INTERACTIONS BETWEEN MEMORY SCANNING AND VISUAL SCANNING IN PROCESS MONITORING

Yili Liu
Department of Industrial & Operations Engineering
The University of Michigan
Ann Arbor, Michigan 48109-2117

Technical Report 95-23

November 1995

Interactions between Memory Scanning and Visual Scanning in Process Monitoring

Yili Liu, Ph.D.
Department of Industrial and Operations Engineering and Center for Ergonomics
University of Michigan
1205 Beal Ave.
Ann Arbor, MI 48109-2117

*Correspondence to Prof. Yili Liu, Department of Industrial and Operations Engineering, University of Michigan, 1205 Beal Avenue, Ann Arbor, Michigan 48109, USA
Phone #: (313)763-0464
Fax #: (313)764-3451
email: yililiu@engin.umich.edu

Keywords: visual scanning, memory scanning, human performance models,
multitask performance, human-machine interface,
monitoring behavior

Abstract

Many real-world tasks require the simultaneous performance of memory scanning of several memorized items and visual scanning of several physically separated sources of information in the visual field. This paper reports a study that was conducted to quantify the possible interactions between memory scanning and visual scanning. A quantitative model was derived to integrate Sternberg's linear model of memory scanning and Neisser's linear model of visual scanning, and the derived model was tested through two experiments. The experiments simulated a process controller's task of monitoring an array of instrument meters to detect if any of them indicated a system error. The subjects were required to keep a number of items in their working memory (the definition of errors) and search through an organized array of instrument meters to decide whether any of the meters carried an item that matched any of the memorized items in their working memory. Two experimental factors were investigated in both experiments: the number of memorized items and the number of circles needed to be searched, and they defined memory scanning demand and visual scanning demand respectively. Experiment 1 employed a different set of memorized items for each experimental trial, whereas Experiment 2 employed the same set for all the trials that had the same experimental condition. The experiments identified both the strengths and the limitations of the derived model. Implications for human-machine interface design and human performance modeling are discussed.

1. Introduction

One common characteristic of many everyday tasks is the need to perform a number of concurrent activities, which may require the simultaneous performance of memory scanning of several memorized items and visual scanning of several physically separated targets in the visual field. For instance, a process controller often need to keep a number of items in her working memory (e.g., the definition of system failures or the emergency status of a number of monitored processes), while constantly scanning the process displays to detect errors or failures that may occur in any of the monitored processes. As an illustration, we can imagine a situation in which a process controller monitors eight widely-separated moving-pointer clock-like displays and a system failure occurs when any of the eight displays shows a three-o'clock or nine-o'clock position. In this illustrative example, the process controller constantly compares the two items in her working memory (the two clock positions that define process failures) with the status of each of the eight widely-separated displays. This active interplay between memory scanning of several items in working memory and visual scanning of physically separated sources of information is also characteristic of the tasks of a quality inspector, a telephone operator, or an air-traffic controller.

Memory scanning and visual scanning are two aspects of human performance that have enjoyed great success in computational modeling when the two types of scanning are treated as separate processes. Computational models have been developed for the two types of scanning processes separately. In the domain of memory scanning, Sternberg (1969) identified memory scanning as a linear, exhaustive search process that shows a linear relationship between the number of memorized items (called the size of the "positive set" in the literature on memory scanning) and the time needed to make a response regarding whether a displayed item is a member of the positive set. An important finding of Sternberg's study and a large number of subsequent studies is that memory scanning time is determined by the size of the positive set, but not by the position of the target within the positive set. For instance, if a subject memorizes a list of 4 letters, say, "a, b, c, d", as the positive set, then the subject's response time would be the same regardless which of the four letters is presented as the target. It appears as if the subject has to perform an exhaustive memory search process to compare the presented target with all the items in her working memory before making a decision. This exhaustive search hypothesis was further supported by the finding that the response time for making a "yes" response to indicate that a target was found was the same as that for making a "no" response when the target did not match any of the positive set items. For instance, if a negative set item such as the letter "e" is presented to the subject, in the illustrative experiment mentioned above, then the subject's response time for making a "no" response would be the same as her response time when one of the four positive items is presented.

In the domain of visual scanning, two representative experimental paradigms have been developed--structured visual search and unstructured visual search. In a typical structured search experiment, visual stimuli are presented as an organized array of items in the visual field, and the subjects are required to search the stimulus array according to some strict serial order (e.g., from left to right, or from top to bottom). The focus of investigation is the relationship between the time needed to find a target and the serial position of the target in the stimulus array. This type of visual search is characteristic of real-world tasks such as reading a panel of displays in strict sequence or reading a name list. One of the first and representative studies of this experimental paradigm was the letter-search paradigm (Neisser, 1963), which required the subjects to search a letter list sequentially to find a target letter. An important finding was that there was a linear relationship between the serial position of the target letter in a letter-list and the time needed to find that letter. This linear relationship suggests that subjects stops further visual search as soon as they find a target. Since then, the process of structured visual search through an organized array of items has been modeled as a linear, self-terminating scanning process.

In a typical unstructured search experiment, in contrast, visual stimuli are usually scattered randomly on the display, and the subjects are not required to follow a strict sequence in searching the display. The primary research interest is how the probability of detecting a target changes as a function of time available for search. This type of visual search is typical of real-world tasks such as quality inspection to detect product flaws. A pioneering study of this paradigm is that of Drury's study on inspection of sheet metal (1975), which showed that the probability of detecting a target will increase if more search time is given. However, this probability will increase at a diminishing rate rather than as a linear function of time available.

Although there has been a great success in modeling memory scanning and visual scanning when the two types of scanning are treated as separate processes, there has been a substantial gap between studies of memory scanning and studies of visual scanning. Models of memory scanning and models of visual scanning have been proposed and investigated from the two perspectives, and have generally been separately applied to the two aspects of human cognition and performance. In a typical memory scanning study, visual stimuli are presented at a fixed location, and thus visual scanning is not needed to perform the task. In a typical visual scanning study, in contrast, there is one target that is to be searched for, and thus memory scanning is not involved. It is not clear if the quantitative relationships observed in these studies of memory scanning and studies of visual scanning would still be valid when both types of scanning are involved at once.

Since memory scanning and visual scanning are often intertwined with each other and simultaneously involved in real-world tasks, as illustrated in the process control task described above, it is important to address the relationship between memory scanning and visual scanning in the context of complex task performance, and to bridge the gap between the two domains of human performance modeling. The objective of the current study was to address this issue by examining the strengths and limitations of an integrative computational model of memory scanning and visual scanning that can be derived from the linear model of memory scanning and the linear model of structured visual scanning.

