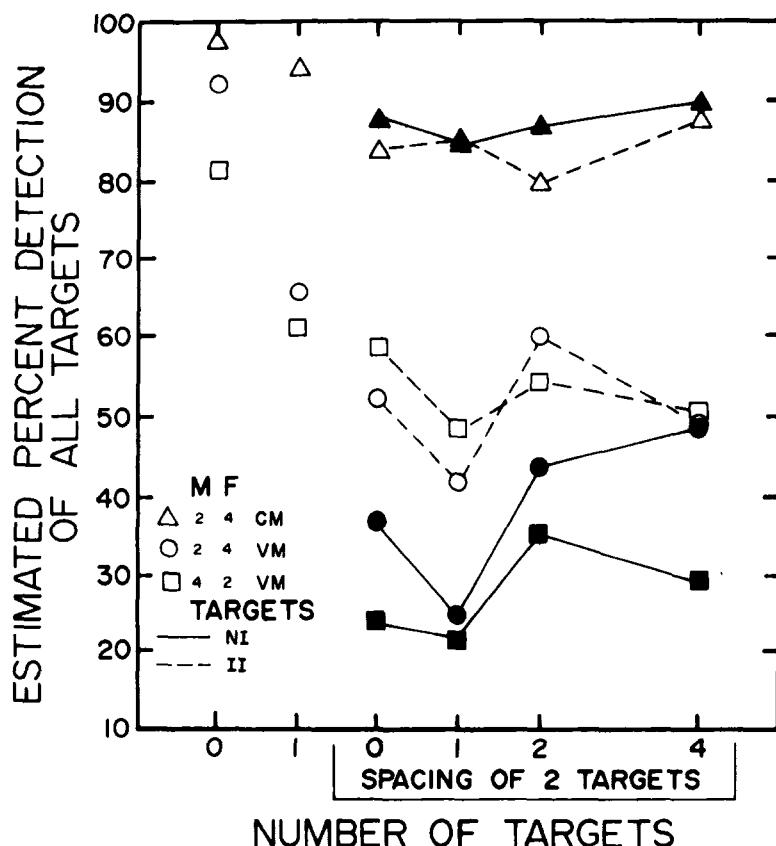


Figure 19. Data from Experiment 3b: Estimated percentages of detection as a function of number and spacing of targets. (CM = consistent mapping; VM = varied mapping; NI = nonidentical items; II = identical items; M = memory-set size; F = frame size.)

performance when  $M = 2$  and  $F = 2$ . However, when  $M = 4$  and  $F = 2$ , the search-order strategy is rational and beneficial at all spacings for the following reason. On 50% of the trials, the second target will be identical to the first. Thus when  $M = 4$ , the second target will match the previously found memory-set item on one half of the trials but will match a particular one of the other three memory-set items on only one sixth of the trials. Thus, when  $M = 4$ , it is clearly optimal at all spacings to search first for the memory-set item already found. When  $M = 2$ , the order of search is irrelevant, since the second target is equally likely to be either of the two memory-set items. In summary, therefore, the data at Spacing 4 indicate that the subjects tend to maintain the search-order strategy in those situations where it is in fact helpful.

When the spacing is zero, it is noteworthy that performance in the  $M = 4$ ,  $F = 2$ , II condition is better than that in the  $M = 2$ ,  $F = 4$ , II condition, but this finding reverses for the NI conditions. Selection on the basis of search order can help explain this result. For example, in the II condition, when  $M = 4$  and  $F = 2$ , following a detection of the first target, the very next comparison will involve the second target. When  $M = 2$  and  $F = 4$ , it could take as many as three additional comparisons before the second target is compared. The converse of this reasoning predicts a reversal of these effects for the NI conditions.

In summary, the VM conditions of Experiment 3b support the conclusions of Experiment 3a and show a qualitative pattern that is becoming a signature for those VM conditions that use multiple targets.

The CM conditions do not present such a nice picture. Although there is a slight superiority of performance in the NI conditions, the greatest difference occurs at Spacing 3 rather than at Spacing 1. A breakdown of the data for individual subjects showed high variability but did indicate that three of the four subjects were near ceiling (over 95% correct detection of two targets). The elimination of masks when  $F$  was raised to 4 may have contributed to this result. Apparently, masks reduce performance at the high speeds

of the CM conditions more than do characters, possibly because letter and digit sets of characters do not overlap at all dot positions. In view of these possible problems in Experiment 3b, we decided to replicate the CM conditions of Experiment 3a, though at a lower frame time. We hoped to verify the nature of the spacing effect and the target-similarity effects in the CM conditions.

### C. Experiment 3c: Multiple-Target Detection Using Consistent Mapping

This experiment was a direct replication of the CM conditions of Experiment 3a except that the frame time was 60 msec and 18 blocks were run for each subject. This change raised the total number of observations per two-target point, summing across subjects, to 648 (compared with 432 in Experiment 3a).

The results are shown in Figure 20. They duplicate the pattern of results in the CM conditions of Experiment 3a to an excellent degree. In this study, the difference between the II and NI conditions persisted up to Spacing 2 (three times 60 equals 180 msec), whereas in Experiment 3a, the difference lasted only up to Spacing 1 (two times 80 equals 160 msec). Thus, the time for which the NI conditions remain superior to the II conditions is roughly the same in Experiments 3a and 3c. Taken together, the results

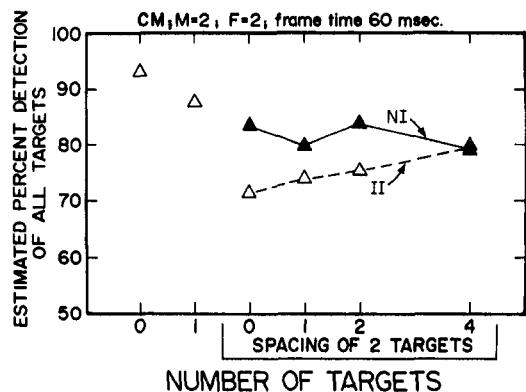


Figure 20. Data from Experiment 3c: A replication of the CM conditions of Experiment 3a, with frame time reduced to 60 msec. (CM = consistent mapping; NI = nonidentical items; II = identical items; M = memory-set size; F = frame size.)

of the two experiments leave no doubt concerning the basic shape of the CM spacing functions: Performance in the NI conditions is superior to that in the II conditions, and the greatest difference occurs when the spacing is zero.

The conclusions from Experiments 3a, 3b, and 3c may be summarized as follows: The two-target spacing functions show a radically different qualitative pattern for the CM and VM conditions. These results suggest that the controlled search utilized in the VM conditions is qualitatively as well as quantitatively different from the automatic detection used in the CM conditions.

#### D. A Selective Review of Multiple-Target Studies

Our review of the literature here will be limited to "multiple-frame" tasks requiring separate detection of multiple targets and using accuracy as a dependent measure. Moray (1975) has summarized a wide range of such studies, most from his laboratory. The general procedure in these studies involved two "streams" of information arriving simultaneously on two distinct channels (e.g., the two ears, two frequencies, two visual locations). Time-sharing conditions required two independent responses for the signals on the two channels—when signals arrived simultaneously, two nearly simultaneous responses were required. There were two control conditions: "dedicated," in which one channel was to be ignored, and "single channel," in which one channel contained no inputs.

The results of several of these studies that are relevant to our present concerns are as follows: (a) Time-sharing detection was as good as dedicated or single-channel detection if no signal was simultaneously detected on the other channel. (b) Time-sharing detection dropped when a simultaneous detection occurred on the other channel. (c) In two studies that varied the temporal relationship of two targets (Shaffer & Hardwick, 1969; Fitter, Note 4), it was shown that one target caused a decrement in the detection of the other target when they occurred within about .5 sec of each other.

These findings are quite consistent with ours, though less detailed. Our studies and Moray's both showed independent detection of two targets well-separated in time and a decrement in detection for two simultaneous targets. The Shaffer and Hardwick (1969) and Fitter (Note 4) studies also gave results similar to ours: When two targets occurred within about .5 sec of each other, a detection decrement occurred.

In at least some of the studies reviewed by Moray, automatic processing may have been the primary detection mechanism. This possibility is suggested by the fact that most of these studies used both consistent mappings and high degrees of practice. Whichever search mode was utilized, however, our results suggest that some type of detection decrement would have been associated with near-simultaneous presentation of two targets, and this was the result found.

#### V. Summary and Conclusions

We have presented a theory of information processing that emphasizes the roles of automatic and controlled processing.

