AFOSR-TR- 88-0726

PREATTENTIVE AND ATTENTIVE VISUAL INFORMATION PROCESSING

BY

Howard E. Egeth

Department of Psychology
Johns Hopkins University
Baltimore, Maryland 21218

Grant No. AFOSR-87-0180
Project No. 2313
Task No. A4

Interim Report

Period Covered: April 1, 1987 through March 31, 1988

June 15, 1988

Grant Monitor: Alfred R. Fregly
Directorate of Life Sciences

Prepared for:

Directorate of Life Sciences
Air Force Office of Scientific Research
Bolling Air Force Base
Washington, D.C. 20332-6448

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE

ADA197670

57

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION	
1b. RESTRICTIVE MARKINGS	
2a. SECURITY CLASSIFICATION AUTHORITY	
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
3. DISTRIBUTION/AVAILABILITY OF REPORT	Approved for public release; distribution unlimited.
4. PERFORMING ORGANIZATION REPORT NUMBER(S)	
5. MONITORING ORGANIZATION REPORT NUMBER(S)	AFOSR-TR- 88-0726
6a. NAME OF PERFORMING ORGANIZATION	Johns Hopkins University
6b. OFFICE SYMBOL (If applicable)	
7a. NAME OF MONITORING ORGANIZATION	Directorate of Life Sciences Air Force Office of Scientific Research
6c. ADDRESS (City, State and ZIP Code)	Baltimore, Maryland 21218
7b. ADDRESS (City, State and ZIP Code)	Building 410 Bolling AFB, D.C. 20332-6448
8a. NAME OF FUNDING/SPONSORING ORGANIZATION	AFOSR
8b. OFFICE SYMBOL (If applicable)	NL
9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER	AFOSR-87-0180
8c. ADDRESS (City, State and ZIP Code)	Building 410 Bolling AFB, D.C. 20332-6448
10. SOURCE OF FUNDING NOS.	PROGRAM ELEMENT NO.: 61102F; PROJECT NO.: 2313; TASK NO.: A4; WORK UNIT NO.:
11. TITLE (Include Security Classification)	Pre-attentive and attentive visual information processing
12. PERSONAL AUTHOR(S)	Howard E. Egeth
13a. TYPE OF REPORT	Interim
13b. TIME COVERED	FROM 4/1/87 TO 3/31/88
14. DATE OF REPORT (Yr., Mo., Day)	June 15, 1988
15. PAGE COUNT	57
16. SUPPLEMENTARY NOTATION	
17. COSATI CODES (FIELD / GROUP / SUB. GR.)	0510
18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)	ATTENTION; LEXICAL ACCESS; PERCEPTION; SEMANTIC PROCESSING; SUBITIZING; TEXTURE PERCEPTION; VISION; VISUAL SEARCH. (SDW)

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

Twelve (12) experiments are described in this report. The first nine (9) are concerned with the hypothesis that the identification of the values of stimulus features in multielement visual displays requires serial processing. Contrary to this hypothesis, the weight of the evidence suggests that feature identification can be carried out by spatially parallel processes. The remaining three (3) experiments are concerned with the ability to extract semantic information from several stimuli in parallel. Both alphanumeric character classification and lexical (i.e., word vs. nonword) decisions can be accomplished by parallel processes, but semantic categorization of words cannot. The implications of these findings for theories of attention are discussed. Keywords.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT	UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT ☒ DTIC USERS ☐
21. ABSTRACT SECURITY CLASSIFICATION	Unclassified
22a. NAME OF RESPONSIBLE INDIVIDUAL	Dr. Alfred R. Fregly
22b. TELEPHONE NUMBER (Include Area Code)	(202) 767-5021
22c. OFFICE SYMBOL	AFOSR/NL

DTIC ELECTE AUG 1 8 1988

DD FORM 1473, 83 APR — EDITION OF 1 JAN 73 IS OBSOLETE.

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE

PREATTENTIVE AND ATTENTIVE VISUAL INFORMATION PROCESSING

BY

Howard E. Egeth

Department of Psychology
Johns Hopkins University
Baltimore, Maryland 21218

Grant No. AFOSR-87-0180
Project No. 2313
Task No. A4

Interim Report

Period Covered: April 1, 1987 through March 31, 1988

June 15, 1988

Grant Monitor: Alfred R. Fregly
Directorate of Life Sciences

Prepared for:

Directorate of Life Sciences
Air Force Office of Scientific Research
Bolling Air Force Base
Washington, D.C. 20332-6448

Abstract

Twelve experiments are described in this report. The first nine are concerned with the hypothesis that the identification of the values of stimulus features in multielement visual displays requires serial processing. Contrary to this hypothesis, the weight of the evidence suggests that feature identification can be carried out by spatially parallel processes. The remaining three experiments are concerned with the ability to extract semantic information from several stimuli in parallel. Both alphanumeric character classification and lexical (i.e., word vs. nonword) decisions can be accomplished by parallel processes, but semantic categorization of words cannot. The implications of these findings for theories of attention are discussed.

I. Introduction

Human vision appears to operate in two rather different modes. One involves processing that requires the allocation of cognitive resources (i.e., attention). This "attentive" mode is generally assumed to operate in a serial fashion. The other mode involves processing that can proceed independently of cognitive resources and in a spatially parallel fashion. Preattentive processing occurs first and provides the basic input for further attentive processing.

The characterization of the preattentive mode as spatially parallel and unlimited in capacity seems to be widely accepted. What is controversial, however, is the functional role preattention plays in visual processing. Early selection theorists see preattention as a collection of relatively low-level processes that serve chiefly to segregate the visual field into perceptual units that are then operated on by focal attentive processes to achieve full analysis and interpretation (e.g., Broadbent, 1971; Neisser, 1967; Treisman, 1964). In contrast, late selection theorists argue that preattentive processing permits full semantic analysis of information in the visual field (e.g., Deutsch & Deutsch, 1963; Duncan, 1980; Norman, 1968). On this account, capacity is limited for behavioral action (e.g., Keele, 1973) or for awareness (e.g., Duncan, 1980), but not for the derivation of form or meaning.

The purpose of the present report is to explore performance in a variety of different stimulus-task domains in an effort to clarify the functional role of preattentive processing. We begin with brief reviews of the literature concerning three widely studied tasks: (a) low-level tasks involving the detection or discrimination of simple visual features, (b) tasks in which a correct response requires the integration of information from two or more separable features, and (3) tasks in which responses are dependent on meaning. We then provide a more intensive and focused discussion of recent research in our laboratory, which has been concerned with the first and third of the aforementioned topics.

II. Background

Simple visual features.

A substantial literature suggests that highly discriminable feature differences can be processed in parallel across the visual field. Some of the evidence for this conclusion comes from studies of visual search in which subjects are asked to determine whether or not a display contains a specific target element where the target, when it is present, differs from the remaining elements of the display with respect to a single simple feature (e.g., a red target among blue nontargets; a horizontal line segment among vertical line segments). Several studies have shown that the time required to detect the presence of such a target is virtually independent of the number of nontargets that are present (e.g., Egeth, Jonides, & Wall, 1972, Exp. 2; Egeth, Virzi, & Garbart, 1984; Treisman & Gelade, 1980; Treisman & Souther, 1985). Additional evidence comes from studies of same-different discriminations. In a discrimination task a subject is typically instructed to indicate if all of a set of stimuli are the same or whether one of them is different. Several studies have shown that with stimuli defined in terms of simple features the time to respond same or different is essentially independent of the number of stimuli in the display (e.g., Connor, 1972; Donderi & Zelnicker, 1969; Egeth, Jonides, & Wall, 1972, Exp. 1).

We should note here that it can be a difficult matter to decide whether processing is parallel or serial. As Townsend has pointed out in a series of important papers (e.g., Townsend, 1971; 1972; Snodgrass & Townsend, 1980), limited-capacity parallel models can produce mean reaction time (RT) predictions that are indistinguishable from those for serial models. However, it is difficult for a psychologically plausible serial model to account for data that show mean RT to be independent of the number of stimuli. It is for this reason that we have (conservatively) adopted as one signature of a parallel process independence of RT and display size.