In both the linear model of memory scanning and the linear model of structured visual scanning, time is the dependent variable. The models predict how response time changes as a function of memory scanning demand or visual scanning demand, which are task-related independent variables. In the model of unstructured visual scanning, time serves as the independent variable, which is used to predict the dependent variable of detection accuracy. The objective of the present study was to integrate the two linear models of response time in an effort to predict how response time changes as a joint function of memory scanning demand and visual scanning demand. Thus, the present study focuses on structured visual scanning. For the ease of description, in the following discussions the term visual scanning is used to refer to structured visual scanning, unless otherwise noted.

It should be noted that several researchers have studied the issue of visual search of multiple targets simultaneously. Prominent among those studies include the experimental paradigm of Schneider and Shiffrin's (1977) and the modeling work of Drury and his colleagues (Arani, Karwan, and Drury, 1984; Morawski, Drury, and Karwan, 1980). However, those studies differ from the present one substantially in several aspects.

Started from the pioneering studies of Schneider and Shiffrin (1977), a large body of literature has been accumulated that examined the relationship between memory scanning and visual search of multiple targets when the visually presented targets are not widely separated in space. A critical feature that distinguishes Schneider and Shiffrin's studies from the present one is that the visual displays in Schneider and Shiffrin's type of visual search tasks do not involve a large visual angle and thus sequential visual scanning is generally not needed to locate the visual targets. Consequently, research within that paradigm does not address the issue of eye movement and sequential visual scanning, which is one of the focus of investigation for the current study. The second major difference is that the current study has an applied motive of developing an approximative and computational model that integrates two isolated linear models. The goal of adopting this computational approach was to help human-machine interface designers when an estimate of performance time in similar task environments is needed. Numerous authors have pointed out that the use of computational models of human performance in system design has become increasingly important as the cost and time required to perform full-scale simulations have increased (Baron, Kruser and Huey, 1990; Elkind, Card, Hochberg and Huey, 1990; McMillan, Beevis, Salas, Strub, Sutton and Van Breda, 1989).

Drury and his colleagues have derived computational models for visual search of multiple targets (see, e.g., Arani, Karwan, and Drury, 1984). As discussed earlier, their focus of investigation differed from the current one in at least two ways. First, their interest was on unstructured rather than structured visual search, and second, the primary purpose of their models was to predict how detection probability changes as a function of time available, rather than how detection time changes as a function of memory and visual scanning demands.

A Computational Model of Memory Scanning and Visual Scanning

According to the results of a large number of studies on memory scanning, when visual scanning is not involved, memory scanning time can be modeled as

$$\text{Memory scanning time} = (a + b \times S) \tag{1}$$

where S is the size of the positive set, a is the time required to initiate memory scanning, and b is the increment in memory scanning time for each additional positive set item (Sternberg, 1969, 1975).

Studies of visual scanning have shown that when the size of the positive set in short-term memory is one, visual scanning time can be modeled as

$$\text{Visual scanning time} = (c + d) \times N \tag{2}$$

where N is the number of eye movements involved, c is the time to initiate an eye movement, and d is the time to perform the actual movement (saccade). Converging evidence from eye movement studies has demonstrated that c has a value of about 220-250 msec, and d has a value of about 30 msec for a 5 degree saccade with each additional degree increase in magnitude requiring about another 2 msec (Bartz, 1962; Brogan, 1990). Since increase in saccade magnitude produces a very small increase in eye movement time, visual scanning time can be approximated as:

$$\text{Visual scanning time} = A \times N \tag{3}$$

where A (= c+d) is the total time to complete an eye movement cycle.

In a process monitoring task involving both memory scanning and visual scanning, a subject needs to search through several physically separated displays to check whether any of the displays carries any of the memorized items (also called targets in the following description) in her working memory. For the ease of description, suppose that there is a total of M displays and the subject found a target in the Nth display she examined ($M \geq N$), but not earlier, then the subject would have completed N iterations of memory scanning, each of which was preceded by an eye movement cycle. In other words, the subject needed to complete a series of N eye movement-memory scanning alternations before finding the target at the Nth display. Because visual scanning is a self-terminating process, there is no reason for the subject to continue to search the rest of the (M-N) displays. Thus the subject's response time is the sum of the time to complete N visual scanning cycles and the time to complete N iterations of memory scanning, that is,

$$\text{Response time} = A \times N + (a + b \times S) \times N, \quad N > 0 \qquad (4)$$

where

$A \times N$ is the time to complete N eye movement cycles,

$(a + b \times S) \times N$ is the time to complete N iterations of memory scanning with S as the size of the positive set.

If the subject found the target in the last of the M displays (i.e., the Mth display) she examined or if none of the M displays contained a target, then the subject would have to complete a series of M eye movement-memory scanning alternations before making a response. But response time under these situations can be analyzed similarly with equation (4) by simply substituting N with M, without requiring any change in the form of the equation.

The critical issue of investigation for the present study was how well this derived equation would be able to predict the joint effects of visual scanning and memory scanning on response time. Two experiments were conducted to address this question. The experiments simulated an operator's task of monitoring an array of clock-like instrument meters to detect if any of them indicates a system error. More specifically, the experimental tasks required the subjects to keep a number of items in their working memory (the definition of errors, also called targets or the positive set) and search through an organized array of clock-like displays to decide whether any of them carried any of the memorized items. As described in detail below, the number of memorized items (i.e., the size of the positive set for memory scanning: S), and the demand for visual scanning (i.e., the number of circles the subjects have to search before making a decision: N) were the two independent variables employed in both experiments.

The difference between the two experiments was that the first experiment employed a different positive set of memorized items for each experimental trial, whereas the second experiment employed the same positive set for all the trials that had the same positive set size. In other words, the definition of system errors varied from trial to trial in the first experiment, but was kept the same for all the trials of similar sessions in the second experiment. Response time and response accuracy were collected as dependent measures for both experiments.

2. Experiment 1

The experiment required the subjects to search through an organized array of circles to decide whether any of the circles carried any of the memorized items in their working memory. The joint effects of two independent variables were examined: the number of memorized items (i.e., the size of the positive set for memory scanning: S), and the demand for visual scanning (i.e., the number of circles that the subjects have to search through before making a decision: N).