Automatic processing is learned in long-term store, is triggered by appropriate inputs, and then operates independently of the subject's control. An automatic sequence can contain components that control information flow, attract attention, or govern overt responses. Automatic sequences do not require attention, though they may attract it if training is appropriate, and they do not use up short-term capacity. They are learned following the earlier use of controlled processing that links the same nodes in sequence. In search, detection, and attention tasks, automatic detection develops when stimuli are consistently mapped to responses; then the targets develop the ability to attract attention and initiate responses automatically, immediately, and regardless of other inputs or memory load.

Controlled processing is a temporary activation of nodes in a sequence that is not yet learned. It is relatively easy to set up, modify, and utilize in new situations. It requires attention, uses up short-term capacity,

and is often serial in nature. Controlled processing is used to facilitate long-term learning of all kinds, including automatic processing. In search, attention, and detection tasks, controlled processing usually takes the form of a serial comparison process at a limited rate.

Through several experiments, we demonstrated the salient characteristics of automatic and controlled processing by examining these processing modes in detection and search tasks. We have shown that *automatic detection* and *controlled search* are two quite different modes of performing such tasks, and we have delineated conditions that allow these two modes to be utilized. A summary of the main findings is given below.

Experiment 1 used a search paradigm (which could equally well be termed an attention paradigm) in which the accuracy of detection was examined in a variety of conditions. The results showed a clear dichotomy between VM (varied mapping) and CM (consistent mapping) presentation conditions that indicated the use of two qualitatively different detection mechanisms: controlled search and automatic detection, respectively.

The CM, automatic-detection conditions showed little effect of load and were all superior to the easiest VM, controlled-search condition. The VM conditions showed enormous effects of load, suggesting that controlled search is a limited serial comparison process.

Experiment 2 utilized a search paradigm with a single frame, with reaction time as a dependent measure. This study showed results comparable to those of Experiment 1: (a) The CM conditions showed little effect of load (i.e., no set-size effects), and (b) the VM conditions showed large and systematic increases in reaction time as load increased. Thus, the results support the distinction between controlled search and automatic detection suggested by Experiment 1.

The VM reaction times from Experiment 2 (replicating results by Briggs & Johnsen, 1973) were linear functions of set size, with negative slopes twice the size of positive slopes. A quantitative model was fit to these VM data with the following assumptions: (a) Controlled search is a serial, terminating comparison process; (b) the order of search

is to compare first a memory-set item to all display items in turn and then to choose a new memory-set item and continue; and (c) in addition to the time for each comparison, a certain amount of time is needed to choose each new memory-set item. This model fit the means and variances of the VM reaction time data successfully and did so better than a number of alternative models.

The model and parameters used to fit the controlled-search reaction time data from Experiment 2 were fit to the controlled-search accuracy data from Experiment 1. The fit was good enough to suggest that the same search and detection mechanisms underlay performance in the two paradigms. Thus, a link was forged between search tasks and attention tasks and between studies using reaction time measures and studies using accuracy measures.

The literature on search and detection tasks was reviewed in light of the present results and theory, and we concluded that the various results were in accord with ours if allowance was made for the possible use of different controlled strategies in certain tasks (e.g., an exhaustive rather than a serial comparison process).

The distinction between controlled search and automatic detection explains a number of puzzling results in the literature, particularly the occasional findings of flattened or curvilinear set-size functions. These can now be seen to be caused by the use of automatic detection that develops in consistent-mapping paradigms.

Multiple-frame search studies (Experiment 3) were carried out using multiple targets per trial. Strong negative interactions between detection of the two targets were seen in the VM conditions, but only small interactions were seen in the CM conditions. More importantly, the pattern of effects differed radically in the two cases. In the VM conditions, performance was hurt most when (a) the targets were in successive frames (the spacing effect), and (b) the targets were not identical (the target-similarity effect). In the CM conditions, performance was hurt most when the two targets were (a) in the same frame, and (b) identical to each other. These differences further supported the assumption

of a qualitative difference between the two detection modes.

Part II of this project (Shiffrin & Schneider, in press) will build on the empirical and theoretical base established in this paper. The present conclusions and findings will be supported, but more importantly, many new areas and implications will be explored. In particular, we shall examine the development of automatic detection; the role of training; the role of categorization; the role of controlled processing in the development of automatic processing; the means by which automatic processing may be modified; the importance of processes that attract attention automatically; and the limits upon the focusing and dividing of attention. A general processing theory will be presented and compared to previous models of search and attention.

### Reference Notes

1. Sternberg, S., & Scarborough, D. L. Parallel testing of stimuli in visual search. In *Proceedings of the International Symposium on Visual Information Processing and Control of Motor Activity*. Sofia, Bulgaria, July 1969.
2. Sternberg, S. *Scanning a persisting visual image versus a memorized list*. Paper presented at the meeting of the Eastern Psychological Association, Boston, April 1967.
3. Sternberg, S. *Estimating the distribution of additive reaction time components*. Paper presented at the meeting of the Psychometric Society, Niagara Falls, Ontario, Canada, October 1964.
4. Fitter, M. *Perceptual independence and divided attention*. Paper presented at the meeting of the Experimental Psychology Society, 1971.

### References

- Atkinson, R. C., Holmgren, J. E., & Juola, J. F. Processing time as influenced by the number of elements in a visual display. *Perception & Psychophysics*, 1969, 6, 321-326.
- Atkinson, R. C., & Juola, J. F. Factors influencing speed and accuracy of word recognition. In S. Kornblum (Ed.), *Attention and performance IV*. New York: Academic Press, 1973.
- Atkinson, R. C., & Juola, J. F. Search and decision processes in recognition memory. In D. H. Krantz, R. C. Atkinson, R. D. Lucek, & P. Suppes (Eds.), *Contemporary developments in mathematical psychology* (Vol. 1). San Francisco: Freeman, 1974.
- Atkinson, R. C., & Shiffrin, R. M. Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation: Advances in research and theory* (Vol 2). New York: Academic Press, 1968.
- Briggs, G. E. On the predictor variable for choice reaction time. *Memory & Cognition*, 1974, 2, 575-580.
- Briggs, G. E., & Johnsen, A. M. On the nature of central processes in choice reactions. *Memory & Cognition*, 1973, 1, 91-100.
- Broadbent, D. E. *Perception and communication*. London: Pergamon, 1958.
- Broadbent, D. E. *Decision and stress*. London: Academic Press, 1971.
- Corballis, M. C. Access to memory: An analysis of recognition times. In P. M. A. Rabbit & S. Dornic (Eds.), *Attention and performance V*. New York: Academic Press, 1975.
- Deutsch, J. A., & Deutsch, D. Attention: Some theoretical considerations. *Psychological Review*, 1963, 70, 80-90.
- Egeth, H., Atkinson, J., Gilmore, G., & Marcus, N. Factors affecting processing mode in visual search. *Perception & Psychophysics*, 1973, 13, 394-402.
- Egeth, H., Jonides, J., & Wall, S. Parallel processing of multi-element displays. *Cognitive Psychology*, 1972, 3, 674-698.
- Ellis, S. H., & Chase, W. G. Parallel processing in item recognition. *Perception & Psychophysics*, 1971, 10, 379-384.
- Eriksen, C. W., & Spencer, T. Rate of information processing in visual perception: Some results and methodological considerations. *Journal of Experimental Psychology Monograph*, 1969, 79(2, Part 2).
- Estes, W. K. Interactions of signal and background variables in visual processing. *Perception & Psychophysics*, 1972, 12, 278-286.
- Estes, W. K., & Taylor, H. A. A detection method and probabilistic models for assessing information processing from brief visual displays. *Proceedings of the National Academy of Sciences*, 1964, 52, 446-454.
- Estes, W. K., & Taylor, H. A. Visual detection in relation to display size and redundancy of critical elements. *Perception & Psychophysics*, 1966, 1, 9-16.
- Feller, W. *An introduction to probability theory and its applications*. New York: Wiley, 1957.
- Gardner, G. T. Evidence for independent parallel channels in tachistoscopic perception. *Cognitive Psychology*, 1973, 4, 130-155.
- Hick, W. E. On the rate of gain of information. *Quarterly Journal of Experimental Psychology*, 1952, 4, 11-26.
- Ingle, N. W. Categorization: A mechanism for rapid information processing. *Journal of Experimental Psychology*, 1972, 94, 239-243.
- Jonides, J., & Gleitman, H. A conceptual category effect in visual search: O as letter or as digit. *Perception & Psychophysics*, 1972, 12, 457-460.

