Perception & Psychophysics
1983, 33 (3), 271-282

Acquisition and decision in visual same-different search of letter displays

D. C. DONDERI
McGill University, Montreal, Quebec, Canada

From 2 to 23 capital As, Bs, and Cs were positioned randomly over visual displays varying in size from 5 to 10 deg square and in luminance from 7 to 250 cd/m². The task was to decide whether all letters were the same or one was different from the rest. Instructions stressed accuracy, and responses were 97% correct. Three experiments with 50 observers varied amount of practice, number of letters (N), and size and luminance of the display. All experiments produced a linear invariance between mean "same" ($\bar{S}$) and mean "different" ($\bar{D}$) response times in seconds with N as the parameter: $\bar{D} \cong \bar{S}/2+.4$. The data are consistent with Krueger's same-different decision theory, and with the separation of acquisition from decision processes.

In a same-different task, the observer searches the display to decide whether the N stimuli are the same or one is different from the rest. There are no targets presented in advance to be remembered and to be compared with each display. Sometimes response times for this task do not increase with N, for example, with 2 to 14 geometric shapes arranged randomly over the display (Donderi & Case, 1970; Donderi & Zelnicker, 1969).

When visual search or same-different times do not vary as a function of N, then all of the display information is processed "in parallel" by automatic processing. Many search decisions (e.g., Rabbitt, Cumming, & Vyas, 1978) and same-different decisions (e.g., Bindra, Donderi, & Nishisato, 1968) take more time as N increases, but this does not necessarily mean that the stimuli are searched in serial order. If information from all stimuli was processed simultaneously but less efficiently as N increased, then the mean decision latency would increase with N (Snodgrass & Townsend, 1980; Taylor, 1978) and, under simple assumptions about the distribution of decision latencies, might increase as the log of N. Thus, "parallel processing" may occur, although decision latency increases as a function of N. A distinction between serial and parallel processing cannot be based on an increase in mean decision latency alone.

A simple serial processing theory leads to the expectation that, in the same-different task, "different" decisions should be faster than "same" decisions, because a correct "different" decision is self-terminating: one display item different from the rest indicates the "different" response, while all items must be searched in order to respond "same." But, while same-different decision latencies are often constant over N, "same" decisions are sometimes faster than "different" decisions (e.g., Bindra, Donderi, & Nishisato, 1968; Donderi & Zelnicker, 1969).

Krueger (1978) developed a theory of same-different decisions occurring within a single "glance" lasting about 200 msec. His theory predicts faster "same" than "different" decisions. Krueger represents each stimulus by a vector of attribute values. A count of the differences between the values of the two stimulus vectors is recorded, and a same-different decision is made on the basis of the count. There are two separate criteria on the distribution of difference counts: one for a "same" decision and one for a "different" decision. If the difference count falls into the region between the criteria, then another data sample must be taken before a decision can be made. Each sample increases the decision latency by a constant duration. By manipulating the moments of the distribution of difference counts for *same* and *different* stimulus pairs, Krueger modeled the evidence that "same" decisions are often faster than "different" decisions when the decision process takes place within one glance, and also that "different" errors to *same* stimuli predominate.

Krueger's theory concerns one "glance" that lasts for about one fixation, and he proposed that displays that require more than one glance are processed by a serial self-terminating search. Other researchers have been careful to keep their stimuli either within the fovea or equidistant from a central fixation point, so that variations in fixation would not influence decision latency (e.g., Egeth, Jonides, & Wall, 1972). The resulting data concern events that

This research was supported by the Natural Sciences and Engineering Research Council of Canada, Grant AO 108 to the author. I thank Ian Berg, Robert Grimes, and Theresa Tillock for their help as research assistants. Justine Sergent and Bart Farell read earlier versions of the paper, and I thank them for their advice. My mailing address is: Department of Psychology, McGill University, Montreal, Quebec H3A 1B1, Canada.

happen in a fraction of a second over small visual fields.

Visual search in the real world occurs over many targets and takes many seconds and many fixations to complete (Bloomfield, 1973). Teichner and Krebs (1974) postulated two different search processes in multiple-target experiments. The most common process was a one-step search among stimuli that differed on only one attribute, for example, searching for a large circle among small ones or for a triangle among squares. The other process was a multistage search of several attributes within a single stimulus, for example, searching among numbers of various colors for the number 3 or for the color blue. They concluded that successive serial scanning was the best way to explain their simple single-attribute search data. In the experiments they reviewed, the efficiency of search increased up to a maximum of about 50 stimuli/sec when the information load in the entire display was between 50 and 100 bits. They did not review enough data to adequately describe the more complex multistage search process.

If there is a constant latency preceding each response in addition to a variable latency that depends on the number of display elements, then the per-element contribution of the constant latency to the total latency will decrease as the number of elements increases, and the apparent efficiency per element of the decision process will increase as the number of display elements increases. Barber (1981) identified this possible effect of a constant latency as one explanation for Teichner and Kreb's efficiency conclusions.

There is an information gap between data from the single-fixation experiments with few briefly exposed targets (e.g., Egeth, Jonides, & Wall, 1972) and data from the large-field, multiple-target tasks reviewed by Teichner and Krebs. What happens when 2 to 20 moderately complex stimuli are displayed for long durations in both same-different and target-search tasks? The latency to respond "same" or "different" over 2 to 14 geometric shapes under 50- or 200-msec exposures is constant (Donderi & Case, 1970; Donderi & Zelnicker, 1969). Does this happen with letters as stimuli? Does changing the field size and brightness change decision latencies when the number and distribution of stimuli remain constant? Data on performance of tasks involving, at the one extreme, very few briefly presented central stimuli and, at the other, many stimuli presented at long exposures over a large field may suggest a consistent interpretation for both single-fixation and multiple-fixation decisions, and for both same-different and target-search tasks.

The three experiments reported here help to fill the information gap. Capital letters were displayed. In all of the experiments, observers viewed each display ad lib, and ended it simultaneously with the decision. The experiments differed in number of days of testing, number of displays presented, and size and luminance of the display screen; but they were identical in the task and response required of the observers: to signal the difference between displays with all letters the same and displays with one letter different from the rest.

GENERAL METHOD

Apparatus and Technique

The display presentation and response recording techniques were always the same. Onset and termination of the display were under the observer's control. The experimenter signaled readiness for a trial by lighting a green lamp on the response panel. The observer first pressed a central button, which extinguished the green ready light, and then released the central button to present a display and start a timer. Then the observer pressed one of two lateral buttons to record the response, stop the timer, and end the display.