2.1. Methods

2.1.1. Stimuli and Memory Scanning: Using the terminology adopted by Sternberg (1975), the set of all the items that might appear as test stimuli is defined as the stimulus ensemble. The stimulus ensemble in this experiment consisted of 16 vectors emanating from the center of a circle. The angular positions of the vectors were at 0°, 22.5°, 45°, ..., 315°, 337.5°, respectively. In other words, the 16 vectors were at positions that divided a circle into 16 equal sections. From among the stimulus ensemble a set of elements is selected arbitrarily and randomly and is defined as the positive set. The remaining items are called the negative set. In each trial of the memory scanning task, the subjects were first presented with a positive set, and then presented with an organized array of circles which carried test items (which are drawn from the stimulus ensemble but may or may not include a member of the positive set--see the following section about how the test items were presented). The subjects were required to make a binary response regarding whether any of the displayed items was a member of the positive set. The size of the positive set was varied at three levels, corresponding to S=1, 2, and 4, respectively, where S is the size of the positive set.

2.1.2. Task Display and Visual Scanning: The task display consisted of eight information display circles and a fixation point. For explanatory purpose, the eight information circles are named circles C1 through C8, respectively, as shown in Figure 1. The names of the circles were not displayed to the subjects. The fixation point was where the subjects were required to look at immediately before the start of each trial, and it was located to the left of circle C1. The visual angle subtending the center of circle C1 and the fixation point was about 10 degrees, which was the same as the visual angle subtending the centers of any two adjacent circles. A chin-rest was used to prevent subjects from making head movements while they were performing the tasks. The display was selected according to results from previous studies on visual scanning (Liu, 1995; Liu and Wickens, 1992; Wickens and Liu, 1988) and observations from testing pilot subjects to ensure that critical details of the task stimuli could not be resolved in peripheral vision. Therefore, visual scanning was necessary to accomplish the tasks.

In a random half of the trials within each experimental block, exactly one positive item would appear, and it could appear in any of the eight circles with equal probability, and thus required a positive response from the subjects. For the ease of description, this type of experimental trials are called "positive set trials". In the remaining half of the trials of each block, none of the circles carried a positive item, and thus required a negative response. These trials are called "negative set trials" in the following descriptions. In the experimental sessions in which visual scanning was involved, the subjects were instructed to search the 8 circles strictly in the order of C1 to C8 to look for a potential positive item. For "positive set trials", the position of the circle that carried a positive item determined the demand for visual scanning. Since there were eight circles, visual scanning demand was varied at eight levels, corresponding to the situation in which the positive item appeared in circle C1, or C2, ..., or C8, respectively. For explanatory purpose, the eight levels of visual scanning demands are referred to as N=1, N=2, ..., and N=8, respectively, where N is the number of circles the subjects had to search before making a decision response. For the "negative set trials" in which none of the circles carried a positive set item, the subjects had to search through all the eight circles before making a response, and thus its visual scanning demand was approximately equivalent to that in the N=8 condition in a "positive set trial".

Insert Figure 1 about Here

2.1.3. <u>Design and Procedure</u>: The experiment consisted of two training sessions, followed by six experimental sessions and two no-scanning baseline sessions. Each session included 3 blocks of 16 trials each. Each of the 3 blocks within a session involved a different positive set size (S=1, 2, 4). Subjects were presented with a different positive set on each trial. At the start of each trial, the positive set was first presented to the subject with a display duration proportional to the size of the positive set (More specifically, display duration = 1.5 sec × size of the positive set). The positive set was followed by a display of the word "READY", which reminded the subjects to fixate at the fixation point. Five seconds after the display of the ready signal, the word "GO" was displayed to ask the subjects to start the visual scanning and memory scanning process. In the experimental sessions in which visual scanning was involved, the subjects were instructed to search the 8 circles strictly in the order of C1 to C8 to look for a potential positive item. In the no-scanning baseline sessions the subjects were required to fixate at circle C1 and decide if C1 carries any of the positive items, which appeared in a random half of the trials. Subjects' response time and response error rate were recorded on a trial-by-trial basis for data analysis. The order of the sessions and the order of the trials within a session were randomly assigned to each subject. The tasks were implemented on an Macintosh IIci computer with a 21 inch monitor, and were conducted in a sound-attenuated experimental room.

2.1.4. Subjects: Twelve right-handed University of Michigan subjects were recruited as subjects and paid for their participation in the experiment. All subjects had normal or corrected to normal vision.

2.2. Results and discussion

The trial-by-trial computer recordings of subjects' response time and response accuracy were first grouped according to the experimental conditions they belong to and then analyzed to determine subjects' performance under the 3 baseline conditions and the 24 experimental conditions. The 3 baseline conditions were defined by the three levels of memory set size when visual scanning was not required for performing the tasks, and the 24 experimental conditions were determined by the cross-multiplication of the three levels of memory set size and the eight levels of visual scanning demand.

2.2.1. Response time data: Response time data for the 3 baseline conditions are plotted as the three data points on the ordinate in Figure 2. This set of data points were collected under no-scanning conditions, which were identical to those in a typical experiment of the Sternberg memory scanning paradigm. The 3 data points represent response time of memory scanning of 1, 2, and 4 items, respectively. Parameters in equation (1) can be estimated by performing a linear regression of the three data points on memory set size, which resulted in the following regression equation:

$$\text{Memory scanning time} = 483 + 77 \times S \qquad (5)$$

The two regression coefficients in Equation (5) provided a value of 483 msec for the parameter "a" and a value of 77 msec for the parameter "b" in equation (1) for the current task.

Response time data for the 24 experimental conditions that involved visual scanning are plotted in Figure 2 as a function of visual scanning demand (on the abscissa) and memory demand (the group of dashed-lines within the panel). In order to estimate the value of parameter "A" in equation (3), we can compare the data point corresponding to ($S=1$, $N=0$) and that corresponding to ($S=1$, $N=1$). The two data points are the two left-most data points on the dashed-line at the bottom of Figure 2. Comparing these two data points revealed that response time in the $N=1$ condition (requiring 1 cycle of eye movement) was 332 msec longer than in the $N=0$ condition (no visual scanning), when both conditions involved the same memory set size of 1. In other words, it took approximately 332 msec to complete an eye movement cycle in the simplest experimental condition of the current experiment. This value of 332 msec provided an estimate for parameter "A" in Equation (3).

Substituting these values of a, b, and A into equation (4), which was derived on the basis of equations (1) and (3), we have the following equation for predicting response time in other experimental conditions:

$$\begin{aligned}\text{Response time} &= 332 \times N + (483 + 77 \times S) \times N \\ &= 815 \times N + 77 \times S \times N \qquad (6)\end{aligned}$$

The critical issue of investigation for the present study was whether this equation was able to predict the experimental data collected in other conditions.