- Kristofferson, M. W. Effects of practice on character classification performance. *Canadian Journal of Psychology*, 1972, 26, 54-60. (a)
- Kristofferson, M. W. When item recognition and visual search functions are similar. *Perception & Psychophysics*, 1972, 12, 379-384. (b)
- Kristofferson, M. W. Types and frequency of errors in visual search. *Perception & Psychophysics*, 1972, 11, 325-328. (c)
- Kristofferson, M. W., Groen, M., & Kristofferson, A. B. When visual functions look like item recognition functions. *Perception & Psychophysics*, 1973, 14, 186-192.
- LaBerge, D. Attention and the measurement of perceptual learning. *Memory & Cognition*, 1973, 1, 268-276.
- Moray, N. A. A data base for theories of selective listening. In P. M. A. Rabbit & S. Dornic (Eds.), *Attention and performance V*. New York: Academic Press, 1975.
- Neisser, U. Decision time without reaction time: Experiments in visual scanning. *American Journal of Psychology*, 1963, 76, 376-385.
- Neisser, U. *Cognitive psychology*. New York: Appleton-Century-Crofts, 1967.
- Neisser, U. Practiced card sorting for multiple targets. *Memory & Cognition*, 1974, 2, 781-785.
- Neisser, U., Novick, R., & Lazar, R. Searching for ten targets simultaneously. *Perceptual and Motor Skills*, 1963, 17, 955-961.
- Nickerson, R. S. Response times with a memory-dependent decision task. *Journal of Experimental Psychology*, 1966, 72, 761-769.
- Rabbitt, P. M. A. Ignoring irrelevant information. *British Journal of Psychology*, 1964, 55, 403-414.
- Rabbitt, P. M. A. Learning to ignore irrelevant information. *American Journal of Psychology*, 1967, 80, 1-13.
- Ross, J. Extended practice with a single character classification task. *Perception & Psychophysics*, 1970, 8, 276-278.
- Schneider, W. *Selective attention, memory scanning, and visual search: Three components of one process*. Unpublished doctoral dissertation, Indiana University, 1975.
- Shaffer, H., & Hardwick, J. Monitoring simultaneous auditory messages. *Perception & Psychophysics*, 1969, 6, 401-404.
- Shiffrin, R. M. The locus and role of attention in memory systems. In P. M. A. Rabbit & S. Dornic (Eds.), *Attention and performance V*. New York: Academic Press, 1975.
- Shiffrin, R. M. Capacity limitations in information processing, attention, and memory. In W. K. Estes (Ed.), *Handbook of learning and cognitive processes* (Vol. 4). Hillsdale, N.J.: Erlbaum, 1976.
- Shiffrin, R. M., & Gardner, G. T. Visual processing capacity and attentional control. *Journal of Experimental Psychology*, 1972, 93, 72-83.
- Shiffrin, R. M., & Geisler, W. S. Visual recognition in a theory of information processing. In R. L. Solso (Ed.), *Contemporary issues in cognitive psychology: The Loyola Symposium*. Washington, D.C.: Winston, 1973.
- Shiffrin, R. M., & Schneider, W. Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review*, in press.
- Simpson, P. J. High-speed scanning: Stability and generality. *Journal of Experimental Psychology*, 1972, 96, 239-246.
- Sorkin, R. D., & Pohlman, L. D. Some models of observer behavior in two-channel auditory signal detection. *Perception & Psychophysics*, 1973, 14, 101-109.
- Sperling, G. The description and luminous calibration of cathode ray oscilloscope visual displays. *Behavioral Research Methods and Instrumentation*, 1971, 3, 148-151.
- Sperling, G., Budiansky, J., Spivak, J. G., & Johnson, M. C. Extremely rapid visual search: The maximum rate of scanning letters for the presence of a numeral. *Science*, 1971, 174, 307-311.
- Sternberg, S. High-speed scanning in human memory. *Science*, 1966, 153, 652-654.
- Sternberg, S. Memory scanning: Mental processes revealed by reaction time experiments. *American Scientist*, 1969, 57, 421-457. (a)
- Sternberg, S. The discovery of processing stages: Extensions of Donder's method. In W. G. Koster (Ed.), *Attention and performance II*. Amsterdam: North Holland, 1969. (b)
- Sternberg, S. Memory scanning: New findings and current controversies. *Quarterly Journal of Experimental Psychology*, 1975, 27, 1-32.
- Swanson, J. M., & Briggs, G. E. Information processing as a function of speed versus accuracy. *Journal of Experimental Psychology*, 1969, 81, 223-229.
- Teichner, W. H., & Krebs, M. J. Visual search for simple targets. *Psychological Bulletin*, 1974, 81, 15-28.
- Townsend, J. T. A note on the identifiability of parallel and serial processes. *Perception & Psychophysics*, 1971, 10, 161-163.
- Townsend, J. T. Some results concerning the identifiability of parallel and serial processes. *British Journal of Mathematical & Statistical Psychology*, 1972, 25, 168-199.
- Treisman, A. M. Contextual cues in selective listening. *Quarterly Journal of Experimental Psychology*, 1960, 12, 242-248.
- Treisman, A. M. Verbal cues, language and meaning in selective attention. *American Journal of Psychology*, 1964, 77, 206-219.
- Treisman, A. M. Strategies and models of selective attention. *Psychological Review*, 1969, 76, 282-299.
- Treisman, A. M., & Riley, J. G. A. Is selective attention selective perception or selective response? A further test. *Journal of Experimental Psychology*, 1969, 79, 27-34.
- Yonas, A., & Pittenger, J. Searching for many targets: An analysis of speed and accuracy. *Perception & Psychophysics*, 1973, 13, 513-516.

## Index to Appendixes

<i>Appendix</i>	<i>Title</i>	<i>Page</i>
A	Glossary	55
B	Some Additional Procedural Details for the Experiments	56
C	The Effects of Masking	57
D	A Whole-Report Study	58
E	Procedural Details for Experiment 2	59
F	Predictions for Controlled-Search Models	59
G	Variance Predictions for Controlled-Search Model 1a	62
H	Predictions for Experiment 1 Using Controlled-Search Model 1a	63
J	A Replication and Interpretation of the Flat VM Set-Size Function When $M = 1$	64
K	Estimation Procedure for the Multiple-Target Experiments	66

## Appendix A

### Glossary (Arranged by Topic)

**Detection task.** Any procedure in which the subject attempts to detect the presence of any one of a specified, previously presented or learned set of characters in a set of one or more *inputs* presented simultaneously or sequentially.

**Search task.** A *detection task* in which the *inputs* are presented above sensory threshold so that each is clearly identifiable.

**Multiple-frame task.** A *detection task* in which the *inputs* are presented sequentially and rapidly, each presentation interval being termed a *frame*. Usually, detection accuracy is the dependent variable.

**Single-frame task.** A *detection task* in which all *inputs* are presented simultaneously. When the number of *inputs* is small, and the *inputs* are above threshold, accuracy is high and reaction time is the dependent variable.

**Threshold detection task.** A *detection task* that uses *inputs* near identification threshold.

**Memory set.** The previously presented or learned set of characters to be detected on a given trial of a *detection task*.

**Memory ensemble.** The set of characters from which the *memory sets* are chosen on various trials of a *detection task*.

**Memory-set size ( $M$ ).** The number of characters in the *memory set*.

**Inputs.** The set of presented characters in a *detection task*, among which the members of the *memory set* are to be sought.

**Display characters (or visual set).** The set of *inputs*, if these are presented visually.

**Frame (or frame set).** The subset of the *inputs* that are presented simultaneously.

**Frame size ( $F$ ).** The number of characters presented in a given *frame*.

**Load.** The momentary information-processing load in a *detection task*, defined as  $M$  times  $F$ .

**Frame time ( $f$ ).** The time from onset of a given *frame* to onset of the next.

**Targets.** Those members of the *memory set* that are actually present among the *inputs* on a given trial.

**Distractors.** Those *inputs* during a given trial that are not *targets*.

**Distractor set.** The set of characters from which the *distractors* are drawn on a given trial.

**Distractor ensemble.** The set of characters from which the *distractor sets* are drawn on various trials.

**Positive trial.** A trial on which at least one *target* is present.

**Positive response.** A response indicating a *target* was detected.

**Negative trial.** A trial on which no *target* is present.

**Negative response.** A response indicating no *target* was detected.

**Hit.** A *positive response* during a *positive trial*.

**Correct rejection.** A *negative response* during a *negative trial*.

**False alarm.** A *positive response* during a *negative trial*.

**Miss.** A negative response during a positive trial.

**Memory search.** A search task in which  $M$  is varied and  $F = 1$ .

**Visual search.** A visual search task in which  $F$  is varied and  $M = 1$ .

**Hybrid search.** A search task in which both  $M$  and  $F$  are varied.

**Consistent mapping (CM).** A search task in which, over trials, some memory-set items are never distractors (consistent target mapping), and/or some distractors are never memory-set items (consistent distractor mapping). These subsets of memory-set items or distractors that never change response assignments across trials are said to be CM items and are said to be "consistently mapped."