The gray wood-surface response panel was 12 in. (30.5 cm) square and inclined 20 deg from horizontal. The buttons were 7/8-in-diam (2.2-cm-diam) Plexiglas circles mounted flush with the surface over spring-loaded microswitches. The center button was mounted 4 in. (10 cm) from the bottom edge. The lateral buttons were centered on a line parallel to and 8 in. (20 cm) from the bottom edge. The linear distance between the central button and each lateral button was 4.5 in. (11.4 cm). The green ready light was mounted at the far end of the panel away from the observer.

Display

Each display slide was photographed from black Letraset sans-serif capital letters A, B, and C mounted on a 5 × 5 in. (12.7 × 12.7 cm) square of graph paper divided into 100 .5-in.^{2} (.127-cm^2) regions (Figure 1). The letters on each display were distributed into N=2, 5, 8, 11, 14, 17, 20, and 23 of these 100 squares, which were chosen at random, independently for each display, by finding N different pairs of horizontal and vertical coordinates from a table of random numbers. There were 24 displays for each N. Four *same* displays were all As, 4 were all Bs, and 4 were all Cs. There were 12 *different* displays for each N. Each letter ap-

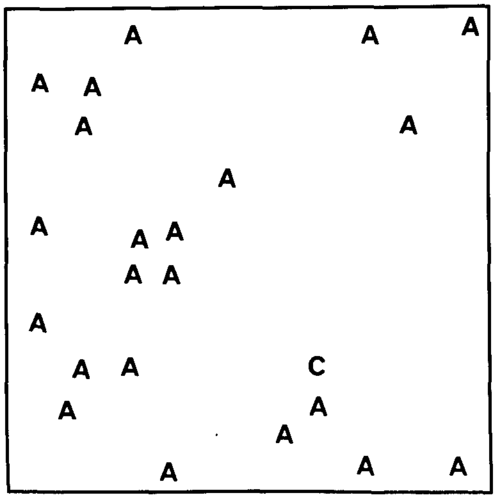

Figure 1. Display of 23 letters, 1 different from the rest. All letter positions, including the position of the 1 different letter, were selected at random in each display.

peared four times as the single different letter, and twice with each of the other two letters as the N−1 identical background letters. The location of the different letter was determined randomly for each display. The complete set contained 192 displays: 12 *same* and 12 *different* at each of eight different values of N.

Instructions

Observers were instructed to self-pace the display presentations following the green ready signal by releasing the center button on the response panel only when ready. Each observer chose and used one hand consistently throughout the experiment. Instructions were to respond with the "same" button if all the letters in the display were the same and with the "different" button if one or more letters were different from the rest. The instructions were to respond "as quickly as you can but with as few mistakes as possible." Ten practice trials selected from the regular display series were given before each experiment.

Observers

Of the 50 observers, who ranged in age from 14 to 27 years, 22 were men. They were volunteers, either university students majoring in psychology, who participated for free, or paid high school and university students referred through a student employment office during the summer. Each observer participated in only one experiment. They were assigned to conditions within experiments in rotation, based on order of appearance in the laboratory. In describing and analyzing the results, no account was taken of differences in age, sex, or education.

EXPERIMENT 1
FIVE DAYS, N = 5, 8, 11, 14

Response times (RTs) to *same* and *different* displays with 5, 8, 11, and 14 letters were obtained over 5 days, so that changes in the relationship between RT, N, and the type of display could be observed with increasing amounts of practice.

Method

There were six observers. They sat 1.9 m in front of a white cardboard screen, 43 cm wide × 33 cm high. The experiment was carried out in dim ambient light. The screen without the display reflected 2.1 cd/m² to the observers' position. The bright background of the illuminated display reflected 253 cd/m², and the letter contrast was 85%. The display subtended a viewing angle of 10 × 10 deg, and each displayed letter was contained within a box .4 deg square.

The 12 *same* and 12 *different* slides for each N were divided randomly, and half were assigned to each of two Carousel slide trays. Each tray contained 48 slides: 6 *same* and 6 *different* slides from each N, arranged in random order. The trays were presented in counterbalanced order each day, and each tray was rotated clockwise or counterclockwise in counterbalanced order over the 5-day series. Ten slides from the second tray were presented as practice trials before the first day's data were recorded. The response type and, to the nearest .01 sec, the RT from the onset of the display were recorded. The positions of "same" and "different" buttons were counterbalanced across the observers but remained constant for each observer throughout the experiment.

Results

The results of the analyses of the mean RTs of correct responses will be presented following the description of each experiment. Analyses of the standard deviations of RTs for correct responses, and analyses of the wrong responses, will be presented in separate sections following the descriptions of all of the experiments.

Correct responses. Each wrong decision was replaced by the mean of the correct RT for the same N, display type, observer, and day.

The mean RT declined from .91 sec on the first day to .73 sec on the fifth [$F(4,5) = 14.88$, $p < .001$]. The decline over days (d) was fitted by the equation $RT = .6\exp(-d) + .73$, $r^2 = .99$, indicating that an asymptotic level had been approached. An exponential function approaching an asymptotic minimum reflects the nature of this decision task, which includes an irreducible minimum physical response duration as an a priori component. Averaged over days, the RT increased monotonically with an increase in N, from .77 to .82 sec [$F(3,15) = 3.74$, $p < .05$]. There was a significant increasing linear trend with N [$F(1,15) = 7.52$, $p < .05$], but no other trend was significant. The display type did not significantly affect RT, nor did it interact significantly with test days or with N, as determined by an analysis of variance with observers as the single random factor. Table 1 displays the RT for each value of N on each of the 5 days of the experiment, as well as SDs for days and N. To assess the interaction of observers with the other factors, the 12 daily trials within each N, display type, and observer type were treated as the random replications factor, and the within-cells degrees of freedom were reduced to compensate for the substitution of means for wrong responses. Using this random factor, the interactions of observers with each of days [$F(20, 2556) = 6.6$, $p < .01$], N [$F(15,2556) = 2.4$, $p < .05$], and display type [$F(5,2556) = 9.9$, $p < .01$] were significant, but no three-way or four-way interaction approached significance. Thus, each observer was significantly different over days, N, and type of display, but the effects were independent. The interaction with display type is the most interesting (Figure 2). Observers who responded rapidly responded more rapidly to *same* than to *different* displays; slower responders did the opposite. The regression of mean "different" ($\bar{D}$) on mean "same" ($\bar{S}$) RT in seconds