Insert Figure 2 about Here

In order to compare the predicted and the observed data, the response time data predicted by Equation (6) are plotted in Figure 2 as the group of solid lines. A close inspection of the relationship between the predicted and the observed data (the dashed lines) in Figure 2 suggests that there appeared to be a close fit between the two sets of data when the task did not have a high memory scanning demand and a high visual scanning demand simultaneously. In other words, the predicted and the observed data appeared to be close to each other when the product of N and S did not have a large value.

This speculation based on visual inspection of the data was confirmed by further data analysis, which revealed that the data predicted by Equation (6) were able to approximate the observed data when $N \times S \leq 8$. When $N \times S > 8$, Equation (6) underestimates response time. Another term must be added to the predictive equation to compensate for this underestimation, which was found to be approximately a linear function of $(N \times S - 8)$. More specifically, the following equation was found to be a good approximation to the observed data:

$$\text{Response time} = 332 \times N + (483 + 77 \times S) \times N + \max\{0, 75(N \times S - 8)\}$$

$$= 815 \times N + 77 \times S \times N + \max\{0, 75(N \times S - 8)\} \qquad (7)$$

In the above equation, the first two terms are the same as those in equation (6), which is based on predictions of the existing linear models of memory scanning and visual scanning, and the third term has a value of 0 when $N \times S <= 8$

Response time data for the 24 experimental conditions were also subjected to a 3 × 8 repeated-measures analysis of variance (ANOVA). Significant main effects on response time were found for both the factor of memory scanning ($F(2, 11)=17.35$, $p < 0.001$) and the factor of visual scanning ($F(7, 11)=23.22$, $p < 0.001$). There was also a significant two-way interaction between the two variables ($F(14, 11)=9.83$, $p<0.001$). These results are consistent with the predictions of equations (6) and (7), according to which an increase in either N or S would produce a corresponding increase in response time, as demonstrated by the main effects of the two variables. Furthermore, the multiplicative term in equations (6) and (7) suggests that the effect of either variable (N or S) on response time would be greater if the other variable is at a higher level, as shown by the observed two-way interaction between the two variables.

2.2.2. Response error: Response error data for the 24 experimental condition are plotted in Figure 3 as a function of visual scanning demand (on the abscissa) and memory demand (the group of lines within the panel). The 3 no-scanning baseline data points are plotted on the ordinate. Response error data for the 24 experimental conditions were subjected to a 3 × 8 repeated-measures ANOVA, which revealed significant main effects on response error for both the factor of memory scanning ($F(2,11)=11.66$, $p<0.01$) and the factor of visual scanning ($F(7,11)=6.23$, $p<0.05$). There was also a significant two-way interaction between the two variables

$(F(14,11)=3.82, p<0.05)$.

Further analysis of the data suggests that this two-way interaction can be explained by the differential effects of visual scanning when memory scanning was at different levels. More specifically, when memory scanning demand was at the level of $S=1$, increases in visual scanning demand had no effect on response error (all ps > 0.10 for the 7 paired-comparisons between data at adjacent levels of visual scanning demand). When $S=2$, visual scanning produced a significant effect on response error only when the level of visual scanning demand was sufficiently high (when $N\geq 7$) (all ps > 0.10 for the 5 paired-comparisons between adjacent data points with $N\leq 6$, and both ps < 0.05 for the paired-comparisons between $N=6$ and $N=7$ and that between $N=7$ and $N=8$). However, when memory scanning demand was at the level of $S=4$, almost every increase in visual scanning demand resulted in an increase in response error (all ps <0.05 for the 6 paired-comparisons between adjacent data points with $N>=2$), with the difference between $N=1$ and $N=2$ as the exception ($p>0.10$ for the paired-comparison between the two data points).

Insert Figure 3 about Here

Since the purpose of the present study was to test the predictive power of a response time model, response time data is of primary interest. The main purpose of collecting and analyzing response error data was to check whether subjects adopted different speed-accuracy tradeoff strategies in different task conditions. It is important to ensure that longer response times were caused by higher processing demands rather than by a shift in subjects' speed-accuracy tradeoff strategies in certain experimental conditions to sacrifice speed for achieving lower error rates (Wickelgren, 1977). The above analysis on response error found no evidence of such shifts in the subjects' speed-accuracy strategy, because the experimental conditions with longer response times in the present data also produced similar or higher, rather than lower, error rates.

In summary, a particularly interesting finding of this experiment was that Equation (4), which was derived a priori based on an existing linear model of memory scanning and an existing linear model of visual scanning, needs to be modified in order to approximate the joint effects of memory scanning and visual scanning demands on response time closely. The modified version, as presented in equation (7), contains an additional term, whose effect on response time was manifest only when the joint demand of memory scanning and visual scanning was sufficiently high, and this effect increased approximately linearly as a function of this joint demand.

Experiment 1 employed a varied-set procedure, in which the positive set varied from trial to trial. The subjects reported that they tended to forget part of the positive set if they were not able to find a positive item in the first few circles of the search sequence. A long sequence of successive comparisons between visually displayed vectors and memorized vectors tended to interfere with their memory. Experiment 2 extends the scope of investigation by employing a fixed-set procedure, in which the same positive set was used for the blocks of trials involving the same set size. An important question is how the subjects' performance would differ from that observed with the varied-set procedure.

3. Experiment 2

3.1. Method

The experimental method was basically the same as that of Experiment 1. The only difference between the two experiments was that Experiment 1 employed a different positive set of working memory items for each experimental trial (called varied-set procedure), whereas Experiment 2 employed the same positive set for all the trials that involved the same positive set size (called fixed-set procedure). In other words, the definition of system errors varied from trial to trial in Experiment 1, but was kept the same for all the trials that had the same positive set size in Experiment 2.

All the other aspects of the experimental methods, including definition of stimulus ensemble, task display, visual scanning requirements, and the experimental design and procedure, were the same as in Experiment 1. Twelve right-handed University of Michigan students (six men and six women) were recruited as subjects and paid for their participation in the experiment. All subjects had normal or corrected to normal vision. None of the subjects had participated in the first experiment.

3.2. Results and discussion

Similar to Experiment 1, the trial-by-trial computer recordings of subjects' response time and response accuracy were analyzed to determine their performance under the 3 baseline conditions in which visual scanning was not involved in performing the tasks and under the 24 experimental conditions when visual scanning was involved. The 3 baseline conditions were defined by the three levels of memory set size with no visual scanning, and the 24 experimental conditions were determined by the cross-multiplication of the three levels of memory set size and the eight levels of visual scanning demand.