**Varied mapping (or VM).** A search task in which, over trials, all distractors are occasionally memory-set items, and all memory-set items are occasionally distractors.

**Set-size function.** In a reaction time task, the function giving mean reaction time as a function of either  $M$  or  $F$  or of  $M$  times  $F$ . (A

similar function may be obtained for the variances.)

**Slope.** The slope of the set-size function (in certain cases interpreted as a comparison time).

**Memory slope.** The slope when  $M$  varies and  $F$  is fixed.

**Frame slope.** The slope when  $F$  varies and  $M$  is fixed.

**Multiple targets** (or redundant targets). The occurrence of two or more targets on a trial.

**Spacing.** When more than one target is presented on a given trial, the spacing is the time between the onsets of any two successive targets. Alternatively, the spacing may be defined as one more than the number of frames intervening between the frames of any two successive targets, but a zero spacing indicates simultaneous targets.

**II.** When two physically identical targets are presented on a trial, they are termed II as a mnemonic for identical item.

**NI.** When two different targets from the memory set are presented on a trial, they are termed NI as a mnemonic for nonidentical item.

## Appendix B

### Some Additional Procedural Details for the Experiments

#### Equipment

A Digital Equipment Corporation PDP-8e computer was programmed to present all trials according to the appropriate randomizations, to collect responses, to give feedback, and to control all timing. The stimuli were presented on a DEC VR-14 cathode ray scope equipped with a P-24 fast decay phosphor. The dots making up any stimulus were refreshed periodically during the presentation time for that stimulus.

#### Stimuli

Each frame consisted of four characters in a square pattern (see Figure 2). The letters used were the consonants *C, D, F, G, H, J, K, L*, and *M*, and the digits were the numbers 1 to 9. The characters were constructed from dots on a rectangular grid 32 dots wide by 48 dots high, subtending a visual angle of  $.44^\circ$  in width and  $.66^\circ$  in height. The average number of dots per character was 43. These characters were highly discriminable and looked very similar to Letter Gothic typeface. They appeared to be con-

structed from continuous lines. Five different random dot masks were created. These were each made by randomly positioning 43 dots on the 32 by 48 dot matrix. These masks were assigned randomly to fill noncharacter display positions. The horizontal spacing between elements in a frame was  $.97^\circ$  visual angle; the vertical spacing was also  $.97^\circ$ . A fixation dot appeared at a position central to the four characters .5 sec prior to the first frame, and it remained there throughout the trial.

The frame generation time was approximately 15 msec, and the refresh rate was 4.2 msec. At the beginning of each frame, the computer required 15 msec to prepare the display, during which time the screen was blank. After generation, the average 173 dots of the four-character display were plotted in 4.2 msec. If all the dots were plotted in less than 4.2 msec, the plotting was not restarted until that period had expired. Thus, the intensity of each dot was independent of the number of dots displayed. The plotting was continued until the end of the frame time. The luminance directional energy per point was .5 candle-microsecond

(see Sperling, 1971). Background luminance of the screen was less than .01 footlambert.

### Procedures

Two subjects were run at one time; they viewed separate screens but received identical displays. Each subject had his own response-button box and headset. The two subjects were run simultaneously in the same room with a partition between them. White noise was presented over the headsets and through a speaker in the room to mask the noise of the other subject's responses. Subjects used chin rests and were about 65 cm from the display scope at all times.

The subject in Experiment 1 was to respond as soon as he or she saw the target, without waiting till the end of the trial. If no target was seen, the subject responded with a different key at the trial's end. When the 20 frames had been presented and when both subjects had responded, the trial was terminated and a feedback tone was given over the headset to any subject who had made an error. The tone remained on until the subjects indicated, by pressing the ready key, that they had memorized the memory set for the next trial. A

"target present" response was scored as correct only if it was given at least 200 msec after onset of the actual target.

Subjects were encouraged to take frequent rests. They could delay the next trial by not responding "ready" after the memory set was presented. Between blocks, subjects were allowed to leave the room if they wished. The subjects were instructed to maintain the highest possible level of accuracy.

The order of presentation of the 36 blocks was randomly permuted, with the constraint that within a session, a block with long frame durations (frame times of 400 msec or above) immediately follow a block with short frame durations (frame times of 200 msec or under).

The time to run a block ranged from 5 to 40 minutes, and sessions ranged from 40 to 80 minutes. Subjects typically finished one or two sessions a day, with at least 15 minutes, and generally several hours, between sessions. If subjects ever felt their performance was deteriorating due to fatigue, they were to terminate the session at the end of the block they were on. Experiment 1 required 14 sessions per subject, excluding practice sessions. The practice sessions included a representative sample of blocks.

## Appendix C

### The Effects of Masking

Does the nature of the elements that are adjacent to, precede, or follow a target affect detection accuracy? That is, do adjacent masks affect detection differently than adjacent characters? In Experiment 1, we did not collect the data in such a way that we could tell which elements were temporally or spatially adjacent to targets. We therefore carried out an additional study to collect data to examine *temporal* masking. The subjects of Experiment 1 were run in the VM conditions at  $f = 160$ ,  $M = 1$ , and  $F = 1, 2$ , or 4.

The results showed false alarms to be under 3% for all conditions. The target-present trials were analyzed with respect to the type of element preceding and following the target. If we let R represent a random dot mask and C a character, then the masking conditions were RR, CR, RC, and CC, where the first symbol denotes the element preceding the target and the second symbol denotes the element following the target (in the same display position as the target). The data are summarized in Figure C1.

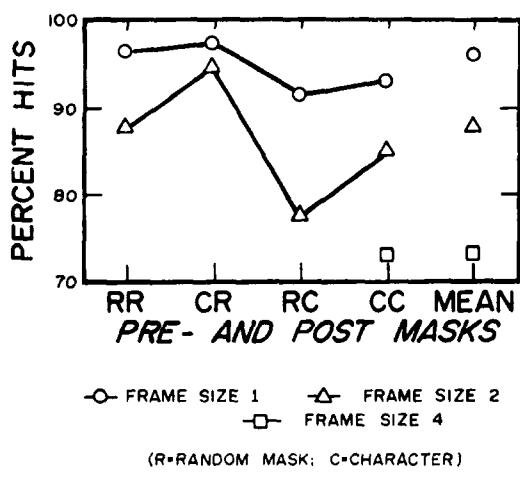


Figure C1. Hits as a function of the elements preceding and following the target, in the display position of the target. The first of each group of two letters on the horizontal axis designates the element preceding the target in time.

If the masking differences are responsible for the differences among conditions, then in a given masking condition (say, CC), performance should be identical for the different frame sizes. Quite clearly, the differences in performance for different frame sizes remain for each masking condition; thus, temporal masking cannot be used to explain the performance differences among the VM conditions that vary in frame size. There is a tendency seen in these data for targets to be detected better either when preceded by characters or followed by random dot masks. The following hypothesis is suggested by this result: If a serial search

occurs through each frame, there may be some tendency to search first the display positions that contained characters on the preceding frame.

We next asked, Do characters adjacent to targets cause lateral-masking effects different from those caused by random dot masks? In a study to be reported in Part II, we used a condition with frames of four characters, only two of which were to be attended. Performance in this condition was identical to that in a condition using two masks and two characters in each frame, suggesting that characters and random dot masks have roughly equal lateral-masking effects in the conditions of our studies.

#### Appendix D A Whole-Report Study

Since we argued that characters are clearly perceivable when presented at the frame times of the VM conditions in Experiment 1, and since we wished to establish that no perceptual difficulties would arise at the frame time of 160 msec utilized in Experiment 2, we carried out a whole-report task to establish the minimum frame time required to identify the characters of a display.

The procedure was similar to that of Experiment 2, except that the subject reported all the display characters in reading order, (rather than making a presence-absence judgment). The frame size was always 4, digits on some blocks of trials and consonants on others. The subject's task was to type the four presented characters in reading order on a keyboard and to guess when unsure. The independent variables were digits or consonants for report, and frame times of 40, 80, 120, and 160 msec were

used. All conditions were permuted between blocks within a session.