Some additional evidence concerning processing mode is available from another paradigm in which redundant targets are displayed. Several studies have shown that when all stimuli are targets the time required to identify a target decreases as the number of targets present in a display is increased (e.g., Van der Heijden, 1975; Van der Heijden, La Heij, & Boer, 1983). In the study by Van der Heijden et al (1983), subjects had to discriminate Es from Fs; this task might well be construed as requiring the determination of whether a simple feature (a horizontal line segment at the bottom of the character) was present or absent. Subjects had to press a response key when one or more Es was present and refrain from responding when one or more Fs were present. (Es and Fs were never shown on the same display.) Subjects responded slowest to one-element displays, somewhat faster to two-element displays, and most quickly to three-element displays. Such a redundancy gain is compatible with self-terminating processing models but not those that assume exhaustive processing. It also appears to be incompatible with models that assume limited-capacity processing (this includes the entire class of serial models). In fact, what remains plausible in the face of this result is a parallel, self-terminating, unlimited capacity model. However, Snodgrass and Townsend (1980) suggest that the word "unlimited" here may be too strong. It may be possible to devise parallel models with some degree of capacity limitation that could still predict a redundancy gain. However, the degree of limitation would have to be less than that of two well-known parallel models that cannot predict a redundancy gain. One of these assumes a fixed capacity that is divided among the target elements present on a trial; the other is similar except that as targets are processed the capacity that was devoted to them is reallocated to the remaining items. For the sake of simplicity, in the remainder of this paper we will interpret a redundancy gain to mean that capacity is unlimited. However, it should be understood that some degree of capacity limitation may still be compatible with the obtained results. This is an issue that will have to addressed in further more detailed research.

The preceding studies demonstrate a capacity for parallel processing of featural information. When does parallel processing of featural information not obtain? One circumstance in which parallel processing appears not to occur involves the special case where the target of search is distinguished by the absence of a feature that is present in all of the nontargets (Treisman & Souther, 1985; see also Neisser, 1963). Thus, if the target is a circle with an intersecting line (like the tail of a Q) in a background of plain circles, search time is independent of the number of nontarget plain circles. But, if the target is a plain circle and the nontargets are all circles with intersecting lines, then search time increases with the number of nontargets. Treisman and Souther (1985) show how this interesting asymmetry can be accounted for in terms of feature integration theory, which assumes parallel processing of featural information. Thus, this finding does not really challenge the general conclusions concerning attentive and preattentive processing discussed above.

A more provocative failure of parallel processing of feature information was reported by Sagi and Julesz (1985a; see also 1985b). They examined the nature of processing in three different tasks. One, like several of the tasks described above, required a same-different discrimination; the stimuli to be discriminated were horizontal and vertical line segments. The displays consisted of a few "targets" (the horizontal and vertical line segments) embedded in a texture composed of diagonal line segments that were all oriented in the same direction. Subjects had to indicate whether all of the targets were of the same orientation (all horizontal or all vertical) or whether one of them was in a different orientation (e.g., one vertical among three horizontal). Stimuli were presented briefly and followed by a pattern mask. Performance was measured in terms of accuracy at different stimulus onset asynchronies. For the discrimination task the SOA required to achieve 95% accuracy increased linearly with the number of targets; the mean slope was 16.6 msec per target element. (At the end of the paper it was briefly mentioned that a similar set of results had been obtained in the color domain.) In striking contrast to the studies discussed earlier, Sagi and Julesz (1985a) concluded, on the basis of these findings, that discrimination, even of highly discriminable single features, is not a spatially parallel process, but is instead accomplished serially. It is important to recognize that their findings about discrimination cannot be dismissed on the grounds that orientation discrimination is difficult and cannot be handled preattentively. There is, in fact, a substantial body of evidence indicating that the required discrimination is a particularly easy one (e.g., Olson & Attneave, 1970; Beck & Ambler, 1972; Pomerantz & Sager, 1976). The procedures employed by Sagi and Julesz differ in possibly important respects from those used in most of the experiments mentioned above. It is important to note that the inclusion of the background diagonals may change the nature of the task. In an ordinary discrimination task the stimuli would contain just horizontals and/or verticals; the subject can respond different as soon as any difference in orientation is noted. When background line segments are introduced it is no longer possible to respond different as soon as a difference in orientation is noticed because all targets differ from the background in orientation. We presume it is for this reason that Sagi and Julesz say that their task requires identification of individual targets.

One of the chief purposes of the empirical research described below is to attempt to clarify the reason(s) for the disparity between the findings of Sagi and Julesz and those of the other investigators mentioned earlier.

Conjunctively defined targets

There is a well-known series of studies by Treisman and her colleagues that have been interpreted as demonstrating serial processing when subjects search for a target defined as a conjunction of features (e.g., Treisman & Gelade, 1980; Treisman, Sykes, & Gelade, 1977). Thus if targets are defined as red squares and the nontargets are a mix of red circles and blue squares, reaction time (RT) will increase markedly with number of items in the display. The increase for negative (target absent) RTs is about twice as great as for positive (target present) RTs, a finding which implies that the search is self-terminating.

Treisman's account of these results (and of the finding of parallel processing for simple features) is known as feature-integration theory. The basic idea is that, "...features are registered early, automatically, and in parallel across the visual field, while objects are identified separately and only at a later stage, which requires focal attention. We assume that the visual scene is initially coded along a number of separable dimensions, such as color, orientation, spatial frequency, brightness, direction of movement. In order to recombine these separate representations and to ensure the correct synthesis of features for each object in a complex display, stimulus locations are processed serially with focal attention (Treisman & Gelade, 1980, p. 98)."

The finding that RT for conjunctively defined targets is strongly dependent on display size is easy to replicate (e.g., Quinlan & Humphreys, 1987). However, in view of recent findings the interpretation of this finding is not as straightforward as when feature-integration was first proposed. Pashler (1987) found that the 2:1 ratio of negative to positive slopes is due to a range effect involving display size. When small display sizes (up to eight elements) are used the display-size functions are parallel (see also Houck & Hoffman, 1986 for a similar result). This might suggest that search is serial and exhaustive, but additional data (a reduction of positive RT when redundant targets were added to displays) suggests that search is not exhaustive here. The data appear to be consistent with a limited-capacity parallel processing model of the self-terminating variety. Pashler proposes that when large displays are presented subjects may search in stages, searching clumps of up to eight items at a time. The parallel self-terminating within-clump searches are embedded in a molar serial search process.

There are some additional problems for feature-integration theory. For one, not all combinations of (seemingly separable) features yield the expected result. Nakayama and Silverman (1985) examined performance with the dimensions of motion and color; RT increased with display size. However, they also found that when one of the dimensions in a conjunctive search is stereo disparity and the other is either color or motion, RT does not increase with display size. Another problem is raised by the results of a study by Egeth, Virzi, and Garbart (1984) who showed that subjects do not search unselectively and serially through display elements in an effort to find the target; there is evidence that they search through subsets of stimuli and ignore other subsets. Thus whole groups of stimuli may be rejected in parallel (see also Treisman, 1985).

Finally, there are two reports that it is possible to correctly conjoin features even under circumstances where it is reasonable to assume that focal attention has not been brought to bear (Allport, Tipper, & Chmiel, 1985; Houck & Hoffman, 1986).

A simple summary statement seems imprudent at this time; search for conjunctively defined targets seems more complex than is suggested by Treisman's feature integration theory, but it is not yet clear what new theory will prevail.

Semantic tasks

The question at issue in this section is whether or not it is possible to apprehend the meanings of two or more stimuli in parallel. More specifically, our inquiry is focused on what should be the most demanding stimulus domain--alphanumeric symbols and words. When one deals with natural stimuli from the real world it is at least possible that meaning is carried directly by just a few perceptual attributes. For example, teeth, claws, and thorns are all sharp; this may directly signify "danger" or some other such meaning. If that is the case, then the principles discussed in the preceding sections may possibly suffice to account for the ability or inability to process such meanings preattentively. However, with verbal stimuli this is not the case. In general, there are no perceptual features that distinguish among categories of words or other alphanumeric symbols. Thus, the words "tooth," "claw," and "thorn" are no more alike than, say, "trout," "crow," and "hedge." Moreover, fairly detailed visual processing is required to ascertain the meaning of a word (compare, e.g., "arc" and "are"). For these reasons it is difficult to believe that parallel processing of semantically defined targets is possible. Be that as it may, there are at least four lines of evidence that lead us to suspect that discriminations beyond the level of single features can be made in parallel.

For one, Egeth, Jonides, and Wall (1972) gave subjects the task of detecting the presence of a single target defined as any digit in a background consisting of several different letters. Reaction time to detect a target was unaffected by the number of nontargets, which suggests that display characters were examined in parallel. The digit-letter classification is not unique; Schneider and Shiffrin (1977) have shown that after extensive consistent-mapping practice search for arbitrary target sets is similarly unaffected by the number of nontargets in a display.

A second line of evidence is provided by Pashler and Badgio (1985). They devised a task that required exhaustive processing to the point of identification. Subjects had to name the highest digit present in an array of digits. The effect of display size was additive with the effect of visual quality (clear vs. degraded). This result implies that subjects did not serially examine the displays to find the highest digit; if they had then the effect of visual quality would have interacted with display size as each comparison would have been slowed down by the same amount in the degraded condition. As it stands, however, the data might only indicate that a very early stage involved perhaps in the "clean up" of degraded stimuli was executed in parallel. This interpretation is contraindicated by the additional finding that visual quality interacted with response factors (e.g., the identity of the highest digit). The authors argue that the overall pattern of results is consistent with parallel encoding of the entire display.