Table 1
Experiment 1: Mean Response Time for N = 5-14 Letters Over Days 1-5

Day	Number of Letters (N): 5	8	11	14	Mean per Day	SD per Day
1	88	90	91	96	91	21
2	80	79	83	83	81	17
3	74	77	76	79	76	16
4	75	70	71	77	73	16
5	72	72	73	75	73	15
Mean per N	78	78	79	82	79*	
SD per N	18	15	16	18		17†

Note—*Entries are in hundredths of a second.*
**Grand mean. †Grand SD.*

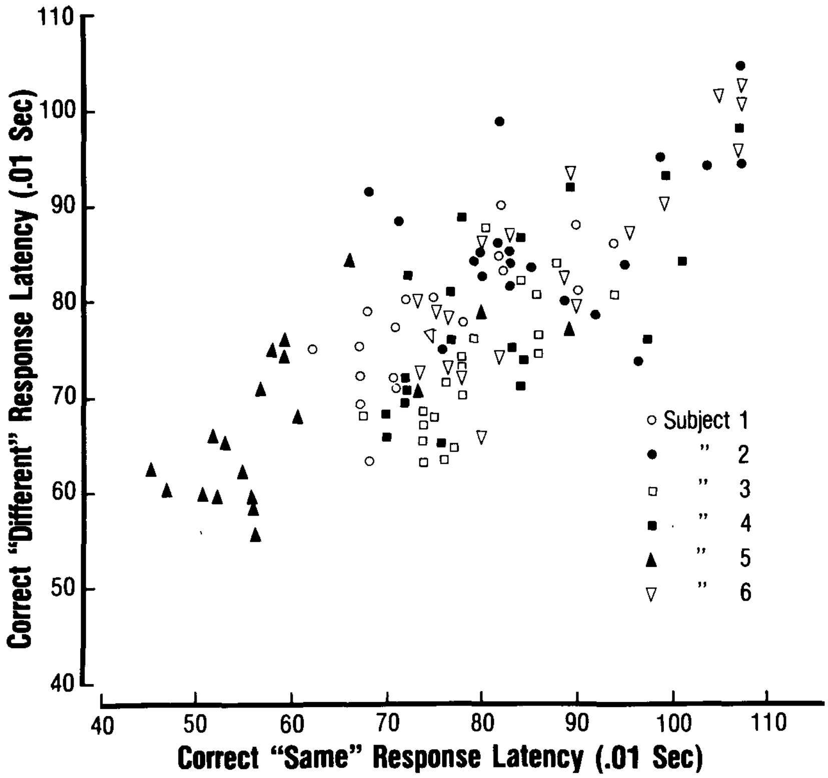

Figure 2. Relationship between mean correct "different" ($\overline{D}$) and "same" ($\overline{S}$) response times at constant N (number of letters displayed) for each observer on each day of Experiment 1. The regression equation is $\overline{D} = .54\,\overline{S} + .36$.

over the 120 separate "same" and "different" averages in the experiment—observers (6) × days (5) × N (4)—was $\overline{D} = .54\overline{S} + .36$, $r^2 = .60$ ($p < .001$).

Distribution of RTs. Each RT was expressed as a deviation from its particular cell mean RT, and all of the deviations were summed in a frequency histogram, which was positively skewed towards longer RTs. In order to study the effect of the positive skew, the log RT data were analyzed in the same way as the original data. There were *no* changes in the pattern or level of significant effects from the analysis of raw RTs to the analysis of log RTs. The skew was not extreme enough to affect any inferences drawn from analysis of variance of the RT data.

Discussion

Two interesting results are, first, the complete insensitivity of mean RTs to differences in display type (*same* or *different*) and, second, the relative insensitivity of mean RTs to changes in N. The increase with N was significant but small, amounting to a change in RT of .04 sec over a display N range of nine letters. Both of these results hold over all 5 days of practice: there are no significant interactions with days, and the variance contributed by the interactions is on the order of the trials-within-conditions variance.

On the other hand, a significant and interesting interaction does occur between observers and display type. Observers whose mean RTs are short respond faster to *same* than to *different* displays; observers with long mean RTs respond faster to *different* than to *same* displays (Figure 2). Observers was the only factor in this experiment that influenced differences between display types.

No serial processing model of reaction time matches the data from Experiment 1. The only parallel model compatible with the data is the parallel model with constant processing times. According to Taylor (1976), total response time increases with N or with log N in every serial model or parallel model with variable processing time, and in all of these models correct "different" responses are faster than correct "same" responses, the differences between display types increasing with N. There was not even the suggestion of an interaction between N and display type in Experiment 1 ($F < 1$). However, in a parallel model with constant processing time, the time for each element may increase with N, but there should be no difference between mean "same" and "different" RT (as in Experiment 1).

In theories of reaction time discussed by Krueger (1978), Schneider and Shiffrin (1977), Snodgrass and

Townsend (1980), Sperling et al. (1971), and Taylor (1978), reaction time either is independent of N or has a clear functional relationship to N. The data from this experiment suggest that N may only *contribute* to RT along with such other equally important factors as display luminance, letter discriminability, and contrast.

EXPERIMENT 2
ONE DAY, N = 5, 8, 11, 14

In the first experiment, the visual displays were large and bright, and they were seen by each observer on 480 separate exposures over 5 days. Under these conditions, display type (*same* or *different*) mean RTs varied significantly across observers, but not across N. The lack of an interaction with N was unexpected, as was the generally small increase in RT with N.

The highly visible displays and the spaced practice may both have contributed to these results. The relationship found between correct "same" and "different" RTs has interesting consequences if it is generalizable. The second experiment was designed to obtain data from the same type of display used in Experiment 1. But each observer was tested for only one session, and the displays were smaller than those used in the first experiment. Experiment 2 was designed to learn whether, under these conditions, (1) there was an N × display type interaction, (2) the empirical relationship between correct $\overline{S}$ and $\overline{D}$ RTs was repeated, and (3) the results were otherwise compatible with those of Experiment 1.

Method

The slides were rear-projected onto a screen located 1.4 m in front of the observers. The display was 4.5 in. (11.4 cm) square, and subtended a visual angle of 5.6 deg square. Ambient luminance from the screen was .7 cd/m², and the letters, which were .21 deg square, had a contrast of 75%.

Each observer responded to 24 displays at each N, N = 5, 8, 11, and 14. Twelve displays at each N were *same*, and 12 were *different*. The order of the displays was random. There were two groups of 48 displays each. Each group of display slides was contained in one Carousel projector tray. There were two replications of the two-tray, 96-display sequence separated by a 10-min rest period. The trays were presented in counterbalanced order over both replications, and the direction of rotation through each tray was also counterbalanced.