3.2.1. Response time data: Response time data for the 24 experimental conditions that involved visual scanning are plotted in Figure 4 as a function of visual scanning demand (on the abscissa) and memory demand (the group of lines within the panel). The 3 no-scanning baseline data points are plotted on the ordinate.

Similar to Experiment 1, parameters in Eq. (4) were estimated by examining the data points in Figure 3. The set of 3 points on the ordinate corresponded to the response time of memory scanning of 1, 2, and 4 items respectively when visual scanning was not required. A linear regression of these three data points on memory set size resulted in the following regression equation:

$$\text{Memory scanning time} = 474 + 71 \times S \tag{8}$$

The pair of regression coefficients in Equation (8) provided a value of 474 msec for the parameter "a" and a value of 71 msec for "b" in equation (1) for the current task.

The data point corresponding to (N=0 and S=1) and that corresponding to (N=1 and S=1) were compared to estimate the value of "A" in equation (3). The two data points are the two left-most data points on the bottom line in Figure 4. This comparison revealed that response time in the N=1 condition was 316 msec longer than in the N=0 condition, when both conditions involved the same memory set size of 1. In other words, it took approximately 316 msec to complete an eye movement cycle in the simplest experimental condition of Experiment 2. This value of 316 msec provided an estimate for parameter "A" in Equation (3).

Substituting these values of a, b, and A into equation (4), we have:

$$\begin{aligned}\text{Response time} &= 316 \times N + (474 + 71 \times S) \times N \\ &= 790 \times N + 71 \times S \times N \end{aligned} \tag{9}$$

In order to compare the observed and the predicted data, the response time data predicted by Equation (9) are plotted in Figure 4 as the group of solid lines. Data analysis revealed that the following equation was able to approximate the observed data closely for practical purposes:

$$\text{Response time} = 790 \times N + 71 \times S \times N + \max\{0, 25(S \times N - 8)\} \tag{10}$$

Because the last term in equation (10) is negligible in most conditions, equation (9) was almost equally good in approximating the observed data.

Insert Figure 4 about Here

Response time data for the 24 experimental conditions were also subjected to a 3 × 8 repeated-measures analysis of variance (ANOVA). Significant main effects on response time were found for both the factor of memory scanning (F(2, 11)=12.51, $p < 0.01$) and the factor of visual scanning (F(7, 11)=9.68, $p < 0.001$). There was also a significant two-way interaction between the two variables (F(14, 11)=5.85, $p<0.01$). Similar to Experiment 1, the main effects of the two variables can be explained by equation (9), according to which an increase in either N or S would produce a corresponding increase in response time. The observed two-way interaction can be explained by the multiplicative term in equation (9), which suggests that the effect of either variable on response time would be greater if the other variable is at a higher, rather than a lower, level.

3.2.2. Response error: Response error data for the 24 experimental conditions are plotted in Figure 5 as a function of visual scanning demand (on the abscissa) and memory demand (the group of lines within the panel), and they were subjected to a 3 × 8 repeated-measures ANOVA. The result of analysis revealed a significant main effect on response error for the factor of memory scanning (F(2,11)=6.38, $p<0.05$). Response error was not influenced by visual scanning, nor did this variable interact with memory scanning (both ps <0.05). Further analysis of the data indicates that error rate was higher when S=4 than when S=1 or S=2, while the latter two conditions showed no differences between themselves (F(1, 11) =2.47, $p>0.10$ for the data set that does not include the S=4 conditions).

Insert Figure 5 about Here

There are two major aspects of the data of Experiment 2 that differ from those of Experiment 1. First, although an additional term must be added to equation (4) in order to account for the response time data of Experiment 1 employing a varied-set procedure, this additional term was basically not necessary for the data of Experiment 2, which employed a fixed-set procedure. Second, although visual scanning demand had a significant effect on response accuracy in Experiment 1 when memory scanning demand was high (when S=4), there was no evidence of a similar effect of visual scanning in Experiment 2. The implications of these results are discussed in the following section.

4. General discussion

A recent report of the Committee on Human Factors of the National Research Council indicated the need for developing comprehensive and computational models of complex human performance (Baron, Kruser, and Huey, 1990). As indicated in the report, although a wide variety of models that focus on some particular aspects of human performance have been developed, the utility of these models in engineering design would increase if more comprehensive models can be developed by integrating existing models. The report pointed out that "models that rely on integrating various submodels should receive the most attention....A gradual extension or aggregation of well-validated models to deal with new or compound situations is recommended" (Baron, Kruser, and Huey, 1990; p.76). The research reported in this article is a step toward that direction. More specifically, this study was conducted to integrate Sternberg's (1969) linear model of memory scanning and Neisser's (1963) linear model of visual scanning.

As described in the introduction section, when an operator such as a process controller carries several items in her working memory and there is only one item being presented in a visual display, Sternberg's (1969) linear model of memory scanning can be used to estimate the amount of time needed by the operator to decide whether the visually displayed item matches any of the items in her working memory. In contrast, when an operator carries only one item in her working memory but the visual display consists of an organized array of widely-separated items, Neisser's linear model of visual scanning is a useful device for estimating the amount of time needed by the operator to find the memorized item in the display through a structured visual search.

An integrated model was derived in the present study to consider the effects of memory scanning and visual scanning simultaneously, and the model attempts to answer the following question: When there are multiple items both in the operator's working memory and in the visual display, how much time does it take for an operator to decide whether any of the displayed items matches any of the items in her working memory if the operator adopts a structured sequential visual search strategy? As described in the introduction section, this type of tasks are characteristic of many real world tasks. The derived model was shown in equation (4), and the predictive power of the model was tested through two experiments.

A salient finding of the two experiments is that the model was able to predict observed data under some, but not all, of the experimental conditions. More specifically, in order to predict subjects' response time, an additional term must be added to the derived model when the items in the subjects' working memory (the positive set) varied from trial to trial (the varied-set procedure), but the additional term played a negligible role when the task did not impose high demands on memory and visual scanning simultaneously or when the positive set was fixed across the trials of the same experimental condition (a fixed set procedure). These experimental findings can be summarized in the following equation:

$$\text{Response time} = A \times N + (a + b \times S) \times N + \max\{0, C \times (N \times S - 8)\}, N > 0 \quad (11)$$

In equation (11), the first two terms constitute equation (4) and are derived on the basis of the existing linear models of memory scanning and visual scanning; the third term is negligible when $N \times S \leq 8$ (the joint demand for memory and visual scanning is not sufficiently high) and when C has a small value (as observed in the fixed-set procedure).