Third, there are some interference effects that occur at the semantic level that might be taken as evidence of parallel processing. A suggestive example comes from a paradigm first developed by Bjork and Murray (1977). They demonstrated the existence of perceptual interference effects between simultaneously presented letters that varied as a function of the similarity of the letters (see also La Heij & van der Heijden, 1983). Maximum similarity, created by repetition of the same letter (e.g., AA) led to the poorest performance. Egeth and Santee further showed that repetition at the name level (e.g., Aa) also led to interference. While not definitive, this finding is at least compatible with the hypothesis that interference is due to similarity of meaning. Moreover, exposures are brief, typically around 50 msec., and followed by a mask; these conditions may make parallel processing more likely.

Additional relevant evidence comes from Stroop-like interference effects. In the usual Stroop interference condition subjects name a color while trying to ignore the meaning of a color word. Both the color and the word carry meaning; the interference effect can thus be construed as evidence for parallel processing of meaning.

Finally, Mozer (1983) and McClelland and Mozer (1986) have explored interactions, specifically letter migrations, between simultaneously presented strings of letters. A stimulus might consist, e.g., of the side by side presentation of SAND and LANE, which might then be followed by a poststimulus cue indicating which of the two words to report. Subjects make a variety of errors.

Especially interesting are migration errors. Thus, in the example above, when cued to report the word on the left they might respond "sane" or "land." The frequency of such errors depends on whether the letter strings form words or not. The explanation offered for the pattern of results assumes both strings simultaneously access high-level structural knowledge about what sequences of letters form familiar words.

Although these lines of evidence suggest parallel processing of semantic information is possible, none of them is completely convincing. As Pashler and Badgio (1985) pointed out, the problem with studies like that of Egeth, Jonides, and Wall (1972) is that they do not require the processing of non-target items to the point of identification. Instead, they permit a response upon detection of the target. Thus conclusions about parallel processing at the level of category or meaning are inappropriate. If we consider the Pashler and Badgio (1985) study, although it's important as far as it goes, it does not speak to the ability to process words in parallel. The problem with interference studies such as those based on the Stroop effect is that although they do establish that mandatory processing of irrelevant information sometimes occurs they are not designed in such a way as to permit the conclusion that relevant and irrelevant material were processed simultaneously and independently. Indeed, Kahneman and Chajczyk (1983) accounted for data from a Stroop task with a capture model which assumed that subjects attended to just one word at a time. Finally, the theoretical interpretation offered for the Mozer (1985) and McClelland and Mozer (1986) studies appears to implicate parallel processing, but only for lexical status (i.e., whether a string of letters is a word or a nonword). It is not clear that word meaning is processed in this way. In view of the inconclusiveness of the existing literature, one of the aims of the empirical research reported in the next section is to come up with an alternative research design that may afford a cleaner test of processing mode with verbal stimuli.

III. Empirical Research

In this section we present a series of nine experiments concerned with the proposal of Sagi and Julesz (1985a,b) that identification of features requires serial processing. We then present a series of three experiments designed to explore whether word meaning can be processed in parallel from more than a single spatial location.

Experiment 1: Discrimination vs. Identification

It is possible that the difference between the Sagi and Julesz findings and the earlier findings is due to the fact that the stimuli in those studies did not contain anything comparable to the diagonal texture elements of Sagi and Julesz. There are several possible ways in which texture might affect performance. For example, the diagonal lines may to some extent stimulate detectors whose tuning curves are centered on horizontal and vertical, thus using up processing capacity that might otherwise be devoted to target elements. This could lead to the increase in processing time with increasing numbers of targets, possibly by inducing serial processing as Sagi and Julesz (1985a) suggest. Another possibility is that diagonal texture elements change the task from one that can be solved by the detection of any feature difference into one that requires identification of specific values of orientation.

The first experiment was conducted in an effort to determine simply if the addition of a textured background to a normal same-different task does indeed result in a qualitative change in performance. The basic design of the Sagi and Julesz discrimination task experiment was recast in a fairly standard "chronometric" paradigm; RT was measured after a 150 ms exposure that was not preceded or followed by any kind of mask and trials were mixed rather than blocked with respect to target numerosity. Displays with and without texture-forming background elements were presented to subjects in separate blocks.

Method

Subjects. Eight paid volunteers were tested initially. Two of these subjects produced data that were markedly slower and more variable than that of the other subjects in this or in any of the subsequent experiments. These two subjects were replaced by two new subjects.

Stimuli. Displays consisted of line segments placed in the imaginary cells of a 6 x 6 matrix that subtended 7 cm vertically and horizontally. Line segments were 8mm long and had a stroke width of approximately 0.7mm. They were drawn in black ink on white cardboard cards with a Digital LVP16 plotter. At the viewing distance of 91.5 cm. the matrix subtended approximately 4.37 deg of visual angle horizontally and vertically, and the line segments subtended approximately .44 deg.

Targets were defined as line segments that were either horizontal or vertical; they were distributed randomly among the 36 cells of the matrix with the constraint that no two targets could appear in cells that were either horizontally, vertically, or diagonally adjacent. Target numerosity was varied randomly from trial to trial. On any given trial two, three, or four targets could appear with equal probability. On half of the trials in each background condition (i.e., no-texture and texture) the targets had the same orientation (all vertical or all horizontal); on the other half of the trials one target differed in orientation from the others (e.g., three vertical and one horizontal). On half of the same trials the targets were vertical; on the other half of the same trials the targets were horizontal. On half of the different trials the discrepant target element was vertical; on the other half of those trials it was horizontal.

In the texture condition cells that did not contain targets contained a negatively sloped diagonal line segment. In the no-texture condition cells that did not contain a target were left blank. Sample stimuli are shown in Figure 1.

Insert Figure 1 about here

Procedure. Displays were presented for 150 ms. in a Gerbrands four-channel tachistoscope. The subject's task was to indicate whether all of the targets had the same orientation or if one differed from the others by pressing with their index fingers one of two appropriately labelled buttons. Reaction time was recorded to the nearest millisecond. The assignment of response buttons to hands was balanced across subjects.

A fixation point was visible at the center of the field whenever a stimulus was not being presented. The sequence of events on a trial was as follows. The experimenter initiated a trial manually. One-half second after initiation of a trial a warning tone sounded for 250 ms; 500 ms after the offset of the tone the stimulus was presented for 150 ms.

The two background conditions were blocked and the order of the conditions was balanced across subjects. In each condition subjects received 16 practice trials followed by two sixty-trial blocks of experimental trials. Thus, in both the no-texture and texture conditions there were 20 same and 20 different trials for each of the three levels of target numerosity (2, 3, and 4). There were short breaks between blocks and between conditions.

Results and Discussion

Mean RTs and error rates are shown in Figure 2. For each condition the function relating reaction time to target numerosity was subjected to a trend analysis the results of which are presented in Table 1 along with the best fitting slopes, intercepts and the percentage of variance attributable to the linear component.

Insert Figure 2 and Table 1 about here

As indicated in Table 1, none of the functions relating mean RT to target numerosity had a slope that differed significantly from zero. Analysis of variance of the reaction time data showed that the 133 ms main effect of background (no-texture vs. texture) was significant, $F(1,7) = 338.35$, $p < .001$, as was the 40 ms main effect of trial type (same vs. different) trials, $F(1,7) = 9.18$, $p < .02$. Neither the main effect of target numerosity (2, 3 or 4) nor any of the interactions were significant.

The mean error rate across conditions was less than 8%. There does not appear to be much evidence of a speed-accuracy tradeoff in these data, as speed and accuracy were positively correlated across conditions. Specifically, the correlation of mean error rate and mean RT was $r(10) = .66$, $p < .05$. In other words, conditions with longer mean RTs had higher error rates.

This first experiment indicates that although the presence of a textured background significantly slows overall reaction time, there is no suggestion that it induces serial processing of display elements. The only hint of such an effect is that in the texture condition the slope on different trials was 13.3 ms per target element--however this slope was not significantly different from zero.

Experiment 2: A Variation in Response Mode

Experiment 1 suggests that texture alone can not explain the discrepancy between the results of Sagi and Julesz (1985a) and those of other investigators. However, our results were not entirely compelling. For one thing we felt obliged to replace two seemingly aberrant subjects. For another, even though a slope of 13.3 ms was obtained in one condition it was not significant. In the present experiment an attempt was made to reduce the variability in the data by requiring a go-no go response. Subjects were instructed to respond when all targets were the same but to refrain from responding when one target differed in orientation from the others. (The same response would appear to require exhaustive processing, and hence would be more likely to show an effect of target numerosity than would different responses.)

Method

Subjects. Eight students at the Johns Hopkins University participated in this experiment.

Procedure. The experiment was essentially the same as Experiment 1 with the following exceptions. Subjects were instructed to respond only on same trials by pressing a button with the index finger of their preferred hand. In both the texture and no-texture conditions there were 16 practice trials followed by two blocks of 75 experimental trials. Of the experimental trials 60% (45) in each block were same and 40% (30) were different. Thus in both the texture and no-texture conditions there were 90 same trials, 30 for each number of target elements (2,3, and 4) and 60 different trials, 20 for each number of target elements.