The results are tabulated in Table D1. There are 200 observations per point. The results may be described very simply: At frame times of 80 msec and above, performance was excellent. In fact, at  $f = 80$ , all four letters were correctly reported in correct order on 84% of the trials, and all four digits were correctly reported on 92% of the trials. At  $f = 40$  msec, however, these numbers dropped to 51% and 64%. Thus, the data indicate that the report probability falls off rapidly between 80 and 40 msec.

A detailed examination of Table D1 shows that the report decrement at 40 msec is primarily due to Display Positions 3 and 4 (the lower two characters and the last two reported). The following hypothesis seems reasonable: At 40 msec, each character is often

**Table D1**  
*Percentages of Correct Reports of Individual Letters (or Digits) and of All Letters (or Digits) in a Single Frame*

Frame time (msec)	Consonants							Digits						
	Single letter				All letters			Single digit				All digits		
	Position				1	2	3	4	Mean	1	2	3	4	Mean
160	98.5	97.5	94.5	94.5	96.3	89			99.5	98.5	99.5	97.5	98.8	97
120	99.0	97.5	93.0	95.0	96.1	86			99.5	99.5	99.5	98.0	99.3	98
80	97.0	94.5	91.5	92.0	93.8	84			99.5	99.0	97.5	94.0	97.5	92
40	93.5	89.5	69.0	68.0	80.0	51			94.5	97.0	80.5	79.5	87.9	64

perceived as an incomplete collection of features. The subject remembers and reports any characters that happen to be given complete processing but otherwise begins at the upper left (reading order) and tries to match the incomplete features against the possible target set before moving to the next display position, and so on. Presumably, the feature-matching process is reasonably successful for the first few display positions, but as the matching continues, features begin to be forgotten, so that the matching process fails badly for the last two display positions to be reported.

Most importantly, the results of this study show that the performance errors in the whole-report task arise at the same frame durations as the false alarms in the CM conditions of Experiment 1. Thus, the two results are probably related. We suggest that the common cause for both effects is the fact that stimuli are often incompletely encoded when frame time drops to 80 msec or less. Conversely, at frame times above 80 msec, the whole-report data suggest that the characters are clearly perceived perceptually (although the search process that matches these items may still fail).

## Appendix E Procedural Details for Experiment 2

Subjects were given at least one practice block in each condition before experimental data were collected. If the subject's accuracy during practice was below 90%, the block was repeated. Subjects required about five blocks of practice on the  $M = 4$ ,  $F = 4$ , VM condition before their accuracy was acceptable. Other conditions usually required only one block to reach acceptable accuracy levels. Thereafter, for each subject, two blocks of each condition were run over five experimental sessions. Typically, seven blocks were presented during each of the 1-hour sessions. The first block of each session had 30 practice trials (all others had 15) for which no data were recorded. After practice trials, each block con-

tained 120 trials, 60 with target present (called *positive trials*) and 60 with target absent (called *negative trials*) in random order. There were nine types of VM blocks and six types of CM blocks. These were presented in random order until the experiment was complete (except that the  $M = 2$  conditions were run later than the others).

Reaction times for positive and negative responses were recorded to the nearest .01 sec, but only mean and variance data were kept for each block. Trials in which reaction times exceeded 1.84 sec were excluded from analysis. If the subject responded incorrectly, he was given a feedback tone (indicating an error) during the presentation of the memory set for the next trial.

## Appendix F Predictions for Controlled-Search Models

Let RTN be the mean negative reaction time and RTP be the mean positive reaction time. The parameters are those already defined in the text in Section III.D.2. It is easiest in the following exposition to assume that memory and display switching time is *always* used prior to the first comparison; however, only the base-time parameters are affected by this assumption, since they are reduced by these switching times.

### *Model 1a*

Search is serial and terminating. The order of comparisons is to compare all frame items against one memory-set item and then to choose a new memory-set item and continue.

*Negative trials.* There are  $MF$  comparisons

and  $MF$  switches to new frame items. There are  $M$  switches to new memory-set items. Hence,

$$\text{RTN} = (c + d + \delta)MF + \mu M + B_N.$$

*Positive trials.* The search is equally likely to terminate on any of the  $MF$  comparisons, called the *target comparison*. Term a switch, decision, or comparison an *Event* ( $E$ ). Term the mean number of events up to and including the target comparison  $E(+)$ ; term the mean number following the target comparison  $E(-)$ ; and let the mean total number of events be  $E$ . By symmetry,  $E(+) = E(-) + 1$ . Also,  $E(+) + E(-) = E$ . Combining these equations,

$$E(+) = \frac{E + 1}{2}. \text{ Hence,}$$

$$\begin{aligned} RTP = & (c + d + \delta) \left( \frac{MF + 1}{2} \right) \\ & + \mu \left( \frac{M + 1}{2} \right) + B_P. \end{aligned}$$

*Model 1b*

Search is serial and the comparison order is the same as in Model 1a, but the search is exhaustive through each frame and terminates after the frame containing the target.

*Negative trials.* Same as in Model 1a:

$$RTN = (c + \delta)MF + (\mu + d)M + B_N.$$

*Positive trials.* Suppose the target is in the  $i^{\text{th}}$  comparison position in the frame and the  $j^{\text{th}}$  comparison position in the memory set. In this model, only  $j$  matters, so the positive reaction time is  $B_P + j[(c + \delta)F + (d + \mu)]$ . Thus, since all admissible values of  $i$  and  $j$  are equally likely,

$$\begin{aligned} RTP = & \sum_{j=1}^M \sum_{i=1}^F \{ j[(c + \delta)F \\ & + (d + \mu)] + B_P \} \frac{1}{MF} \\ = & (c + \delta) \left( \frac{MF + F}{2} \right) \\ & + (d + \mu) \left( \frac{M + 1}{2} \right) + B_P. \end{aligned}$$

*Models 2a and 2b*

These models are the same as Models 1a and 1b, respectively, except that the roles of memory-set size and frame size are reversed. Hence, in Model 2a for negative trials,

$$RTN = (c + d + \mu)MF + \delta F + B_N,$$

and in Model 2a for positive trials,

$$\begin{aligned} RTP = & (c + d + \mu) \left( \frac{MF + 1}{2} \right) \\ & + \delta \left( \frac{F + 1}{2} \right) + B_P. \end{aligned}$$

In Model 2b for negative trials,

$$RTN = (c + \mu)MF + (d + \delta)F + B_N,$$

and in Model 2b for positive trials,

$$\begin{aligned} RTP = & (c + \mu) \left( \frac{MF + M}{2} \right) \\ & + (d + \delta) \left( \frac{F + 1}{2} \right) + B_P. \end{aligned}$$

*Model 3*

Search is serial and terminating. The order of comparisons is randomly chosen without replacement.

*Negative trials.* There are always  $MF$  comparisons. The number of switches to new memory-set items may be calculated as follows. Label the memory-set items by the symbols  $a_1 \dots a_M$ . Label each of the  $MF$  comparisons by the symbol for the memory-set item it involves, and arrange these  $MF$  symbols in a randomly permuted row. By symmetry, the mean number of switches,  $S$ , must equal  $M$  times the mean number of switches to any one memory-set item, say  $a_1$ . Therefore, relabel all comparisons not involving  $a_1$  as  $a_2$ . The number of switches to  $a_1$ 's is equal to the number of runs of  $a_1$ 's in the new sequence. In general, the mean number of runs of  $a_1$ 's when  $r_1$   $a_1$ 's and  $r_2$   $a_2$ 's are arranged in a random row is

$$\frac{r_1(r_2 + 1)}{r_1 + r_2}.$$

This answer is given in Feller (1957, p. 446) as the solution to Problem IX.28, which in turn depends on Problem II.23. (The solution is straightforward, depending only upon extracting from the sum representing the mean, a series of terms representing a hypergeometric distribution.) In our case,  $r_1 = F$  and  $r_2 = (M - 1)F$ , since there are  $F$  comparisons labeled  $a_1$ . Hence, the mean number of switches to  $a_1$  equals

$$\frac{F[(M - 1)F + 1]}{F + (M - 1)F}.$$

Thus, the mean total switches is  $M$  times this quantity, or

$$\frac{MF[(M - 1)F + 1]}{F + (M - 1)F} = MF - F + 1.$$

By similar reasoning, the mean number of switches to new frame items must be  $MF - M + 1$ . Thus,

$$\begin{aligned} RTN = & (c + d)MF + \mu(MF - F + 1) \\ & + \delta(MF - M + 1) + B_N. \end{aligned}$$

*Positive trials.* As explained earlier, if  $E$  is the mean total number of events (comparisons, decisions, switches) on a negative trial, then the mean number on a positive trial is

$$E(+) = \frac{E + 1}{2}.$$

Hence,

Table F1  
*Parameter Estimates for Model 3*

Study	$B_N$	$B_P$	$c + d$	$\mu$	$\delta$	Standard error of prediction
Schneider & Shiffrin (Experiment 2)	623	552	13	45	0	21.3
Briggs & Johnsen (1973)	569	500	38	29	0	23.7

$$\begin{aligned} RTP = & (c + d) \left( \frac{MF + 1}{2} \right) \\ & + \mu \left( \frac{MF - F + 2}{2} \right) \\ & + \delta \left( \frac{MF - M + 2}{2} \right) + B_P. \end{aligned}$$

*Best fit.* The best-fitting parameter estimates for this model for the data from the two studies are shown in Table F1.