In both experiments, subjects made more errors when they had to memorize a larger positive set than when they were required to memorize a smaller set. But the effects of visual scanning demand on error rates were found only in Experiment 1 in which the varied-set procedure was used. A significant two-way interaction between memory scanning demand and visual scanning demand was observed in the response error data in Experiment 1, but not in Experiment 2.

Subjects reported that when the positive set varied from trial to trial, they tended to forget part of the positive set if they had to remember a large positive set and if a positive item was not found in the first few circles of the search sequence. A long sequence of successive comparisons between visually displayed vectors and memorized vectors tended to disrupt their memory, when they had to remember a new positive set for each trial. Subjects did not report this type of interference when the fixed-set procedure was used. This appears to offer an explanation to the aforementioned finding that an additional term was needed to account for the response time data when the joint demand of memory scanning and visual scanning was sufficiently high, and that the joint effects of the two variables on response time and response error were much more pronounced in the varied-set procedure

(Experiment 1) than in the fixed-set procedure (Experiment 2).

In summary, equation (11) appears to offer a potentially useful tool to interface designers for estimating response time in similar task situations. Furthermore, the current results suggest that integrative models derived from existing componential models can provide close approximations to complex task performance only in certain task situations. The integrative models must be evaluated to reveal their strengths and limitations and to identify their boundary conditions for applications. Considerations of additional behavioral and task-related factors may become necessary for making more accurate data predictions.

Acknowledgments

This research was supported by a Faculty research grant from the University of Michigan and by a grant from the Intelligent Transportation Systems Research Center of Excellence. The author acknowledges Richard Gong, Kristen Korzecke, Kelli Skinner and Kirk Ziehm for their assistance in computer programming, data collection and analysis. Correspondence concerning this article should be addressed to Prof. Yili Liu, Department of Industrial and Operations Engineering and Center for Ergonomics, University of Michigan, Ann Arbor, MI48109, U.S.A.

References

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

Baron, S., Kruser, D. and Huey, B. (Eds.) 1990, *Quantitative modeling of human performance in complex, dynamic systems*, (National Academy Press, Washington, D.C.).

Bartz, A. E. 1962, Eye movement latency, duration and response time as a function of angular displacement, *Journal of Experimental Psychology*, **64**, 318-342.

Brogan, D. (Ed.) 1990, *Visual search*, (Taylor and Francis, London).

Drury, C.G. 1975, Inspection of sheet metals--Model and data, *Human Factors*, **17**, 257-265.

Elkind, J., Card, S. Hochberg, J. and Huey, B. 1990, *Human performance models for computer-aided engineering*, (Academic Press, Boston, MA)

Liu, Y. 1995, Quantitative assessment of effects of visual scanning on concurrent task performance, to appear in *Ergonomics*.

Liu, Y. and Wickens., C. 1992, Visual scanning with or without spatial uncertainty and divided and selective attention, *Acta Psychologica*, **79**, 131-153.

McMillan, G., Beevis, D., Salas, E., Strub, M., Sutton, R. and Van Breda, L. 1989, *Applications of human performance models to system design*, (Plenum Press, New York).

Morawski, T., Drury, C.G. and Karwan, M.H. 1980, Predicting search performance for multiple targets, *Human Factors*, **22**, 707-718.

Neisser, U. 1963, Decision time without reaction time, *American Journal of Psychology*, **76**, 376-385.

Schneider, W. and Shiffrin, R. 1977, Controlled and automatic human information processing: I. Detection, search and attention, *Psychological Review*, **34**, 1-66.

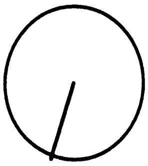
Sternberg, S. 1969, Memory scanning: Mental processes revealed by reaction-time experiments, *American Scientist*, **57**, 421-457.

Sternberg, S. 1975, Memory scanning: New findings and current controversies, *Quarterly Journal of Experimental Psychology*, **27**, 1-32.

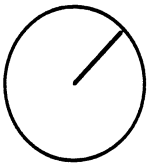
Wickelgren, W. 1977, Speed accuracy tradeoff and information processing dynamics, *Acta Psychologica*, **41**, 67-85.

Wickens, C. D. and Liu, Y. 1988, Codes and modalities in multiple resources: A success and a qualification, *Human Factors*, **30**, 599-616.