Results and Discussion

Mean RTs and error rates are shown in Figure 3; results of trend analyses and summary statistics for the best-fitting straight lines for the different conditions appear in Table 1.

Insert Figure 3 about here

It is worth noting that our effort to reduce variability by introducing a go-no go response seems to have worked; the mean within-condition standard deviation was 89.2 ms in the present experiment, down from 109 ms in the first experiment, $t(14) = 2.36$, $p<.025$.

In the presence of a texture the function relating mean RT to number of targets had a nonsignificant negative slope of 16.1 ms per item. With no texture the 9.3 ms per item slope was positive and significant. The reaction time data were subjected to a within-subjects analysis of variance with background (no-texture vs. texture), and target numerosity (2, 3 or 4) as factors. The only clearly significant effect was the 135 ms. main effect of background, $F(1,7) = 49.10$, $p <.001$. However, the interaction between target numerosity and background almost achieved a conventionally acceptable level of significance, $F(2,14) = 3.58$, $p < .06$.

The mean error rate across conditions was approximately 6%. There is little indication that subjects traded speed for accuracy as the correlation of mean error rate with mean RT was positive, although not significantly so, $r(4) = .74$, $p.>.05$.

This experiment, just like Experiment 1, showed that texture produces an increment in overall reaction time, but again there was no indication that the presence of a textured background causes serial processing of target elements. If anything, there is a hint in the data that texture may produce a negative slope in the function relating same reaction time to target numerosity. (In other experiments we have, in fact, observed significant negative slopes in this condition.) It is difficult to know what to make of the significant positive slope in the no-texture condition. Including pilot work, we have examined approximately 20 functions relating RT to number of targets in no-texture conditions. We have found one significant positive slope (the present study), one significant negative slope, and the rest near zero and nonsignificant. We suspect that the true value is close to zero; the present result may well be Type I error.

Experiment 3: Range Effects in Discrimination

The results of Experiments 1 and 2 make it clear that the difference between the results of Sagi and Julesz (1985a) and those of previous studies is not due simply to the use of textured backgrounds. We turned our attention next to another aspect of their experiment that differed from what might be considered standard chronometric procedures. The number of target elements (2, 3, and 4) was blocked in the Sagi and Julesz experiment. In most studies in which reaction time is the dependent variable of chief interest display size is normally randomly mixed within blocks of trials. Egeth, Jonides, and Wall (1972, pp.688-690) have previously demonstrated that such a design difference can have a major impact on the results of visual search experiments. Specifically, they found that when display size varied randomly from trial to trial reaction time on target-present trials was virtually independent of display size. However, when display size was blocked (actually, different size displays were presented to different subjects) reaction time increased significantly with display size. For a more complete discussion of range effects see Poulton (1982).

In the present experiment we used the standard same-different discrimination task (i.e., no textured background) but with target number blocked (within subjects).

Method

Subjects. Six subjects served in this experiment. The stimuli were the same as in Exper. 1.

Procedure. Each subject was given three blocks of trials, one for each of the three levels of number of targets (i.e., 2, 3, and 4). The order of these conditions was determined by a 3 x 3 Latin Square that was replicated once. (The natural design here with six subjects would be to use all possible orders of the three conditions. A Latin Square was used because the experiment being reported here was embedded in a larger experiment. The results from the other conditions in that larger experiment will be reported elsewhere.)

Each block of trials consisted of 16 practice trials followed by 60 experimental trials, half same and half different. There were short breaks between blocks. Subjects indicated same with their preferred hand and different with their other hand.

Results

Mean RTs and error rates are shown in Figure 4; results of trend analyses and summary statistics for the best-fitting straight lines for the different conditions appear in Table 1.

Insert Figure 4 about here

The difference between same and different RTs was not significant at the .05 level, $F(1, 5) = 4.67$. More important, neither the number of targets nor the interaction of same-different and number of targets was significant, both Fs < 1.0. We conclude that the independence of mean RT and display size found when simple features are discriminated (e.g., Donderi & Zelnicker, 1969) obtains when display size is blocked as well as when it is mixed.

Experiment 4: Range Effects in Identification

We turn now to a chronometric version of the Sagi and Julesz discrimination task in which number of targets is blocked as in their original experiment.

Method

Subjects. Twelve subjects participated in this experiment.

Stimuli. Stimuli were constructed in the same manner as in the first three experiments. All displays contained background texture elements.

Procedure. As in Experiments 1 and 3 (but unlike Experiment 2) two buttons were provided for the response, one for same trials the other for different trials. The subject's preferred hand was always used to indicate that targets were the same. The temporal sequence of events on individual trials was identical to that used in the first two experiments. Subjects were presented with three blocks of 60 trials. Within a given block, same and different trials were equiprobable. Each block was preceded by 16 practice trials. There were short breaks between blocks.

The number of target elements was constant throughout a 60-trial block (and its concomitant set of practice trials). The order in which the three levels of target numerosity (2, 3, or 4) were presented was balanced across subjects, as all possible orders were used equally often.

Results

Mean RTs and error rates are shown in Figure 5; summary statistics for the best-fitting straight lines and trend analyses appear in Table 1.

Insert Figure 5 about here

Both the same and different functions are positively sloped. However, while the 37.1 ms/element slope for the different function differs significantly from zero, $F(1,11) = 6.69$, $p < .05$, the variability in the data was sufficient to render the 24.8 ms/element slope of the same function not significantly different from zero, $F(1,11) = 2.87$, $p > .05$.

Trial type did not interact with target numerosity, $F < 1$. Thus the slopes of the same and different functions did not differ significantly when target numerosity was blocked.

The mean error rate across conditions was just over 3%. Again there is little evidence of a speed-accuracy tradeoff as the correlation of mean error rate and mean RT was $r(10) = .76$, $p<.01$.

The results of this experiment suggest that we have been able, under certain conditions, to obtain results similar to those of Sagi & Julesz using a reaction time paradigm. In the presence of texture, with blocked presentation mode, RT increased directly with target numerosity. (This was clearly the case for different trials; the slope for same trials, although substantial, did not differ significantly from zero.)

Discussion

Despite the linearity of the relation between RT and number of targets (at least for different trials), it is not clear that our results unambiguously support the contention that the identification of orientation required the serial allocation of focal attention. It has been pointed out that increases in reaction time with display size are not a unique signature of a serial process (Townsend, 1971; Snodgrass & Townsend, 1980). Other models of this same/different search task could give rise to increases in reaction time with display size. For example, a parallel limited-capacity search, or a parallel unlimited-capacity search that is exhaustive (and that has nonzero variances for its compact subprocesses), could give rise to increasing RT functions.

Experiment 5: A Further Test of Serial Processing

The purpose of this experiment was to test for evidence of spatially serial processing in feature identification. More specifically, it tested the hypothesis, proposed by Sagi & Julesz (1985a), that the variation in performance as a function of target number with textured stimuli is due to serial allocation of focal attention.

The design and logic are similar to that used in the study by Pashler & Badgio (1985) that as described earlier (cf. Background). In their experiment, Pashler and Badgio attempted to distinguish between parallel and serial models of alphanumeric character identification. They had subjects perform detection tasks in which the size (i.e., the number of stimuli) and the visual quality (e.g., contrast) of the display were varied factorially. The authors reasoned that if identification of alphanumeric characters requires a serial, element by element scan, then any increase in processing time due to contrast reduction should be added to each item in turn, resulting in a multiplicative interaction of visual quality with display size. If, however, the identification process is parallel, then visual quality should be strictly additive with display size. The results indicated a highly significant main effects of display size and visual quality, but no interaction between the two factors. These results could not be accounted for by a serial model and were consistent with a parallel, limited-capacity model of character identification.

As in Experiment 4, subjects in the present experiment made same-different decisions about the orientations of horizontal and vertical line segments embedded in diagonal texture elements, with number of targets a blocked variable. In addition to varying the number of targets, the visual quality of the displays was manipulated by superimposing visual noise (a dot mask) on half of the trials. If the increase in reaction time between two and four targets observed in Experiment 4 (at least on different trials) is due to the serial encoding of feature identity, and visual noise affects the efficiency of this encoding process, then the effect of visual quality should be overadditive with display size, yielding an interaction between those two variables.

Method

Subjects. Eight subjects participated in this experiment.

Stimuli and Apparatus. For this experiment displays were similar to those of the preceding experiments, however they were generated by an IBM AT computer and displayed for 200 ms. on a Hewlett Packard 1345A digital display module with a P31 phosphor. The display scope was mounted into one channel of an Iconix four channel tachistoscope. A second channel in the tachistoscope provided a constant, dim, background illumination. The computer collected all reaction time and error data.