Results

The mean and SD of the correct responses were calculated for each observer for *same* and *different* displays at each N. There was no difference between the results for either measure on the first and second replications. The mean RTs were longer as N increased, and RTs to *same* displays were longer than RTs to *different* displays. The mean and standard error for each display type at each N is shown in Figure 3.

Analysis of variance of the mean RTs using the ap-

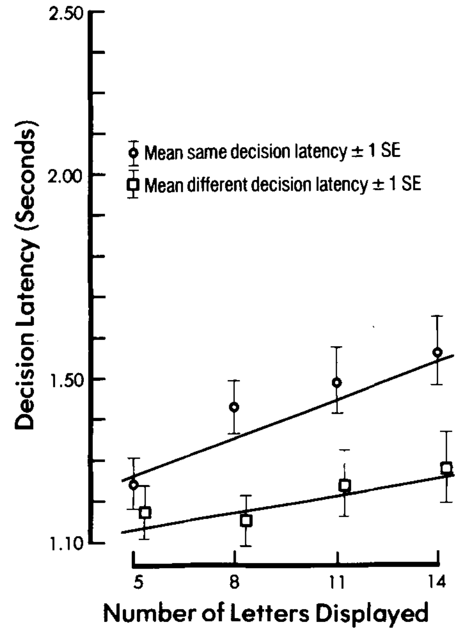

Figure 3. Mean and SE of correct "same" and "different" response times as a function of N (number of letters displayed) for Experiment 2. The regressions are linear with no significant residuals.

propriate observer interaction as the error term indicated that both main effects of N and display type were significant at $p < .001$, and the interaction was significant at $p < .01$. The simple effects of N for both display types were linear at $p < .01$ when tested by orthogonal components of variance. There was no significant residual from the linear trend in any case. The equations of the linear best fit to the mean RTs in seconds for each condition were: with *same* displays, $\overline{S} = .03N + 1.11$, with *different* displays, $\overline{D} = .01N + 1.08$.

With fewer trials than in Experiment 1, "same" and "different" RTs were significantly different, and interacted differently with N. In Experiment 1, N played a minor role in determining RT. The regression of RT on N for the first day of Experiment 1 was $RT = .008N + .83$ (seconds), a flatter slope with a lower intercept than either display type in Experiment 2. In Experiment 2, under different display and practice conditions, N interacted with display type to generate linear RT functions which were different for *same* and *different* displays. As a result of the differences between Experiments 1 and 2, two questions needed to be pursued further: the effects of (1) the range of N and (2) variations in the display conditions.

EXPERIMENT 3
N = 2 to 23

As N increases, the density of letters on the display increases also, and the difference is visibly quite noticeable between N = 14 and N = 23. In order to assess the importance of area, luminance, N, and their interactions with *same* and *different* displays, all variables must be compared over a range of values. The following experiment studied a larger range of N, at several display areas and luminances.

Method

Each observer responded to a range of *same* and *different* displays containing from 2 to 23 letters. Display size varied between observers, and display luminance varied within observers. The mean and standard deviation of the RT for correct responses as well as the number and mean latency of wrong responses were recorded for each condition (N × display type × display size × display illumination).

Displays were rear-projected onto a screen. The large display measured 10 in. (25.4 cm) square, and the small display was the one used in Experiment 2. At the 1.4-m viewing distance, the large display subtended a viewing angle of 10.4 deg square. Sixteen observers responded to the large display, and 16 responded to the small display.

There were 12 *same* and 12 *different* displays with 2, 5, 8, 11, 14, 17, 20, and 23 letters displayed. They were divided into four sets of 48, each set containing 3 *same* and 3 *different* displays at each N. The display slides in each set were arranged in random order in a Carousel projector tray. Two sets in succession were presented at a bright illumination level, and two in succession were presented at a dim level. The order of presentation of the bright and dim sets was counterbalanced across subjects. The specific sets shown in the bright and dim conditions were also counterbalanced across observers.

Area of the displays was altered by varying the distance from the projector to the screen. Screen luminance varied with the projection distance and was further controlled by changing the angle between two Polaroid filters in the projector beam. The filters did not completely compensate for the change in projector distance. The high and low illumination levels of the small display area were more intense than the high and low illumination levels, respectively, of the larger display area. For the small displays, the high illumination level was 172 cd/m², and the low level was 27 cd/m², with letter contrast of 75%. For the large displays, the high level was 45 cd/m², the low level was 6.9 cd/m², and the letter contrast was 75%. Ambient illumination from the screen between displays was .85 cd/m².

Each observer responded to 192 slides in a single session, divided halfway by a 5-min rest. The session was preceded by 10 practice trials drawn from the regular displays. The entire task was completed in about 45 min.

Results

Figure 4 relates the mean and SE of RTs for correct "same" and "different" responses to each N, averaged over the four area × luminance display conditions. Response time increased with an increase in N, and correct "same" decisions took longer than correct "different" decisions. Table 2 presents the mean and SE of correct RTs to the four combinations of area and luminance, averaged over N. Responses took longer when the displays were large and dim.

Four factors contribute to correct RT: display area (A), display illumination (bright or dim, L), the number of letters displayed (N), and the type of display, *same* or *different* (D). Analysis of variance of the mean RTs was carried out using the appropriate observer error term or condition × observer interaction: the summary is presented in Table 3. The pattern of significant main effects and interactions demonstrates that there are two independent sets of significant factors: the display *constant* factors A and L, and the display *variable* factors N and D. The factors within each set interact, but the significant constant factor interaction A × L did not interact with the significant variable interaction N × D. In other words, the joint effects of area and luminance—constant across displays for each observer—were independent of the joint effects of number of letters and type of display—variable across displays for each observer.

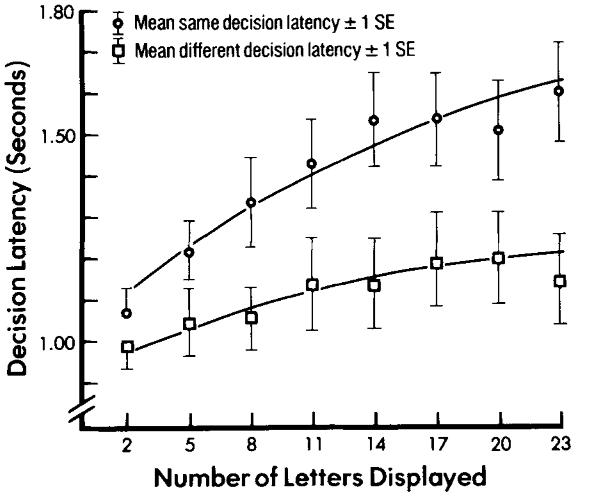

Figure 4. Mean and SE for correct "same" and "different" RTs, averaged over all display conditions, in Experiment 3, as a function of N (number of letters displayed). The least squares regression equations are $\bar{S} = 15.9\ N^{1/2} + .88$ and $\bar{D} = 7.19\ N^{1/2} + .88$.