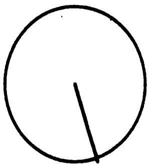
List of Figures



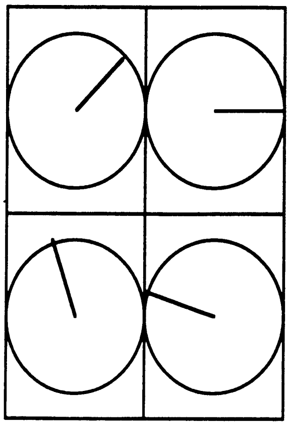
C2

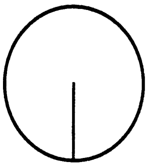


C3

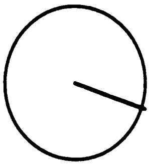


C4

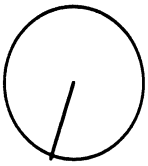




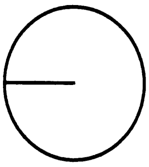
C1



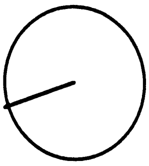
C5



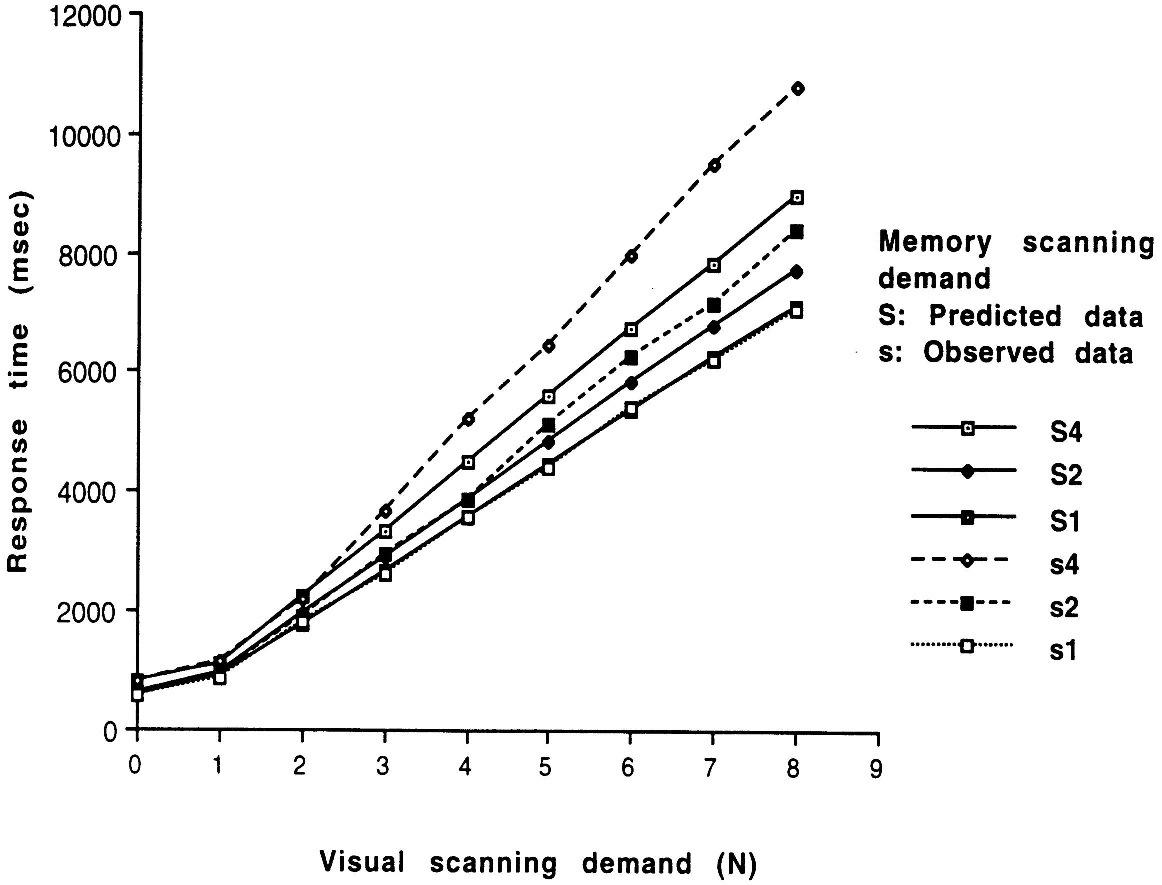
C8

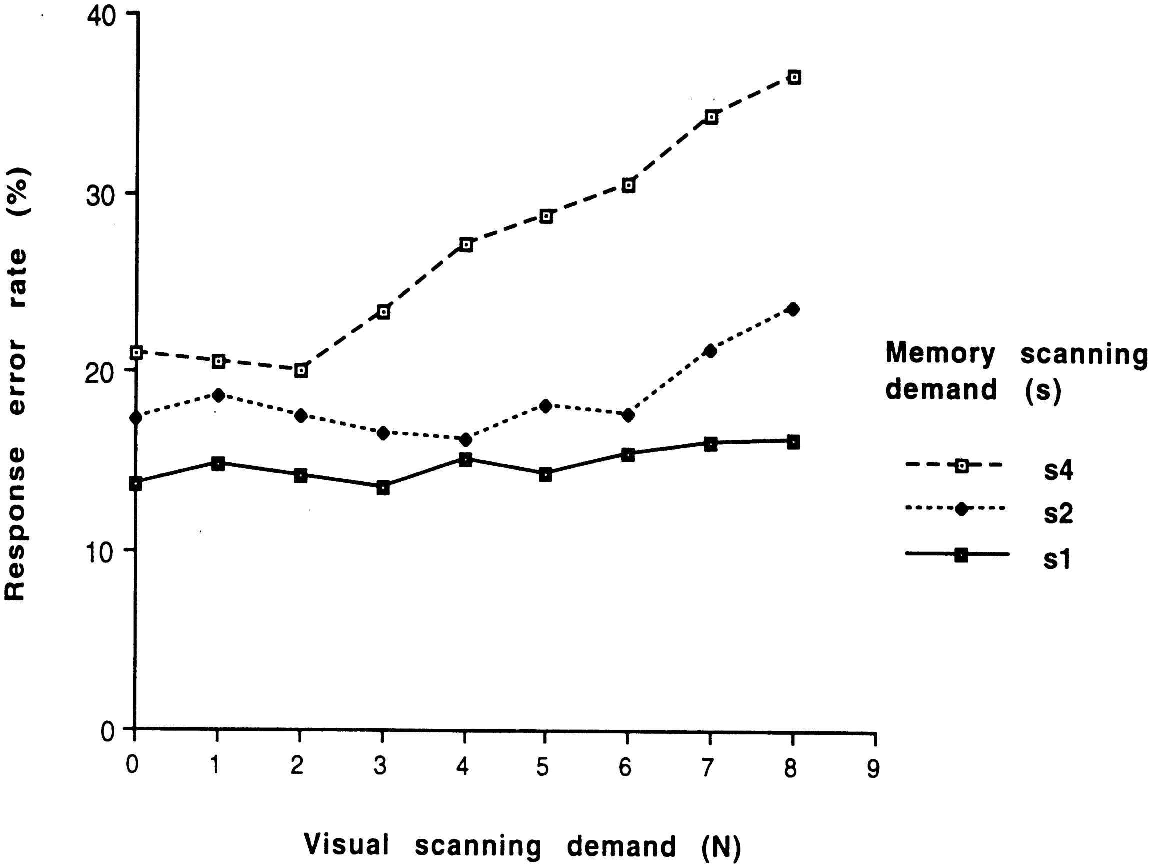


C7

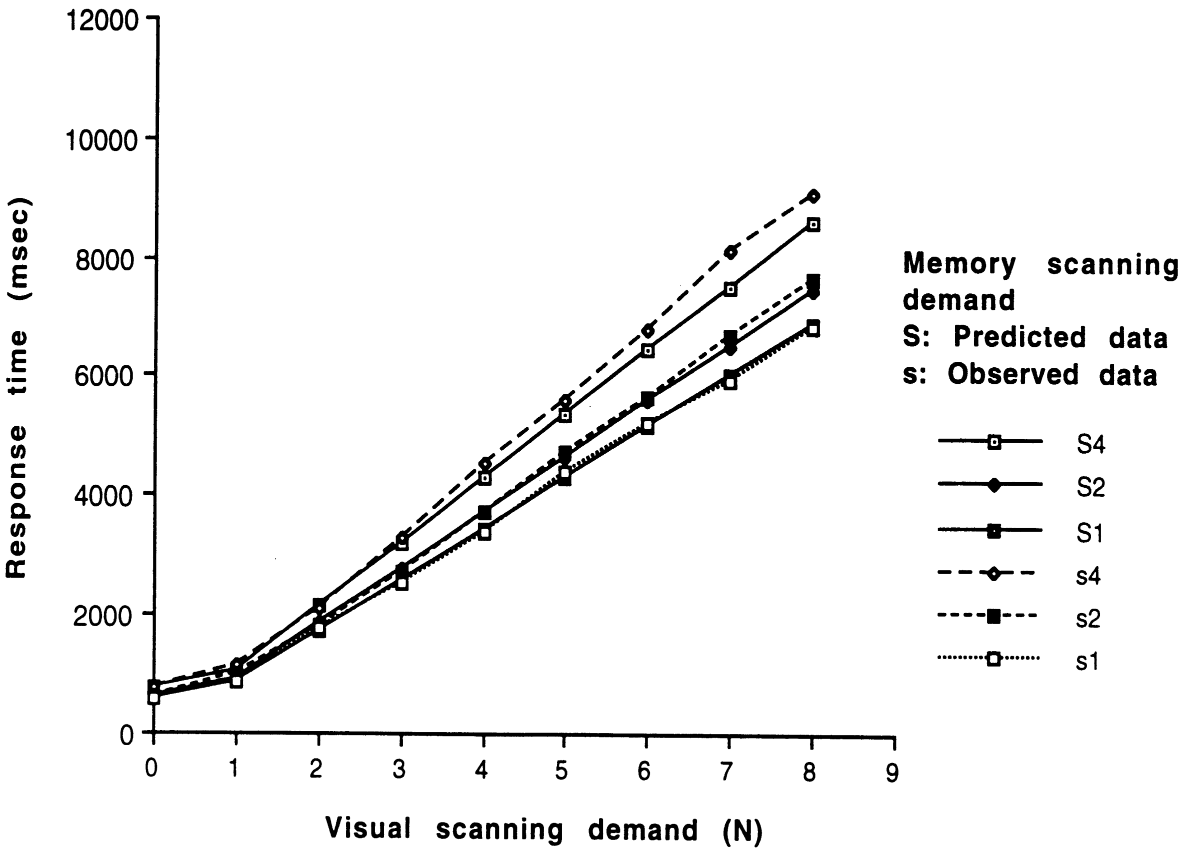


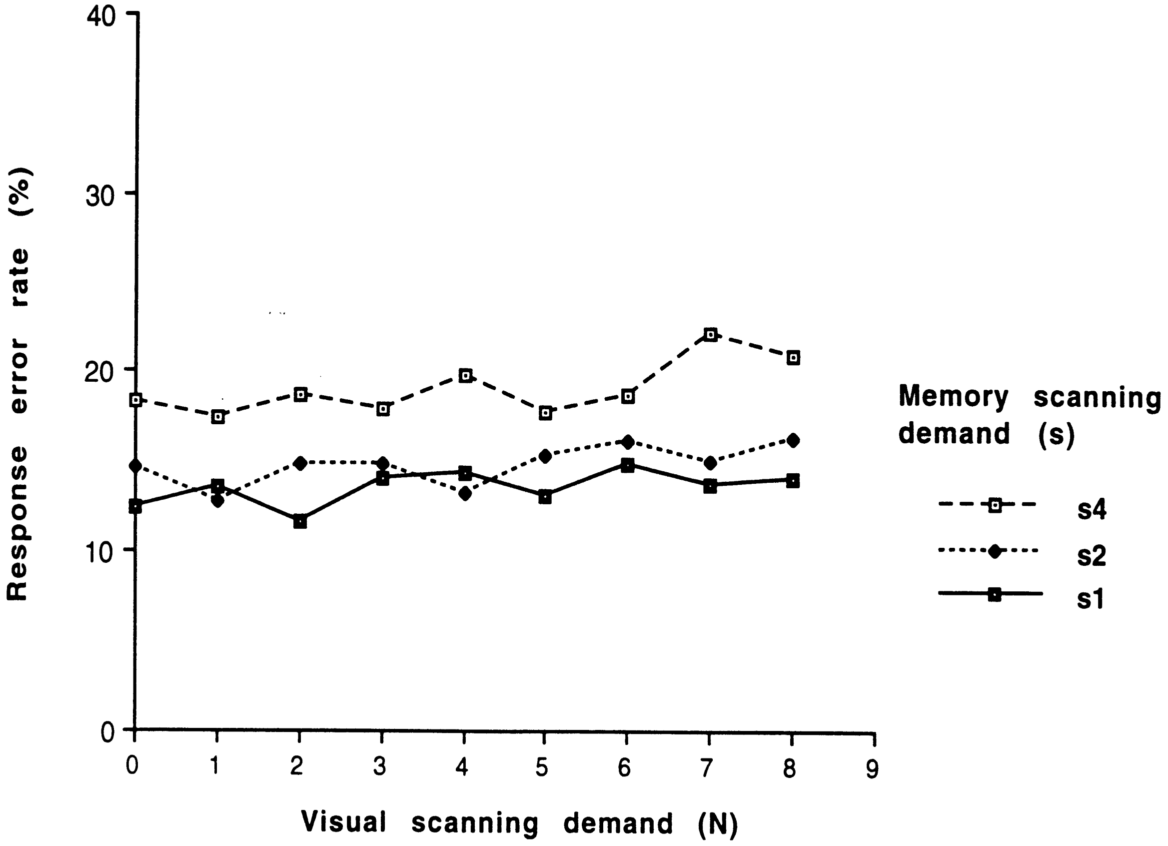
C6


12000
10000
8000
6000
4000
2000
0
Response time (msec)
0
1
2
3
4
5
6
7
8
9
Visual scanning demand (N)
Memory scanning demand
S: Predicted data
s: Observed data
S4
S2
S1
s4
s2
s1



Response error rate (%)
40
30
20
10
0
0
1
2
3
4
5
6
7
8
9
Visual scanning demand (N)
Memory scanning demand (s)
s4
s2
s1


12000
10000
8000
6000
4000
2000
0
Response time (msec)
0
1
2
3
4
5
6
7
8
9
Visual scanning demand (N)
Memory scanning demand
S: Predicted data
s: Observed data
S4
S2
S1
s4
s2
s1



Response error rate (%)
40
30
20
10
0
0
1
2
3
4
5
6
7
8
9
Visual scanning demand (N)
Memory scanning demand (s)
s4
s2
s1