Displays consisted of line segments placed in the imaginary cells of a 8 x 8 matrix that measured 7.5 cm vertically and horizontally. Line segments were 6mm long and had a stroke width of approximately 0.7mm. At the viewing distance of 68 cm the matrix subtended approximately 6.86 deg of visual angle horizontally and vertically, and the line segments subtended approximately .53 deg.

Each line segment was independently displaced 3mm (0.26 degrees visual angle) from the center of its imaginary cell in one of eight randomly chosen directions (up, down, right, left, upper right, lower right, upper left, lower left). This "jitter" was added to the displays to reduce the formation of global patterns that might form from the linking of collinear line segments.

Targets were defined as line segments that were either horizontal or vertical; they were distributed randomly among the 64 cells of the matrix with the constraint that no two targets could appear in cells that were either horizontally, vertically, or diagonally adjacent. Each of the 64 cells that did not contain a target contained a negatively sloped diagonal line segment. Displays were generated "on line" and thus no two subjects saw exactly the same set of displays. On _same_ trials all targets had the same orientation: half the _same_ trials contained vertical targets, half contained horizontal targets. On _different_ trials one target differed in orientation from the rest: half the _different_ trials contained a discrepant target that was vertical, half contained a horizontal discrepant target.

Procedure. As in previous experiments, the subject's task was to determine if the non-diagonal targets in a display were all the same orientation or not. Only two levels of target number were used--two and four. Subjects responded _same_ with the index finger of their dominant hand, and _different_ with the index finger of their non-dominant hand.

Both visual quality and target numerosity were blocked. Half the subjects received the masked condition first and the unmasked condition second, half the opposite order. Within each visual quality condition, subjects received four blocks of trials. Each block consisted of 12 practice trials and 60 experimental trials. There were two blocks at each level of target numerosity. The two blocks of each target number were always contiguous (e.g., 2244 or 4422, but not 2424). The order of target number was counterbalanced across subjects.

Results

Mean correct reaction time and error rates for _same_ and _different_ responses at each level of target number and visual quality are shown in Figure 6. Mean correct RTs for each

---

Insert Figure 6 about here

---

subject were entered into a within-subjects analysis of variance with response type (same vs. different) target number (two vs. four) and visual quality (masked vs. unmasked) as factors. The dot mask proved effective in influencing processing, as the 92 ms main effect of visual quality was highly significant, $F(1,7) = 20.31$, $p.<.005$. The main effect of response type was also highly significant, $F(1,7) = 49.20$, $p.<.001$, indicating that same responses were faster than different responses. Finally, responses to displays with four targets took significantly longer than responses to displays with only two targets, as the main effect of target number was highly significant $F(1,7) = 45.60$, $p.<.001$. The only interaction that approached significance was that of response type by target number, $F(1,7) = 5.0$, $p<.06$, reflecting the fact that across the two levels of visual quality different responses were affected to a greater extent by an increase in target number than same responses.

Similar analyses were performed on the error data. The results revealed a significant main effect of visual quality, $F(1,70 = 10.07$, $P.<.02$, a marginal effect of response type, $F(1,7) = 4.78$, $p.<.07$, and a significant interaction between response type and target number, $F(1,7) = 9.85$, $p.<.02$. Overall, there is little indication of a speed-accuracy trade-off as error rates were positively correlated with reaction time, $r = 0.87$.

Discussion

The results of this experiment are important in two respects. First, they replicate the results of Experiment 4. With target number blocked, reaction time to respond same or different increased with target number. Moreover, the effect of target number is again greater for different trials than for same trials. Second, the lack of interaction between target number and visual quality suggests that if masking the stimulus did indeed affect the extraction of target identity, then this process proceeds in parallel. If the increase in reaction time with target number were due to a serial process, then the effects of visual quality should have been overadditive with target number. There is no hint in the data of such overadditivity. Thus, the results of the present experiment suggest that feature identification does not involve a serial process as Sagi & Julesz have contended.

This result, standing alone, cannot be considered definitive, as one could argue that the mask affected a stage of processing operating prior to the processing of feature identity. However, one model of feature identification that can be unambiguously rejected on the basis of the present results is one in which focal attention operates serially on the actual degraded display elements or some low-level, veridical representation or "primal sketch" (Marr, 1981) of the degraded display.

Experiment 6: Redundancy Gain in Identification

Experiment 5 suggests that identification can be carried out in parallel. However, as mentioned above, the results cannot be considered definitive. This suggests the need for a fresh approach to the question of whether feature identification is a serial or parallel process. An approach that recommends itself is the redundancy gain paradigm mentioned earlier (e.g., Van der Heijden, La Heij, & Boer, 1983, among others).

The present experiment was modeled closely after the research of Van der Heijden et al. (1983) in which subjects had to discriminate Es from Fs; in the present study horizontal and vertical line segments take the place of Es and Fs. A possibly important difference between the Van der Heijden et al. (1983) study and the present one is that we are requiring subjects to distinguish between two levels of a dimension, whereas in the earlier research subjects might have been responding on the basis of the presence or absence of a feature (cf. Garner 1978). Note also that in the present experiment there are no diagonal background elements. This permits a valuable simplification of the subject's task. Recall that it was argued that the background texture may force identification of stimuli in Sagi & Julesz's version of a same - different task because one cannot respond simply on the basis of whether or not a feature gradient is detected. For all its simplicity the present task shares that virtue. No display contains a "feature gradient" as all of the stimuli in a display are identical in orientation; correct response requires identification of that orientation.

Method

Subjects. Eight subjects participated in this experiment.

Stimuli and apparatus. The equipment was the same as in Exp. 5, as was the timing of the events on each trial. On each trial the stimuli were one, two, or three vertical line segments or one two, or three horizontal line segments. The centers of the line segments were located on an imaginary circle 1.3 deg. in diameter. The locations used were at 12:00, 4:00, and 8:00.

Procedure. There were six blocks of 54 trials each. The first block was considered practice. In each block half of the trials contained only vertical stimuli, the other half contained only horizontals. There were equal numbers of trials with one, two, or three stimuli. Stimuli appeared equally often at each of the three possible locations equally often in each condition. Stimuli were randomized within blocks and thus all factors were mixed rather than blocked.

Half of the subjects were instructed to respond with a dominant-hand button press when vertical lines were shown and to refrain from responding when horizontal lines were shown; the other half of the subjects were instructed to respond only when horizontal stimuli were displayed. (Reasons for the use of a go-no go response are provided by Van der Heijden et al., 1983.)

Results

The mean RTs for one, two, and three targets were 451,435, and 410 msec., respectively. These means differed significantly, $F(2, 14) = 87.44$, $p < .001$, which suggests that processing of displays was parallel and of unlimited capacity. No analysis of errors could be carried out, as there were only five errors in the entire experiment.

It is not sufficient to simply compare overall mean RTs, as there are certain artifacts that can produce a similar pattern of results. Suppose, for example, that for each subject there is a particular favored position in the display that is processed more quickly than the others, perhaps because it is inspected first in a serial scan. The greater the number of targets the greater the probability that one of them will be in the favored position and thus the faster the mean RT. (Similar artifacts have been dealt with in other domains; see e.g., Biederman and Checkosky, 1970 and Santee and Egeth, 1982.)

If the obtained redundancy gain exceeds that expected on this artifact then the mean RT on three-target trials should be faster than the mean RT for the fastest position (determined separately for each subject). However, as Van der Heijden et al. (1983) point out, this comparison is biased against finding a redundancy gain. It is appropriate only if there is a fixed favored position. To determine if this analysis was appropriate we analyzed the data for single-target trials. Specifically, for each subject we compared the mean RT for his or her fastest position with the mean RT for the fastest subset resulting when trials were divided randomly into three subsets. There was no significant difference, which we take to be direct evidence that there was no fixed favorite position. (In other words, the data suggest that the "fastest position" was not really any faster than one would expect on the basis of random variability.)

An even more insidious artifact is possible; there may be a favored position that is not fixed, but varies randomly across trials. For displays with a single target that target will be in the favored position on about one third of the trials. On this fraction of trials a relatively fast response can be expected. The greater the number of targets the greater the probability that one of them will be in the favored position (whichever that happens to be on a given trial) and thus the faster the mean RT. An analysis suitable for testing this artifact was suggested by Van der Heijden et al. (1983). They note that for displays with three targets there is a target in the favored position on every trial, and these trials should not contain any particular subset of fast trials. Thus if the RTs in that condition are ordered from slowest to fastest and then divided into three equal-sized subsets (slowest third, intermediate third, and fastest third) then the differences among the means of those three subsets should simply reflect random variation. For the one-target displays the variability among the means might or might not reflect just trial-by-trial variability. The alternative is that the fast subset is fast because of the random favored position artifact. If that were the case there ought to be more variability among the subset means for one-target than for three-target displays.

By way of analysis, for each subject we ordered the 45 RTs in the one-target condition from slowest to fastest and then partitioned them into three equal sized subsets and calculated the means of these subsets. We did the same for the three-target displays. These means are shown in Table 2.