The strongest constant effect was that of luminance. Table 2 shows that mean RT was a decreasing function of display luminance, while display area contributed relatively little, except as it was confounded with luminance, to either the mean or the SD of the RT among conditions. The correlation between $\overline{RT}$ and log luminance over the four conditions was −.96.

The interaction of the two variable display factors N and D was averaged across all of the constant A ×

Table 2
Mean and Standard Deviation of Correct RT (in Seconds) per Observer for Each Area × Luminance Display. Averaged Over Observers in Experiment 3

	Area = 5.6 deg square		Area = 10 deg square	
	Luminance		Luminance	
RT	27 cd/m²	172 cd/m²	6.9 cd/m²	45 cd/m²
Mean	1.33	1.04	1.55	1.12
SD	.30	.23	.43	.24

Table 3
Analysis of Variance of Experiment 3:
Mean Correct Response Time

Source of Variance	Mean Square	df	F
Area (A)	325,541	1	16.5†
Illumination (L)	59,414	1	24.3†
Number of Letters (N)	21,461	7	64.1†
Type of Display (D)	219,961	1	92.0†
Observers Within Area S(A)	19,671	30	
A × L	13,470	1	5.5*
A × N	1,844	7	5.5†
L × N	1,052	7	4.3†
A × D	11,143	1	4.7†
L × D	5,023	1	9.9**
N × D	4,850	7	20.6†
S × L (A)	2,443	30	
S × N (A)	335	210	
S × D (A)	2,392	30	
A × L × N	189	7	.8
A × L × D	433	1	.9
A × N × D	823	7	3.5**
L × N × D	630	7	2.6*
S × L × N (A)	245	210	
S × L × D (A)	508	30	
S × N × D (A)	235	210	
A × L × N × D	163	7	.7
S × L × N × D (A)	241	210	

*$p < .05$. **$p < .01$. †$p < .001$.

L display factors, with which N × D did not significantly interact. Orthogonal trend analyses were carried out on the averaged data. For the *same* displays, the linear and quadratic components were significant at $p < .001$, and the residual was significant at $p < .05$. The quadratic term was contributed by the negatively accelerating aspect of the RT function between 2 and 23 letters; the significant residual was probably contributed by the reversal between $N = 14$ and $N = 17$. For the *different* displays, both the linear and the quadratic terms were significant at $p < .001$, while the residual was not significant. The "different" latencies also show a negatively accelerating increase with N.

The nonlinear RTs in Experiment 3 may be caused by the increasing letter density as N increases over a uniform display area. In a square display of area A with N uniformly distributed letters, each letter is located in the center of a square with area A/N. The distance between one letter and the next is equal to the distance between the centers of the squares, which is equal to the side of a square, which is equal to $(A/N)^{1/2}$. A is constant, so $(A/N)^{1/2} = K \cdot N^{-1/2}$. Suppose that the time to search from one letter to the next is directly proportional to the distance between letters. In order to make a correct "same" decision, N letter must be searched. The search time will be proportional to $N \cdot K \cdot N^{-1/2}$, or to $K \cdot N^{1/2}$. In order to make a correct "different" decision, an average of N/2 letters must be searched, and the total search time will be proportional to $(N/2) \cdot K \cdot N^{-1/2}$, or to $(K \cdot N)^{1/2}/2$. This reasoning applies directly to uniformly spaced letters, and by approximation to the randomly spaced letters of Experiment 3. It predicts that the slope coefficient for correct "different" RT will be one-half of the coefficient for correct "same" RT.

The linear least squares regression of correct "same" latency against $N^{1/2}$ gives $\overline{S} = 15.9N^{1/2} + .88$, $r^2 = .95$ (Figure 5). This relationship is clearly linear in $N^{1/2}$. The RT for $N = 23$ is below the trend of the other values. Without the "different" RT value at $N = 23$, $\overline{D} = 7.18N^{1/2} + .88$, $r^2 = .98$. This equation has the same intercept and half the slope, within the limits of experimental error, as the equation for $\overline{S}$. Thus, the decision latencies of Experiment 3, with the exception of one data point, can be explained by assuming that serial self-terminating search time is proportional to the distance between letters. Bloomfield (1973) also reports that search time is a square root function of the number of uniformly spaced nontargets in a constant-area display containing many more nontargets than targets.

To summarize the results of Experiment 3: (1) $\overline{RT}$ increased as a square root function of N from 2 to 23, (2) the slope of the "different" function was about half of the slope of the "same" function, and (3) small, bright displays produced shorter RTs than did large dim displays.

GENERAL RESULTS AND DISCUSSION

Standard Deviations of Correct RTs

Experiment 1. Standard deviations (SDs) were calculated over the 12 trials for each cell (observer × display type × day × N). The SDs were significantly larger for *same* displays (18.0) than for *different* (16.0) [$F(1,5) = 10.3$, $p < .05$], they decreased significantly from the first (20.9) to the fifth (14.6) day [$F(4,20) = 6.2$, $p < .01$], and they varied significantly over N, with the SDs for $N = 8$ (14.4) and $N = 11$ (16.0) being smaller than the SDs for $N = 5$ (18.5) and $N = 14$ (18.1) [$F(3,12) = 5.4$, $p < .05$] (see Table 1). There were no significant interactions. The appropriate observer × condition or observer × interaction mean square was used as the error term.

Experiment 2. Analysis of variance of the standard deviations indicated significant differences among N [$F(3,66) = 15.5$, $p < .001$]. There was no significant effect of display type (*same* or *different*) or significant interaction with this factor.

Experiment 3. Analysis of variance was used to evaluate the factors contributing to differences among the SDs of correct RTs within each display type × N × area × illumination block for each observer. SDs were larger for large displays [$F(1,30) = 16.1$, $p < .001$] and for the dimmer displays [$F(1,30) =$

20.6, $p < .001$]. The interaction between area and illumination was also significant [$F(1,30) = 14.2$, $p < .001$], because the SDs varied as a function of luminance over the large (dimmer) but not the small (brighter) displays. These data are summarized in Table 2. SDs increased from 21.2 to 35.3 as N increased from 2 to 23 [$F(7,210) = 9.7$, $p < .001$]. There was no significant difference between the SDs to *same* or *different* displays. The relationship between display type and N was different at the two illumination levels (N × D × L interaction) [$F(7,210) = 3.05$, $p < .05$]. This interaction highlights a relative decrease in the SDs for *different* displays with N when the display was brighter. The striking feature of the SD data was the constancy across *same* and *different* displays, the increase with an increase in N, and sensitivity to change in the low illumination levels of the large display.