*Discussion of Model 3.* It is clear from comparing Table F1 with Table 1 that the best fitting model is Model 3. Model 3 provides the best fit because it comes closest to predicting the small slope of the visual set-size function when  $M = 1$  (as seen in the first panel in Figure 6). In Model 3, the slope when  $M = 1$  depends only upon  $c + d$  (since  $\delta$  is estimated to be zero), yet the slope when  $M = 2$  or  $M = 4$  depends also upon  $\mu$ . Thus,  $c + d$  is estimated to be a small quantity (13) and  $\mu$  is estimated to be a large quantity (45). Thus Model 3 simultaneously predicts a small slope when  $M = 1$  and a large slope when  $M = 2$  or  $M = 4$ . In short, the switching time, not the comparison time, is largely responsible for the observed slopes, and no switches occur when  $M = 1$ . In Models 1 and 2, the same parameters are involved in the slope predictions for all the values of  $M$ , so these models cannot predict a small slope for  $M = 1$  without underpredicting the slopes when  $M = 2$  and  $M = 4$ .

Despite the fact that Model 3 provides the best fit to the data, we do not feel this model offers a reasonable description of the subject's actual search process. The reasons are as follows. First, there is a possibility that the low slope when  $M = 1$  reflects a change in search strategy: a shift from serial controlled search

for  $M > 1$  to some sort of parallel controlled search when  $M = 1$ . This possibility is explored in Appendix J.

Second, other researchers who have replicated the conditions of  $M = 1$  with  $F$  varying (e.g., Atkinson, Holmgren, & Juola, 1969) have not found small slopes but, rather, slopes near 30 to 40 msec per item, a result in accord with the predictions of Model 1a. The reason our subjects may have differed from previous subjects is also discussed in Appendix J.

Third, the  $M = 1$ ,  $F = 1$  negative data point from our study appears a bit aberrant. For this point, both the mean and the variance are higher even than for the  $M = 1$ ,  $F = 2$  negative data point, a finding difficult to predict for any model. This point may be excessively high due to sampling error or contamination by some factor of which we are unaware. Whatever the reason, this point should not be weighted too heavily in deciding among models. Yet the advantage of Model 3 over Model 1a is largely due to this point. About one third of the total sum of squared deviations for Model 1a applied to our data arises from this point alone.

Fourth, Model 3 is implausible as a realizable search mechanism. Random search without replacement requires an incredibly efficient record-keeping device to prevent repetition of previously carried out comparisons. Surely such record keeping would require a great deal of time if it were carried out accurately. Furthermore, even if such record keeping were possible, the subject could avoid the problems and time associated with the record-keeping process and could reduce his reaction time considerably through a reduction in the number of switches required, simply by adopting an ordered search like that proposed for Model 1a.

## Appendix G

## Variance Predictions for Controlled-Search Model 1a

*Variance of a Probabilistic Mixture*

$$\text{Since } \rho_i^2 = \sigma_i^2 + \mu_i^2,$$

Suppose the random variable  $Y$  is a probabilistic mixture of the random variables  $X_1, X_2, \dots, X_n$ . Then we can say that

$$P(Y = w) = \sum_{i=1}^n a_i P(X_i = w)$$

where

$$\sum_{i=1}^n a_i = 1.0, \quad 1 \geq a_i \geq 0.$$

Then,

$$E[Y] = \sum_{i=1}^n a_i E[X_i] = \sum_{i=1}^n a_i \mu_i$$

where  $\mu_i = E[X_i]$ . Let  $E[X_i^2] = \rho_i^2$  and  $\text{Var}[X_i] = \sigma_i^2$ . Therefore,

$$\begin{aligned} \text{Var}[Y] &= E[Y - E(Y)]^2 \\ &= \int (Y - E(Y))^2 p d(Y) dY \end{aligned}$$

$$= \int [w - (\sum_{i=1}^n a_i \mu_i)]^2$$

$$\begin{aligned} &\times \sum_{i=1}^n a_i p d(X_i = w) dw \\ &= \int [w^2 + (\sum_{i=1}^n a_i \mu_i)^2 - 2w \sum_{i=1}^n a_i \mu_i] \end{aligned}$$

$$\times \sum_{i=1}^n a_i p d(X_i = w) dw$$

$$= \sum_{i=1}^n a_i \int w^2 p d(X_i = w) dw$$

$$+ (\sum_{i=1}^n a_i \mu_i)^2 \sum_{i=1}^n a_i \int p d(X_i = w) dw$$

$$- 2(\sum_{i=1}^n a_i \mu_i) \sum_{i=1}^n a_i \int w p d(X_i = w) dw$$

$$= \sum_{i=1}^n a_i \rho_i^2 + (\sum_{i=1}^n a_i \mu_i)^2 \sum_{i=1}^n a_i$$

$$- 2(\sum_{i=1}^n a_i \mu_i) \sum_{i=1}^n a_i \mu_i$$

$$= \sum_{i=1}^n a_i \rho_i^2 - (\sum_{i=1}^n a_i \mu_i)^2.$$

$$\text{Var}[Y] = \sum_{i=1}^n a_i (\sigma_i^2 + \mu_i^2)$$

$$- (\sum_{i=1}^n a_i \mu_i)^2. \quad (G-1)$$

In many cases of interest,  $a_i = 1/n$ . Then,

$$\text{Var}[Y] = (1/n) \sum_{i=1}^n (\sigma_i^2 + \mu_i^2)$$

$$- (1/n^2) (\sum_{i=1}^n \mu_i)^2$$

$$= (1/n) [\sum_{i=1}^n \sigma_i^2] + (1/n)$$

$$\times \left\{ \sum_{i=1}^n \mu_i^2 - \frac{(\sum_{i=1}^n \mu_i)^2}{n} \right\}$$

$$= (1/n) [\sum_{i=1}^n \sigma_i^2] + \sigma_K^2 \quad (G-2)$$

where "Variance" of the  $\mu_i$ 's is

$$\sigma_K^2 = \frac{\sum_{i=1}^n \mu_i^2 - \frac{(\sum_{i=1}^n \mu_i)^2}{n}}{n}.$$

*Model 1a*

This model assumes serial terminating search, first through the frame for each memory-set item. We shall assume that the first switch does not occur until the change to the second memory-set item. Let  $O_N^2$  be the observed negative trial variance and  $O_P^2$  be the observed positive trial variance. All the components of the process are assumed to be independent.