Insert Table 2 about here

The difference between the mean RTs of the one-target and three-target trials is significant even when just the fast subsets are considered, $F(1,7) = 92.41$, $p < .001$. However, the difference increases as we move from the fast subset through the medium subset to the slow subset, $F(2, 14) = 15.96$, $p < .001$. This pattern is consistent with the random favored-position artifact. Thus, while the redundancy gain observed in this experiment may well be real (the significant and large effect for the fast subset data is suggestive here), we cannot rule out the possibility that the data indicate only a random favored-position artifact.

(We have done a Monte Carlo simulation of the parallel unlimited-capacity model. Under conditions like those of the present experiment, we found that RTs may diverge as in the top section of Table 2 even when no favored position is assumed. For this reason we offer the analysis based on the work of Van der Heijden et al. [1983] somewhat tentatively; it is not clear to us that an ideal test of the random favored-position artifact has been devised yet.)

The results of this study are somewhat ambiguous because of the possibility that the data can be explained in terms of a random favored-position artifact. However, in view of the results of the preceding experiment and the lack of any clear evidence for a random favored-position artifact in the following experiments, our best guess is that feature identification is a spatially parallel process that has unlimited capacity. One might well wonder if the results would be different if we had included the diagonal texture elements used by Sagi and Julesz (1985a,b) and in several experiments in the present report. We replicated the above experiment but with texture elements instead of a blank background. The results were essentially identical.

Discussion

One issue that is in need of clarification is the nature of the differing capacity limits suggested by Exps. 5 and 6. The redundancy gain in Exp. 6 suggests unlimited capacity, while the increase in RT with target numerosity in Exp. 5 suggests limited capacity (although not so limited as to require serial processing). One possible resolution is suggested by Folk's (1987) claim that the RT increase found in a discrimination task when target numerosity is blocked and a textured background is present, as in Exp. 5, is due to decision-level effects. In brief, the argument is that subjects solve the discrimination task by comparing the amount of activity in horizontal and vertical "maps" (the map terminology is borrowed from Treisman & Souther, 1985, among others). Each horizontal and vertical target contributes activation to its corresponding map. Applying the principle of coarse coding (e.g., Hinton, McClelland, & Rumelhart, 1986), it is further assumed that each diagonal contributes some relatively small degree of activation to both of those maps (as well as a much larger contribution to the diagonal map, of course). When two targets are present, there should be equal activity in the horizontal and vertical maps on different trials and unequal activity in those maps on same trials. Reaction time depends on how long it takes the subject to compare the activation in the two maps. When four targets are present the level of activity in the horizontal and vertical maps is unequal for both same and different trials. (Recall that a different trial consisted of three stimuli of one orientation and one of the other orientation.) Reasonable assumptions about the amount of activation contributed to the two maps by target and texture elements lead to the expectation that the difference in activation levels between same and different trials should be harder to resolve when there are four targets than when there are two targets, thus accounting for the increase in mean RT as target numerosity increases. The reader is referred to Folk's (1987) thesis for a fuller description of the decision-level account. The model accounts for the independence of mean RT and target numerosity when there is no background; it also accounts for the absence of increasing functions when target numerosity is mixed. Nevertheless, at this point his model is qualitative; further quantification is necessary.

Experiment 7: Increased Target Numerosity

This experiment represents our initial effort to explore the same-different discrimination task with target numbers greater than four. In this study, instead of target numbers 2, 3, and 4 we used 2, 4, and 6.

Method

Subjects. Subjects were twelve undergraduate students at the Johns Hopkins University ranging in age from 17 to 21. All participated to fulfill a course requirement.

Stimuli and Apparatus. The equipment was the same as in Exps. 5 and 6 (computer-driven oscilloscope). The displays were of the same type but were always shown without any mask.

Procedure. The subject's task was to indicate whether all of the targets had the same orientation or if one differed from the others by pressing one of two appropriately labelled buttons on a response box placed on the table in front of the subject. The left and right index fingers were used to make the response. Subjects were instructed to respond as quickly as possible, while minimizing errors. Reaction time was recorded to the nearest millisecond. Subjects responded to same trials with the dominant hand, different trials with the non-dominant hand.

The sequence of events on a trial was as follows. A fixation cross appeared in the center of the display for 500 ms. A display then appeared after a variable blank interval of 500 to 1500 ms and remained on for 150 ms. At this exposure duration, the possibility of eye movement was unlikely. If a response error was made, the computer beeped to let the subject know. All trials on which errors were made were followed by a "buffer" trial that was *not* counted in the data analysis. If no response was made after 1500 ms, the trial was counted as an error. The fixation cross for the next trial appeared after a 2 second intertrial interval.

Procedure. Subjects were presented with three levels of target number (2, 4, and 6) in separate blocks. There were a total of 120 trials for each level of target number, broken into two 60-trial blocks. Thus, each subject completed six blocks of trials. Within each block, half the trials were same trials and half were different. The order of blocks was counterbalanced across subjects. A set of twelve practice trials was presented before each block. Subjects were required to get at least seven of the twelve correct before moving on to experimental trials; otherwise, another 12 practice trials were run. Short rests were allowed between blocks of trials.

Results

Mean reaction time (RT) and error rates for same and different responses at each level of target number are shown in Figure 7. Mean correct RTs for each subject were entered

---

Insert Figure 7 about here

---

into a within-subjects analysis of variance with response type (same vs. different) and target number (2, 4, and 6) as factors. The overall advantage of same responses over different responses was significant; for the main effect of response type $F(1,11) = 8.64$, $p<.01$. This "fast-same" effect is a fairly typical result and is generally attributed to a conservative response criterion on different trials that guards against false alarms based on the detection of "spurious differences" (Krueger, 1978). The overall effect of target number just failed to reach the conventionally accepted level of significance, $F(2,22) = 3.08$, $p<.06$. The interaction of response type with target number was significant, $F(2,22) = 3.68$, $p<.05$.

The data for same and different functions were separately subjected to orthogonal trend analyses. The inverted "v" trend for the different function was confirmed by a significant quadratic component, $F(1,11) = 5.65$, $p<.05$. The quadratic trend for the same function, however, failed to reach significance, $F(1,11) = 2.26$, $p>.05$. Neither function contained a significant linear trend (for same and different functions, $F(1,11) = 1.52$, $p>.05$ and $F(1,11) = .38$, $p>.05$, respectively).

As is evident in the figure, error rates tended to mimic reaction time suggesting that subjects did not trade accuracy for speed. This is substantiated by a positive correlation of reaction time with error rate, $r(4) = .76$, $p<.05$. Mean error rates for subjects were entered into a within-subjects analysis of variance. Neither of the main effects were significant, $F(1,11) = 3.37$, $p>.05$; $F(1,11) = 2.66$, $p>.05$, for response type and target number, respectively, nor was their interaction, $F(1,11) = 3.19$, $p>.05$.

Discussion

The results of this experiment successfully replicate the our earlier findings (e.g., Exps. 4 and 5) with displays containing up to four targets, but show that the effect is not generalizable to larger target numbers. At larger target numbers, reaction time is actually reduced, suggesting a qualitative change in processing strategy.

Target numbers 2 - 4. The results of the present experiment, with displays containing 2 or 4 targets are similar to those obtained in Exp. 4. The slopes of the functions in the two experiments are comparable: 28.5 and 42.5 ms/target for same and different, respectively, in the present experiment, versus 24.5 and 37.5 ms/target for the analogous functions in the earlier experiment. Moreover, as in Exp. 4, only the increase in reaction time for the different function in the present experiment was significant, $t(11) = 2.38$, $p<.05$; for the same function $t(11) = 1.95$, $p>.05$. Thus it appears that different responses are affected to a greater degree by target number than <u>same</u> responses.

To provide a direct test of the differential effect of target number on response type, an interaction comparison was performed on the data for just the two and four target conditions. The interaction between response type and target number just failed to reach significance, $F(1,11) = 4.33$, $p<.06$. Given the similar result in Exp. 5, however, it is difficult to accept the null hypothesis in this case. If the effect is indeed real, it has important implications. Since no plausible model predicts such an interaction, the possibility arises that target number may affect a stage other than the stage responsible for feature identification. This does not, however, rule out the possibility that the feature identification stage is also affected by target numerosity. The observed interaction trend may be the result of the combined effects of target numerosity on two different stages. For example, target number may affect <u>same</u> and <u>different</u> functions equally at the feature identification stage and differentially at a decision stage.

<u>Target number 6</u>. The significant quadratic trend for the <u>different</u> function suggests that reaction time was significantly faster when six targets were present than when four targets were present. A similar trend is seen in the <u>same</u> function, although it was not significant. This reduction in reaction time from four to six targets suggests that with increasing target number, performance in the task may change qualitatively.