Table 4
Mean Response Time (in Seconds), and Response Frequency (in Percent) for All Display-Response Combinations

	Display			
	Same		Different	
Response	RT	Frequency	RT	Frequency
	Experiment 1			
"Same"	.79	98	.65	7
"Different"	.83	2	.79	93
	Experiment 2			
"Same"	1.65	91	1.15	3
"Different"	1.32	9	1.29	97
	Experiment 3			
"Same"	1.41	97	.88	2
"Different"	1.60	3	1.12	98

Wrong Responses

Experiment 1. Of the 2,880 responses, 84 (3%) were wrong. The number and latency of wrong responses did not vary with N. There were more wrong "same" responses to *different* displays than the opposite, and the average RT of the wrong responses decreased over days.

Experiment 2. There were 2,304 responses, of which 142, or about 6%, were wrong. There was a tendency for the latency of wrong responses to increase with the number of letters per display, and the more numerous wrong "different" responses took longer than the wrong "same" responses. Wrong "same" responses were faster than correct "same" responses, and wrong "different" responses were slower than correct "different" responses.

Experiment 3. There were 129 wrong responses among 6,144 responses, for an error rate of 2%. There were more wrong "same" responses made to *different* displays than the reverse, and the wrong "same" responses increased in latency as the number of letters increased. There was no relationship between the number of letters in the display and mean RT for wrong "different" responses, and the frequencies of errors to either the *same* or *different* displays did not increase with an increase in the number of letters on display. The effects of display size and luminance were small. There were fewer errors made to the large, bright display (25) than to the other three displays, for which the error totals were: large dim, 40; small bright, 39; and small dim, 35.

A summary of the mean RT and percentage data for both the correct and the wrong responses from Experiments 1, 2, and 3 is presented in Table 4.

Correct Responses

The data from Experiments 2 and 3 were simplified by displaying them in the same way as the data from Experiment 1 (Figure 2). The average correct "different" latency ($\overline{D}$) was plotted as a function of the average correct "same" latency ($\overline{S}$) for each N in every condition. The plot (Figure 5) contains 36 points, eight from each of the four area and brightness combinations of Experiment 3 and four from Experiment 2. The data include results from two display sizes and two display brightnesses.

The data give a linear plot described by the regression equation $\overline{D} = .54\,\overline{S} + .36$, $r^2 = .83$. The standard error of the regression coefficient is .04, and the standard error of estimate is 7.3. This equation is the same as the regression equation of Experiment 1. The function relating mean "same" and "different" RTs was calculated, and Figure 2 was plotted, from six observers' data in Experiment 1, while 44 observers are averaged over the points used to calculate the equivalent relationship for Experiments 2 and 3, and to plot the data in Figure 5. The relationship between mean correct "same" and mean correct "different" RTs over Experiments 2 and 3 is again characterized by two constants: (1) a constant bias of about .4 sec in favor of "same" decisions, and (2) a constant proportion of about .5 between the "different" and the "same" decision latencies. The linear relationship between correct "same" and "different" RTs is invariant over N, while the display and task conditions of each experiment determine both the function relating response time to N, and the mean RT.

The .4-sec bias in favor of "same" responses is consistent with Krueger's theory that within a single glance an accurate "same" decision can be made faster than an equally accurate "different" decision. As the response time increases, the contribution of the additive constant decreases relative to the contribution of the proportional constant (Barber, 1981). The .5 proportion may result from a serial scan of each letter until a different letter is found or until all the letters have been searched, assuming that either equal time is taken to scan each letter during

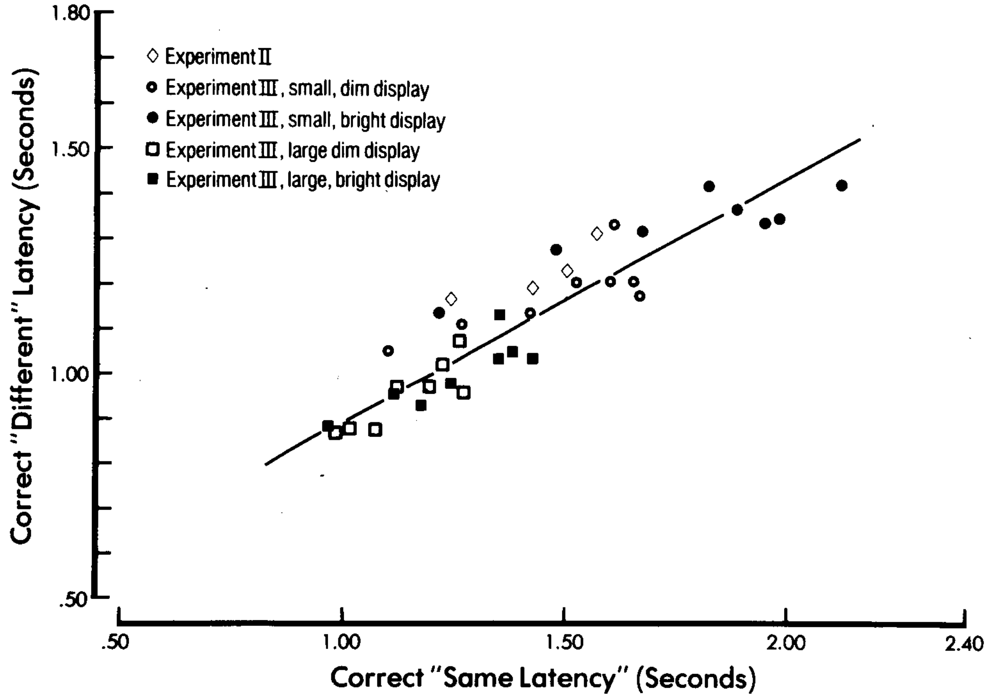

Figure 5. Relationship between mean correct "different" ($\overline{D}$) and "same" ($\overline{S}$) response times at constant N (number of letters displayed) for Experiments 2 and 3. The regression equation is $\overline{D} = .54\,\overline{S} + .36$.

each part of the search *or that equal time is taken to scan equivalent display areas.* If there is a different letter, it will be found on the average after half of the time. Therefore, however long it takes to search N letters, searching half of that number will take half of the time.