*Negative variance.* There are always  $MF$  comparisons, decisions, and visual switches, and there are always  $(M - 1)$  switches to new memory-set items. Hence,

$$O_N^2 = MF\sigma_c^2 + (M - 1)\sigma_\mu^2 + \sigma_N^2. \quad (G-4)$$

*Positive variance.* There is a probabilistic mixture. The probability is  $1/MF$  that there are exactly  $i$  comparisons,  $i$  ranging from 1 to  $MF$ . When there are  $i$  comparisons, there are exactly  $[i - 1/F] = j(i)$  switches to new

memory-set items (where the brackets denote the largest integer contained in the enclosed fraction). So the  $i$ th element of the mixture has a variance of  $i\sigma_c^2 + j(i)\sigma_\mu^2$  and a mean of  $ic + \mu j(i)$ , where  $c$  is an abbreviation for  $(c + d + \delta)$ . Using Equation G-2, then,

$$\begin{aligned}\sigma_K^2 &= \frac{1}{MF} \left\{ \sum_{i=1}^{MF} [ic + \mu j(i)]^2 - \frac{1}{MF} \left[ \sum_{i=1}^{MF} (ic + \mu j(i)) \right]^2 \right\} \\ &= \frac{c^2(MF+1)(2MF+1)}{6} + \frac{\mu^2 F(M-1)M(2M-1)}{6} + \frac{2c\mu}{MF} \left( \frac{F^2(M-1)M(2M-1)}{6} \right. \\ &\quad \left. + \frac{F(F+1)M(M-1)}{4} \right) - \frac{1}{M^2F^2} \left( c \frac{MF(MF+1)}{2} + \mu \frac{MF(M-1)}{2} \right)^2 \\ &= \frac{c^2}{12} (MF+1)(MF-1) + \frac{\mu^2}{12} (M-1)(M+1) + \frac{\mu c}{6} (M-1)F(M+1).\end{aligned}$$

Substituting into Equation G-5,

$$\begin{aligned}O_P^2 &= \frac{1}{MF} \left\{ \sigma_c^2 \frac{MF(MF+1)}{2} + \sigma_\mu^2 \frac{FM(M-1)}{2} \right\} + c^2 \left\{ \frac{M^2F^2-1}{12} \right\} + \mu^2 \left\{ \frac{M^2-1}{12} \right\} \\ &\quad + \mu c \left\{ \frac{(M^2-1)F}{6} \right\} + \sigma_P^2 \\ O_P^2 &= c^2 \left\{ \frac{M^2F^2-1}{12} \right\} + \mu^2 \left\{ \frac{M^2-1}{12} \right\} + \mu c \left\{ \frac{(M^2-1)F}{6} \right\} + \sigma_c^2 \frac{MF+1}{2} \\ &\quad + \sigma_\mu^2 \frac{M-1}{2} + \sigma_P^2.\end{aligned}\quad (G-6)$$

## Appendix H

### Predictions for Experiment 1 Using Controlled-Search Model 1a

We shall assume the following model for the VM conditions of Experiment 1, based directly upon Model 1a.

1. The available time used for searching each frame is  $e = f - B_A$ .

2. If a target is present but is not found during the available time, then an error (a "miss") will occur.

3. The search scheme used during each frame is that of Model 1a described earlier. The entire display is searched for each memory-set item in turn. The subject terminates the search when a target is found.

4. Each comparison and decision takes a time,  $K(t)$ , that has a mean,  $c + d + \delta$ , of 40 msec (from the Experiment 3 model) and that has a gamma distribution with variance  $\sigma_o^2$ . The mean is abbreviated as  $c$ .

5. Each switch to a new memory-set item

takes a time,  $S(t)$ , that has a mean,  $\mu$ , of 42 msec (from the fit of the model to Experiment 2) and a gamma distribution with variance  $\sigma_\mu^2$ .

6. A switch to a new memory-set item is not necessary at the start of the search of each frame. Switches occur prior to the recruitment of the second and subsequent memory-set items in the search of any frame.

To quantify the model, using the typical form of the gamma distribution:

$$K(t) = \frac{1}{\Gamma(\nu_1)} \alpha_1^{\nu_1} t^{\nu_1-1} e^{-\alpha_1 t}, \quad t \geq 0, \quad (H-1)$$

where  $\frac{\nu_1}{\alpha_1} = c = 40$  msec, and  $\alpha_1 > 0, \nu_1 > 0$ .

$$S(t) = \frac{1}{\Gamma(\nu_2)} \alpha_2^{\nu_2} t^{\nu_2-1} e^{-\alpha_2 t}, \quad t \geq 0, \quad (H-2)$$

where,  $\frac{\nu_2}{\alpha_2} = \mu = 42$  msec, and  $\alpha_2 > 0$ ,  $\nu_2 > 0$ .

Then, due to the convolution properties of the gamma distribution, the distribution of the sum of  $i$  samples, from Equation H-1, is

$$Q_i(t) = \frac{1}{\Gamma(i\nu_1)} \alpha_1^{i\nu_1} t^{i\nu_1 - 1} e^{-\alpha_1 t}, \quad (\text{H-3})$$

and the distribution of the sum of  $j$  samples, from Equation (H-2), is

$$R_j(t) = \frac{1}{\Gamma(j\nu_2)} \alpha_2^{j\nu_2} t^{j\nu_2 - 1} e^{-\alpha_2 t}. \quad (\text{H-4})$$

According to the model, if there are  $i$  total comparisons up to a given point in the search, there will be exactly  $[i - 1/F]$  total recruitments of new memory-set items up to that same point in the search, where the brackets denote the largest integer contained in the fraction enclosed by the brackets. Thus we set  $j = [i - 1/F]$  in the following equations. The total time utilized to complete  $i$  comparisons and  $j$  switches is then the sum of the times in Equations H-3 and H-4 (the asterisk denotes convolution):

$$T_k(t) = Q_i(t) * R_j(t) = \int_0^t Q_i(t-y) \times R_j(y) dy. \quad (\text{H-5})$$

Finally, suppose that for a given frame, the subject compares items in some order, an order that will not be completed if available time happens to run out. Due to the experimental

design and the assumptions of the model, a target is equally likely to be in any of the  $MF$  possible comparisons positions. Thus, the

$$P(\text{Hit} | f, M, F) = \sum_{k=1}^{MF} P(\text{correct} | \text{target is } k^{\text{th}} \text{ in}$$

the order of scanning)  $\times P(\text{target is } k^{\text{th}} \text{ in the order of scanning})$ . The last term is equal to  $1/MF$ , so

$$P(\text{Hit} | f, M, F) = \frac{1}{MF} \sum_{k=1}^{MF} \left\{ \int_0^e T_k(t) dt \right\}. \quad (\text{H-6})$$

Equation H-6 holds since  $P(\text{Hit})$  is just the probability that the total time until the target is compared in the search is less than the total search time available, and the total available time is  $e$ , by hypothesis.

If  $B_A$  is assumed to be zero, and the estimates for the means and variances are carried over from the fit to Experiment 2, then this model may be fit directly to the 18 hit rates in the VM curves of Figure 3. When  $B_A$  was fit to the data, it was estimated to be 45 msec. When  $B_A$ ,  $\sigma_e^2$ , and  $\sigma_\mu^2$  were fit to the data, and the means were carried over from the fit to the data of Experiment 3, the estimates were  $B_A = 45$  msec;  $\sigma_e^2 = 4,200$  msec $^2$ ;  $\sigma_\mu^2 = 11,428$  msec $^2$ .

The fit was carried out through a numerical grid searching process that minimized chi-square across the hit rates. Note that the mean of gamma is  $\nu/\alpha$  and the variance is  $\nu/\alpha^2$ , so these quantities were set equal to the appropriate values from the fit of the data from Experiment 2.

## Appendix J

### A Replication and Interpretation of the Flat VM Set-Size Function When $M = 1$

Why does the  $M = 1$ , VM, memory-set-size function in Experiment 2 have a near zero slope? (The slopes for each subject are given in Table 2.) This question is particularly important not only because our model fails to predict such a finding but also because the accuracy data from Experiment 1 show considerable decrements (as  $F$  increases) when  $M = 1$ . We will first replicate the finding, and then explore the hypotheses (a) that a switch in search strategy occurs when  $M = 1$  and (b) that the constancy of reaction times might be maintained through a trade-off of time for errors.

We designed an experiment with a procedure similar to that of Experiment 2, but the frames following and preceding the target frame were

filled with varying numbers and types of distracting characters. Define Frame 3 as the target frame and Frames 1, 2, 4, and 5 as the nontarget frames. The subject was told that only the characters presented on the target frame were relevant. Indeed, if there was a target present it was always in the target frame.  $M$  was set equal to 1 for all conditions, while  $F$  (for the target frame) was equal to 1, 2, or 4 in different blocks.