There are several possible explanations for a qualitative change in processing six targets present. One that we find particularly appealing is that grouping effects begin to emerge with six targets. If targets form a perceptual group, then subjects may be able to complete the task by detecting a difference (or not) in that group without having to explicitly process the identities of the individual members. In other words, the formation of a perceptual group may allow the targets to emerge from the texture elements as a unit that can then be processed as if no texture elements were present. This type of process is supported phenomenologically by many subjects who reported that when six targets were present, they seemed to "pop out" as a group from the background texture.

A similar possibility is that with increasing numbers of targets, a static analogue of an optic flow field is formed. Farell & Julesz (1986) have recently provided evidence that global directional flow of spatially distinct display elements can be extracted preattentively and independent of local information required for the recognition of individual elements. In the present experiment, with high target number, the task may be performed by noting whether there is a "break" or not in the directional flow field generated by the targets.

Clearly the possible effects of perceptual grouping or flow field generation are interesting in and of themselves and are worthy of further investigation. However, they do not shed light on the primary concern of this project which is whether visual feature identification requires limited capacity attentional resources or not.

Experiment 8: Target-Noise Similarity

The model proposed by Folk (1987), which is only sketched out in the Discussion of Exp. 6, while accounting for the dissociation between present-absent and same-different tasks observed with textured stimuli, is admittedly post-hoc and complex. The specific hypotheses are difficult to test. However, the broad notion, that the effect is specific to diagonal texture elements creating background activity in feature maps, should be testable. For example, if a same-different orientation-discrimination task were required with texture elements that do not generate background activity in either of the target maps, one would expect that, as if no texture elements were present, performance could be based simply on the presence of activity as opposed to the magnitude of the difference in activity between the two maps, yielding an invariance of reaction time with target number.

Experiment 8 was conducted to test this possibility. The diagonal line segments used as texture elements in the previous studies were replaced with circles. Assuming circles generate little or no activity in horizontal and vertical feature maps, reaction time should not vary with target number. If, however, the effect of target number in a same-different task is not specific to the nature of the texture elements, (e.g., if same-different discrimination is simply difficult with any type of textured displays) then one might expect results similar to those found in Experiments 4 and 7.

Method.

Subjects. Subjects were 12 undergraduate students at the Johns Hopkins University ranging in age from 19 to 22. All participated to fulfill a course requirement.

Stimuli, Apparatus, and Procedure. Stimuli, apparatus and procedure for this experiment were identical to that used in Experiment 7 with the following exception. On all displays, diagonal texture elements were replaced by circles, subtending .53 degrees of visual angle.

Results

Mean correct RTs and error rates for each response type at the three levels of target numbers are shown in Figure 8. The RT data were subjected to a within-subjects analysis of

---

Insert Figure 8 about here

---

variance with response type (same vs. different) and target number (2, 4, 04 6) as factors. The main effect of response type was the only significant effect, $F(1,11) = 25.49$, $p<.001$.

The data for the <u>same</u> and <u>different</u> functions were then separately subjected to orthogonal trend analysis. For the same function, neither the linear nor the quadratic components were significant, $F(1,11) = 0.41$, $p>.05$; $F(1,11) = 0.52$, $p>.05$, respectively. For the <u>different</u> function, the linear component was significant but the quadratic was not, $F(1,11) = 10.71$, $p<.01$; $F(1,11) = 0.09$, $p>.05$, respectively.

Error rates tended to mimic RT as the two were positively correlated, $r(4) = .66$, $p>.05$. Analysis of variance of error data revealed no significant effects.

<u>Discussion</u>

The results suggest that changing the background texture from diagonal line segments to circles resulted in a dramatic change in task performance. In Figure 9 the results of Experiment 1 are superimposed on the results of the present experiment.

Insert Figure 9 about here

In an effort to statistically assess the change in performance due to the change in backgrounds, the results from the two experiments were entered into a single analysis of variance with background (diagonals vs. circles) as a between-subjects variable. The analysis confirmed what is apparent in the figure. The three-way interaction between response type, target number, and background was significant, $F(2,44) = 4.97$, $p<.01$, indicating that background significantly influenced the effect of target number, and this influence varied with response type. Surprisingly, the main effect of background was not significant, ($F(1.22) = 1.51$, $p>.05$.

It is clear that with circles as texture elements, target number had no effect (2.75 ms/target slope) on <u>same</u> trials and little effect (11.5 ms/target slope) on <u>different</u> trials, compared to large, non-monotonic effects with diagonal texture elements. Thus, these results suggest that the target number effects observed in Exps. 4 and 5 are specific to the use of diagonal texture elements. In addition, the results are consistent with the hypothesis that diagonal texture elements may have influenced decision processes in Exps. 4 and 5 by generating background activity in target feature maps.

It could also be argued, however, that the reduction in slope with circles as texture elements is not specific to the removal of noise from feature maps, but is instead due to a decrease in the general similarity between the targets and background. Perhaps as similarity between targets and background decreases, targets have a greater tendency to emerge from the background as a group, allowing responses to be based on the simple detection of differences among the features in the newly emerged group.

This is a reasonable alternative that requires further research. However, it should also be pointed out that the hypothesis of grouping based on "similarity" outlined above is difficult to disconfirm. The "similarity" of targets and background is hard to define a priori. If a set of background elements yields no target number effect, one concludes that the elements were dissimilar enough to allow grouping of targets. If, however, there is a target number effect one can always conclude that the elements were not dissimilar enough for grouping to emerge.

Experiment 9: Subitizing

Before turning to our investigations of verbal materials there is one further study that uses line segments that needs to be reported. It is not precisely on the subject of discrimiantion or identification, although it is on a highly related topic.

One of the classical findings of experimental psychology is that in judging the number of discrete elements presented simultaneously there is a discontinuity in performance at about six items. Kaufman, Lord, Reese, and Volkmann (1949) used the term subitizing to refer to the rapid, confident, and accurate report of six or fewer elements. The numerosity of arrays larger than this is presumably ascertained by a qualitatively different process. With long or unlimited exposure duration and an emphasis on accuracy subjects will tend to count the stimuli; with more than six stimuli present reaction time (RT) increases by several hundred milliseconds for each additional stimulus that is counted (see Klahr & Wallace, 1976 and Mandler & Shebo, 1982 for reviews). With an exposure duration that is too brief to permit counting subjects will simply estimate numerosity; this can be accomplished with little increase in RT.

The six-item limit on subitizing has been taken to be an indication of the number of items that the mind can grasp at once (e.g., Miller, 1956), which, in turn, suggests the possibility that numerosities of up to six items can be ascertained directly (i.e., without serial counting). More detailed analysis, however, shows that even below six items the latency of subitizing judgments increases somewhat with the number of items in the display. It is common to find increases in latency in the range from 25 to 100 ms for each additional stimulus up to about four stimuli, (for reviews see Klahr and Wallace, 1976 and Mandler & Shebo, 1982). Such a latency increase suggests that processing of numerosity even within the subitizing range may involve a serial component. Models incorporating a serial component for determining the numerosity of even small arrays have been proposed by Klahr and Wallace (1976) and by Ullman (1984).

A study that challenges previous subitizing results and suggests parallel processing of numerosity was reported by Sagi and Julesz (1985a,b). These authors examined the nature of processing in three different tasks. In the task relevant to our present purpose subjects had to indicate the number of targets present in a field of nontargets. The relevant details about the Sagi and Julesz experiments are as follows. Targets were mixtures of a few horizontal and vertical line segments embedded in a texture composed of diagonal line segments that were all oriented in the same direction. Subjects had to determine the number of targets irrespective of orientation (thus a stimulus containing one vertical and three horizontal targets should get the response "four"). Target numerosity was blocked; in each block of trials the subject had to decide between two numerosities (1 vs. 2, 2 vs. 3, 3 vs. 4). Stimuli were presented briefly and followed by a pattern mask. Performance was measured in terms of accuracy at different stimulus onset asynchronies. The finding of chief interest was that the processing time needed to achieve a given level of accuracy remained constant as target number increased; a result consistent with parallel processing of target numerosity.

The research by Sagi and Julesz differs methodologically in several ways from most of the subitizing literature. For one thing, it is not common to have subjects assess numerosity in the face of distracting nontargets (although it should be said that it is not clear how this particular difference could lead to parallel processing). For another, the blocking of numerosity is unusual in this kind of research. Finally, they measured accuracy of response to masked displays with well-trained psychophysical observers instead of reaction time with subjects serving for just an hour or two.

It is possible that the particular way in which numerosity was blocked (especially in conjunction with the use of highly practiced subjects) may have had a profound effect on the outcome of the experiment. Specifically, when numerosity is blocked the task can become one of judging whether a stimulus is greater than or less than a specific criterion. This task can be solved by a mechanism as simple as a one-layer perceptron (see Minsky & Papert, 1969, Ch. 1). However, as is typically the case, when a unique response is required for each of several different numerosities (i.e., when the system must count inputs) a one-layer perceptron fails (see Minsky & Papert, Ch 8). In short, the difference between Sagi and Julesz's results and most previous results may reflect a basic difference in the "computational geometry" of the blocked and mixed presentation modes. A series of studies was conducted in an attempt to resolve the empirical discrepancy (see Folk, Egeth, & Kwak, in press); of which just one is reported in this technical report.