The basic process of Krueger's theory occurs in a single glance, where Krueger suggests that several characters can be compared. Although several characters can be compared, only two attribute vectors are generated. Each vector contains the attributes of a string of characters. A "same" or "different" decision is made by counting the number of attributes which differ between the two vectors. There is only one comparison process per glance, regardless of the number of characters.

In Krueger's theory, a multiple-glance decision is necessarily serial and self-terminating. Successive glances register "same" either until one glance registers "different," ending the trial with a "different" response, or until all of the characters have been compared, ending the trial with a "same" response. The results of Experiments 1-3 are consistent with this theory. At the single-glance, rapid end of the decision latency range, there is a comparison process in which "same" is faster than "different," while at the slower, multiple-glance, serial self-terminating end of the RT range, "different" is faster than "same." Thus, the results are qualitatively consistent for both short and long RTs.

Krueger's comparison process needs some elaboration to explain the data from Experiments 1-3. Suppose, as Krueger's theory suggests, that several letters are included within a single glance. If a display of 23 letters is presented, four (say) of them are selected on the first glance. Either all of the letters are the same, requiring another glance, or one is different from the rest, terminating the trial with a "different" response. The letters in Experiments 1-3 were scattered randomly over the displays. Krueger's glance process must select an even number of these scattered letters, divide them arbitrarily into two strings, calculate an attribute vector for each set, and then compare the vectors on successive passes. The data from Experiments 1-3 give a constant .4-sec increment for correct "different" over correct "same" decisions, consistent, as Krueger suggests, with a faster within-glance processing of "same" comparisons. The data also show self-termination on multiple glances: the proportionality constant between "different" and "same" RTs on longer (multiple-glance) trials is close to 1/2, which would be expected for a serial self-terminating search. Thus, a short response time means a single-glance decision, while a long response time means a multiple-glance decision.

Standard Deviations

The SDs of correct "same" and "different" responses were regressed separately against the "same" and "different" means for each data point in Experiments 1-3. The two regression lines were $SD_s = .24\ \overline{S} - .08$ for the correct "same" responses, and $SD_d = .44\ \overline{D} - .19$ for the correct "different" decisions (data in seconds). When these two equations are solved for SD_s and SD_d at equal N by substituting the empirical relationship between the means, $\overline{D} = .54\ \overline{S} + .36$, we find that $SD_s \cong SD_d$. Thus, the standard deviations are identical, within the limits of experimental error, for *same* and *different* displays presented under the same experimental conditions, and having the same N.

The *different* display search over N letters should be intrinsically more variable than the *same* display search, because the *different* search can end, with probability 1/N, on any one of N letters, while the *same* search is always of basic length N. Thus, it is surprising that, for equal N, the SD of "same" RTs matches the SD for "different" RTs.

Schneider and Shiffrin (1977, p. 62) derive the formula for the variance of a probabilistic random variable, one example of which is the *different* response times of the experiments reported here. The formula is expressed in terms of the means and variances of the component decision processes whose weighted average forms the random variable. In our experiment, these component decision processes would be the searches of successive letters, of length 1 to N, which are carried out until the single *different* letter is found. Our data require that the variance of the correct *different* searches be equated to the variance of the correct *same* searches for equal N. Using the Schneider and Shiffrin formulation for the variation of a probabilistic mixture in the *different* case, and simple assumptions about the independence of variance for searches of successive letters in the *same* case, no simple expression of the variances of the component searches as linear functions of N will equate the "same" and "different" RT variances. This suggests that the variance of the response time is a function only of the total N for both *same* and *different* displays, and so is not a function of the number of letters which must be scanned during the search. The variance, in other words, seems to depend on processes involved in the acquisition of the display, rather than on processes involved in decisions about the display.

Wrong Decisions

There was a total of 335, or 3%, errors over all 11,328 trials in Experiments 1-3. The rate was 6% in Experiment 2 but below 3% in the others.

The numbers and RTs of wrong decisions were generally consistent across all of the experiments. Wrong "same" responses to *different* displays were usually faster and less frequent than wrong "different" responses to *same* displays. In Experiment 1, this was true on all 5 days, as both the mean RTs and the total number of wrong decisions decreased. In Experiment 2, it was true for N = 5, 8, and 14 letters. At N = 14, there were many fewer wrong "same" responses to *different* displays, but the wrong "different" responses to *same* displays were slightly faster than the wrong "same" responses to *different* displays. In Experiment 3, there were consistently fewer wrong "same" responses to *different* displays at all N, but for N = 5 and 20 the general trend of RTs was reversed, and the wrong "different" responses to *same* displays were faster.

These results are consistent with those reported by Krueger (1978, p. 290) from a survey of same-different experiments. Forty-two of 65 appropriately selected experiments produced fewer "same" responses to *different* displays than vice versa, and in 35 of 57 cases with recorded RTs, the "same" responses to *different* displays were faster. This was interpreted by Krueger to mean that the distribution of mismatching elements in *same* displays either had a greater variance or was more positively skewed than the distribution of mismatching elements in *different* displays. Thus, both the correct RT results and the error frequency and RT results of these experiments were consistent with the process postulated by Krueger, although these results extend that consistency to RTs as long as 2 sec.

The relationship between $\overline{RT}$ and response frequency for both correct responses and errors is displayed in Table 4. The fastest RTs were "same" responses to *different* displays. "Different" responses to *different* displays were next, followed (with one exception) by correct "same" responses and then by "different" responses to *same* displays.

These data immediately disconfirm a simple signal detection theory (SDT) interpretation as proposed by Bindra, Williams, and Wise (1965) or by Donderi and Case (1970). In a SDT model, the small percentage of "same" responses to *different* displays (failures to detect the different letter) means that the very small "error" tail of the *different* display intensity distribution (analogous to the S + N distribution in SDT) falls close to the criterion. This, in turn, implies long RTs to the incorrect "same" responses. However, these RTs were the shortest recorded in Experiments 1-3.

The Separation of Acquisition and Decision

In Experiment 1, response time varied slightly as a linear function of N. In Experiment 2, the relationship was strongly linear; and in Experiment 3, RT was linear in $N^{1/2}$. Nevertheless, the additive and proportional constants relating "different" to "same" response time were invariant across experiments. While clearly not universal (e.g., Egeth, Jonides, & Wall, 1972), a linear serial self-terminating search model with a constant increment for "different" decisions adequately describes the data from all of the same-different experiments reported here. Schneider and Shiffrin (1977, p. 27) concluded that their serial search processes were linear and self-terminating. Barber (1981) proposed that visual search tasks required an inspection strategy that was independent of the subsequent processing strategy. The data from Experiments 1-3 are consistent with these proposals. Inspection or acquisition strategies vary widely from one experiment to the next, while the decision process remains the same.