The novel procedure in this study was the introduction of distracting characters in the nontarget frames. Each nontarget frame contained the same number of distracting characters, termed  $F_n$ .  $F_n$  was equal to 1, 2, or 4 in different sessions. These conditions were all

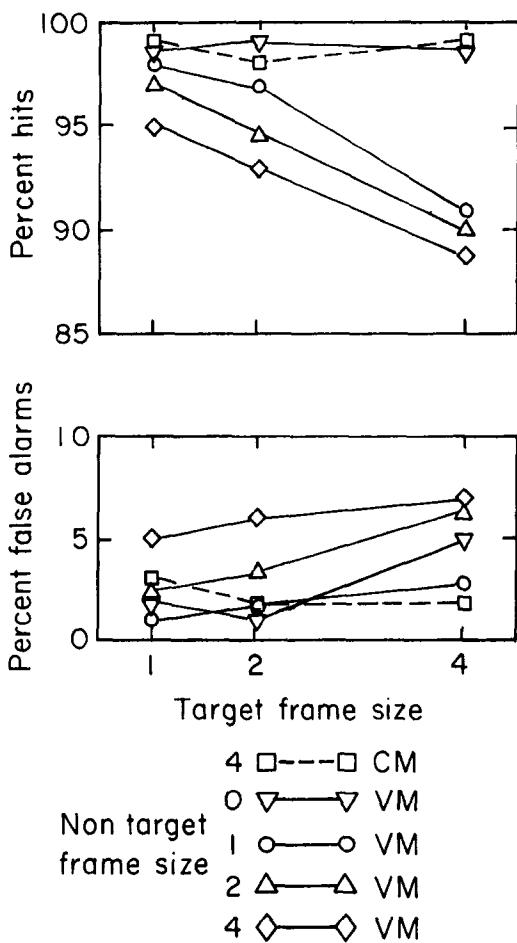


Figure J1. Hits and false alarms as a function of the number of characters in the target frame, for conditions in which the number and type of characters in the frames preceding and following the target were varied. (CM = consistent mapping; VM = varied mapping.)

run under the VM procedure. One CM session was also run in which  $F_n$  was always equal to 4 and  $F$  was equal to 1, 2, or 4 in different blocks. In the CM session, the target, if one appeared, was of a different class than the other frame items (e.g., a number target in letter frames).

Nine blocks of 50 trials each were run in each of five sessions. In each session there were three blocks at each value of  $F$  (in random order). Sessions 1 and 2 used  $F_n = 2$ , Sessions 3 and 4 used  $F_n = 4$ , and Session 5 used  $F_n = 1$ . Only Session 4 used the CM procedure. Both reaction times and percentages of correct responses were recorded.

The percentages of correct detections when a target was present (the hit rate) and the percentages of incorrect detections when a target was not present (the false-alarm rate) are given in Figure J1. The reaction times for hits and for correct rejections are given in Figure J2. The inverted triangles show data from Experiment 2 for comparison purposes (in Experiment 2,  $F_n = 0$ ).

The  $F_n = 0$  condition shows a rise in false alarm to 5% when  $F = 4$ . This was the finding from Experiment 2 that suggested the possibility of a time-accuracy trade-off (see the flat reaction time function for  $F_n = 0$  in Figure J2). The conditions of the present experiment make such an explanation even more plausible, because the reaction time functions have very small slopes for all values of  $F_n$ , but only at the cost of a considerable increase in errors. Thus when  $F_n = 4$  and  $F = 4$ , the percentage of hits drops to 88% and the percentage of false alarms rises to 7.5%. The  $F_n = 4$ ,  $F = 4$  condition is particularly important because it closely parallels the  $F = 4$  condition in Ex-

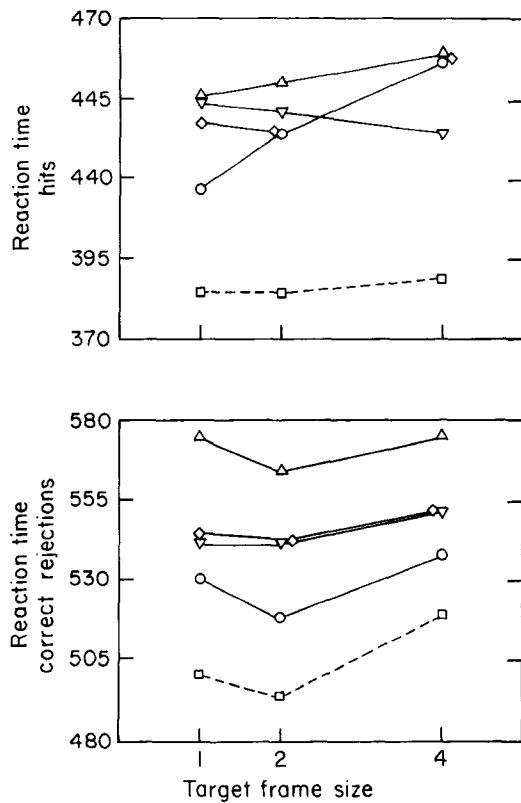


Figure J2. Mean reaction times for hits and for correct rejections, for conditions corresponding to those in Figure J1.

periment 1, in which every frame contains 4 characters. Such a high error rate per frame could not have been tolerated in the 20-frame procedure of Experiment 1, so the subjects were probably led to adopt a more conservative search strategy, perhaps the serial, terminating comparison process we have proposed.

It should next be asked what strategy the subjects adopt that allows them to produce relatively flat reaction time functions (for  $M = 1$ ) at the expense of an increase in errors. It is possible that the subjects are attempting to use some form of parallel search in spite of the VM presentation conditions. This parallel process obviously gives reaction times (but not error rates) similar to those arising from the automatic parallel process that occurs in the CM condition. However, we wish to suggest that this is a controlled parallel search that requires attention, rather than an automatic detection. We shall term this possible parallel search in the VM conditions *controlled parallel search* to distinguish it from automatic detection. We wish to distinguish these two types of parallel process both because the  $M = 1$  conditions differ as  $F$  varies in the multiple-frame task and because the subjects report that all the VM conditions, including  $M = 1$ , require a great deal more effortful attending than any of the CM conditions. (It is interesting to note that the switch in strategy in the VM conditions from serial to controlled parallel search does not seem to improve performance beyond the level to be expected from a serial search—see Figures 10 and 12. In fact, the switch in strategy seems to harm performance slightly—but the subjects may not be aware of this fact.)

The possibility that some form of (inefficient) parallel controlled search might be usable when  $M = 1$  is suggested by certain findings of LaBerge (1973). He found that subjects matched two unknown characters (to each other) more slowly than two known characters, if the subjects were unprepared for the characters and the matching test. However, the matching times for the known and unknown pairs of characters were about equal when the subjects were prepared for the test and the characters. It could be argued that our search task when  $M = 1$  corresponds to the “prepared” conditions in LaBerge's study. He argued that the features making up an unknown character could be temporarily “unitized” when attention was directed to that character. In a roughly similar vein, we could argue that attention placed on a single character in memory might temporarily allow some type of inefficient parallel search to take place, even in VM conditions.

In summary, then, we have tentatively suggested that for the  $M = 1$  conditions of Experiment 2, subjects are adopting a special search strategy (very possibly a controlled parallel search) in which reaction times are kept constant across frame sizes at the cost of an increase in errors. Such a strategy cannot be utilized, however, in the 20-frame procedure of Experiment 1, since the resultant error rates would be too large. In Experiment 1, therefore, subjects utilize the usual serial, terminating strategy even when  $M = 1$ , with the result that accuracy drops as  $F$  increases. Obviously, additional research is necessary to test these hypotheses.

## Appendix K

### Estimation Procedure for the Multiple-Target Experiments

Let  $S_{i,j}$  be the real detection of  $i$  targets out of  $j$  targets presented ( $i \leq j$ ). Let  $R_{i,j}$  be the report of  $i$  targets, given  $j$  targets presented. Let  $R_k|S_i$  be the report of  $k$  targets, given  $i$  targets detected ( $k \geq i$ ). Then

$$P(R_{k,j}) = \sum_{i=0}^j P(R_k|S_i)P(S_{i,j}).$$

We assume  $P(S_{0,0}) = 1$ . Simple calculations then show that

$$P(S_{0,1}) = P(R_{0,1})/P(R_{0,0})$$

$$P(R_1|S_1) = \frac{P(R_{1,1}) - P(R_{1,0})P(S_{0,1})}{1 - P(S_{0,1})}$$

$$P(S_{1,1}) = 1 - P(S_{0,1})$$

$$P(S_{0,2}) = P(R_{0,2})/P(R_{0,0})$$

$$P(S_{1,2}) = \frac{P(R_{1,2}) - P(R_{1,0})P(S_{0,2})}{P(R_1|S_1)}$$

$$P(S_{2,2}) = 1 - P(S_{0,2}) - P(S_{1,2}).$$

The plotted values in Figures 18, 19, and 20 are  $P(R_{0,0})$ ,  $P(S_{1,1})$ , and  $P(S_{2,2})$ . The raw data are reported in Schneider (1975).

Received July 21, 1976 ■