The duration of the acquisition process depends on display and presentation variables including, but not limited to, N. Experiment 1 also demonstrates that the duration of the acquisition process also varies from one person to the next. For *same* displays, the "same" decision requires the complete acquisition process since all the information must be acquired before a correct "same" response can be made. One-half of the acquisition time is required, on the average, before a correct "different" response can be made. This implies a search in which the probability of finding the single different letter is equal at every point in the search. The search time is not directly dependent on N, but is also a function of the area searched, the display luminance, and the symbol density.

Same-different decisions were made to displays of from 2 to 14 geometric shapes haphazardly located over the display (Donderi & Case, 1970; Donderi & Zelnicker, 1969). In the Donderi and Zelnicker experiments, either all of the geometric shapes were the same or one shape in the display was different from the rest. The shapes were exposed for .05 sec at 44 cd/m^2. Correct "same" RTs were longer, at 1.36 sec, than correct "different" RTs, at 1.29 sec, for all N. If we assume, as for Experiments 1-3, that on the average the different shape was found after half of the acquisition time, and if we postulate a fixed increment (T_d) for the "different" decision, then we have, as in the example for Experiment 1: "same" RT = 1.36 sec and "different" RT = 1.29 sec = $(S/2) + T_d$. Solving, we have T_d = .61 sec, a longer value than for the letters of the present experiments.

Donderi and Case presented all shapes the same, one shape different from the rest, or three shapes different from the rest over a range of 2 to 14 shapes. The exposure was .20 sec at 44 cd/m^2. Response times were essentially constant over N: .775 sec for correct three-"different," .782 sec for "same," and .825 sec for correct one-"different" responses. If we assume that again, for one-*different* displays, on the average the single different shape is located after half of the acquisition time, we have S = .782 sec, and one-"different" RT = .825 sec = $(S/2) + T_d$. Solving, T_d = .434 sec, a value similar to those obtained from the present experiments.

We now have estimates for the acquisition time and the constant "different" decision time in the Donderi and Case experiments. A priori, the single different shape is acquired on the average after one-half of the acquisition time. When, on the average, is any one of the three-*different* shapes acquired? Set three-"different" RT = .775 sec = $(S/x) + T_d$. Given T_d = .434 sec from the one-"different" condition and solving, x = 2.29. The first of the three "different" letters is acquired earlier than halfway through the duration of a "same" search: 1/2.29 = .436 of the way through. Again, acquisition time is independent of N.

Making a same-different decision when the unique different letter changes from trial to trial takes a long time. In Schneider and Shiffrin's (1977) terms, it is an example of controlled information processing. When the controlled decision process is at its most efficient, "same" responses are faster than "different" responses for a fixed N. This bias is expressed as a constant term in the linear equation relating "different" to "same" RTs with N as a parameter. The proportionality coefficient of the equation is close to .5, meaning that when decisions are slower, "different" responses are completed in about half the time it takes to complete a "same" response for equal N. This, in turn, strongly suggests that the decision follows a self-terminating search for one different letter. As long as the probability of finding a different letter remains constant across the search, the different letter will be found, on the average, halfway through the search.

The linear invariance relating "same" and "different" RTs at fixed N is a between-observer effect in these experiments. It depends on the fact that there are large differences among the average RTs of different observers. We do not know whether the same invariance can be found within one observer over a wide range of RTs. The six people tested in Experiment 1 (Figure 2) do not show consistent individual invariances. However, the within-observer range of RTs was small compared with the between-observer range.

The function relating RT to N varies from experiment to experiment. This strongly suggests that the time course of acquisition of information from the display is sensitive to variations in the experimental method, while the decision process that uses that information is independent of the display factors that influence speed of acquisition.

In summary, these experiments say three things about same-different decisions made to letter displays: (1) Information acquisition is separate from decision, since linear and square root acquisition functions of N were identified under different experimental conditions; (2) there is a constant .4-sec increment for "different" over "same" responses; and (3) the acquisition process is a self-terminating search, which is not exclusively dependent on N.

REFERENCES

Barber, P. Visual search and number of stimuli reexamined. *Psychological Bulletin*, 1981, **89**, 176-182.

Bindra, D., Williams, J. A., & Wise, J. Judgments of sameness and difference: Experiments on decision time. *Science*, 1965, **150**, 1625-1627.

Bindra, D., Donderi, D. C., & Nishisato, S. Decision latencies of "same" and "different" judgments. *Perception & Psychophysics*, 1968, **3**, 121-130.

Bloomfield, J. R. Experiments in visual search. In The Committee on Vision, National Research Council (Eds.), *Visual search*. Washington, D.C: National Academy of Sciences, 1973.

Donderi, D. C., & Case, B. Parallel visual processing: Constant same-different decision latency with two to fourteen shapes. *Perception & Psychophysics*, 1970, **8**, 373-375.

Donderi, D. C., & Zelnicker, D. Parallel processing in visual same-different decisions. *Perception & Psychophysics*, 1969, **5**, 197-200.

Egeth, H., Jonides, J., & Wall, S. Parallel processing of multi-element displays. *Cognitive Psychology*, 1972, **3**, 674-698.

Krueger, L. E. A theory of perceptual matching. *Psychological Review*, 1978, **85**, 278-304.

Rabbitt, P., Cumming, G., & Vyas, S. Improvement, learning, and retention of skill at visual search. *Quarterly Journal of Experimental Psychology*, 1978, **30**, 319-322.

Schneider, W., & Shiffrin, R. M. Controlled and automatic human information processing: I. Detection, search and attention. *Psychological Review*, 1977, **84**, 1-66.

Snodgrass, J. G., & Townsend, J. L. Comparing parallel and serial models: Theory and implementation. *Journal of Experimental Psychology: Human Perception and Performance*, 1980, **6**, 330-354.

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.

Taylor, D. A. Identification and categorization of letters and digits. *Journal of Experimental Psychology: Human Perception and Performance*, 1978, **4**, 423-489.

Teichner, W. H., & Krebs, M. J. Visual search for simple targets. *Psychological Bulletin*, 1974, **81**, 15-28.

(Manuscript received June 4, 1982;
revision accepted for publication November 29, 1982.)