In this study stimuli were displayed briefly and followed by a mask. The dependent variable of chief interest was accuracy. Each subject served in six one-hour sessions.

Method

Subjects. Six subjects served in this experiment for pay. They were students at Johns Hopkins University.

Apparatus. Displays were generated by an IBM AT computer and displayed on a Hewlett Packard 1345A digital display module with a P31 phosphor. The display scope was mounted into one channel of an Iconix four channel tachistoscope. A second channel in the tachistoscope provided a constant, dim, background illumination. The computer collected all reaction time and error data.

Stimuli. Displays were similar to those used in the previous experiments. Line segments were placed in the imaginary cells of a 8 x 8 matrix that measured 7.5 cm vertically and horizontally. The segments were 6mm long and had a stroke width of approximately 0.7mm. At the viewing distance of 65 cm the matrix subtended approximately 6.86 deg of visual angle horizontallv and vertically, and the line segments subtended approximately .53 deg.

Targets were defined as horizontal and/or vertical line segments. Target positions were distributed under the same constraints as in the previous experiments. All displays contained negatively-sloped diagonal texture elements. Displays were generated "on line" and thus no two subjects saw exactly the same set of displays.

Design. Each subject served for six one-hour sessions. Within each session a subject served in three blocks of 96 trials each. Each block was devoted to a different quantitative discrimination, i.e., 1 vs. 2, 2 vs. 3, and 3 vs. 4. The proportions of specific trial types (e.g., number, orientation and heterogeneity of targets) within each block were identical to those used in Experiment 2. The order of presentation of the three blocks was balanced across days for each subject and across subjects on each day by means of a Latin Square design.

Procedure. On each trial subjects first saw a fixation point for 1000 ms. It then disappeared and the field remained blank for 500 ms, at which time the stimulus appeared for an amount of time that was determined separately for each subject on each day. The stimulus was followed immediately by a post mask similar to that used by Sagi and Julesz (1985a,b). It consisted of randomly oriented Vs, each V consisting of two arms meeting at a 45 deg angle. One V appeared in each cell of the matrix. The post mask was shown for 500 ms or until a response was made, whichever was less. The response was manual; subjects depressed one of two keys on a box that was on a table in front of them. To indicate the smaller of the two numerosities in a particular block the subject pressed the key on the left; to indicate the larger numerosity the subject pressed the key on the right. Accuracy was measured as the average percentage correct of the two alternatives within each number set. This was done to eliminate contamination of the data by a subjective preference toward one alternative or the other.

Exposure duration was set in a separate block of trials at the beginning of each session. The stimuli used to set duration always required the 3 vs. 4 discrimination. Subjects were always started at 200 ms. A simple up-down staircase algorithm was used to set stimulus exposure duration such that subjects were responding correctly on about 70% of the trials (Levitt, 1970). Responses on successive pairs of trials were tallied and exposure duration adjusted according to the following rules. When both trials were correct duration was decreased by 10%. When the first member of the pair was correct and the second incorrect, or when both were incorrect, duration was increased 10%. Finally, if the first was incorrect and the second correct, no change was made. This procedure was carried out until five reversals in the direction of change occurred. It took less than five minutes. Once an exposure duration was established it was used without further change in all conditions for that subject's entire session.

Results.

Exposure duration. As expected, subjects improved with practice. Over sessions mean exposure duration (in ms) decreased as follows: 163, 131, 127, 103, 92, 110. Exposure duration appears to reach an asymptotic value of roughly 100 ms. after about four days.

Performance. The purpose of the algorithm for selecting an exposure duration was to achieve a fairly stable overall level of performance (approximately 70% correct) from session to session in the face of changes in task proficiency with practice (and the concomitant reduction in exposure duration). The method we used seemed effective despite its brevity. Mean overall percentages correct across days were: 77, 73, 79, 75, 73, and 79.

Figure 10 shows the main results of the experiment. Mean accuracy rates for each number set, day and subject were entered into a two way, within-subjects analysis of variance with size of number set (1 vs.2, 2 vs. 3, and 3 vs. 4) and days (1 through 6) as factors. Performance decreased significantly as target numerosity increased, $F(2, 10) = 46.06$, $p < .01$. Moreover, this effect is the same across the six days of the experiment, as target numerosity did not interact with days, $F(5, 25) = 0.53$.

Insert Figure 10 about here

Discussion. Consistent with the previous two experiments, and in contrast to the results of Sagi and Julesz (1985a,b), performance in the present experiment varied directly with target numerosity. This effect was evident with accuracy as the dependent measure and at all levels of practice. The design of this experiment, however, was not identical to the experiment of Sagi and Julesz (1985a,b). In particular, in their experiment stimuli were displayed for 5 ms. and then followed by a blank field for a variable SOA before the 10 ms. presentation of the mask. The SOA was varied systematically in 50 trial blocks, thus permitting the experimenters to sweep out a complete psychometric function. The independence of performance and number of targets was illustrated in a figure (Figure 2 in 1985a, Figure 2 in 1985b) in which the SOA required for 95% accuracy was plotted against number of targets. The functions were essentially flat for the three subjects in the study. One possible problem, then, in comparing our study with Sagi and Julesz's is that the levels of accuracy were too discrepant to make a comparison meaningful. Note, however, that their summary figure merely focuses on and replots the 95% accuracy data from the psychometric functions for the individual subjects (Figure 2 in 1985b). What we can see from the full psychometric functions is that essentially the same independence of performance and number of targets holds throughout the entire range of SOAs. In particular, at the same level of performance as in the present experiment, i.e., at about 75% correct, it also seems to be the case that the SOA required to achieve that level of performance does not vary systematically with number of targets.

As an additional check on the possible effect of overall accuracy on differences among numerosity levels we took advantage of the variability of our data around the mean in the following bit of data snooping. Recall that the staircase adjustment procedure was used to set an exposure duration for the entire session. The procedure worked well as a means for getting subjects to operate at about the same level of overall accuracy, however, it by no means eliminated all variability among subjects, or among sessions for a given subject. We took the data for each subject and ordered his or her performance for the six sessions in terms of overall daily accuracy. These six values were then collapsed into three categories, low accuracy (the two worst sessions), medium accuracy (the two sessions intermediate in accuracy) and high accuracy (the two best sessions). Figure 11 shows accuracy as a function of size of number set separately for the three levels of overall accuracy. These data were entered into an analysis of variance in which both overall accuracy level (low, medium, and high) and size of number set (1 vs. 2, 2 vs. 3, and 3 vs. 4) were within-subject variables. This analysis confirmed that the main effect of size of number set was significant, $F(2, 10) = 42.48$, $p < .001$. It also showed significant differences among the three levels of overall accuracy, but this is entirely unsurprising given the way the data were organized for the purpose of this analysis, $F(2, 10) = 27.67$, $p < .001$. The important point of this analysis is that there was no interaction between size of number set and level of overall accuracy,

$F(4, 20) = 33$. Thus our conclusion that performance gets worse as size of number set increases seems to be unaffected by differences in overall level of performance.

Insert Figure 11 about here

Discussion

The finding of Sagi and Julesz (1985a,b) that counting or numerosity judgement did not take more time for larger numbers was provocative and seemed well worth investigating. Unfortunately, the results of the present series of experiments have to be considered largely negative. Our results consistently supported the generalization that increasing numerosity leads to increases in the time required to process the stimulus. In Experiment 1 with target numerosity mixed we found that mean RT increased for both texture and no texture displays. In Experiment 2 the same general pattern emerged even though target numerosity was blocked. Finally, in Experiment 3 accuracy was the dependent variable and performance was measured over six days of practice. Accuracy decreased as target numerosity increased regardless of the level of practice. The consistency of our results across substantial differences in methodology is quite striking. At this time we cannot explain the discrepancy between our results and those of Sagi and Julesz (1985a,b). There remain unexplored differences in procedure and subject population between our research and theirs that may be responsible for the discrepancy in outcomes.

At the outset we thought that Sagi and Julesz's result might be due to their decision to block numerosity and require discrimination between N and N + 1 targets within each block. As we pointed out in the Introduction, even a mechanism as simple as a one-layer perceptron could solve such a problem. In more traditional studies of numerosity judgment the number of different levels of numerosity that must be judged is much larger. A simple perceptron cannot solve such a problem. However, our own results do not suggest that the difference between the two-choice and multi-choice judgement paradigms leads to any crucial qualitative difference in performance.

Sagi and Julesz concluded that knowing how many targets are present in a display is accomplished by means of parallel processing. (This conclusion is based in part on an additional experiment reported in the 1985b paper, which we shall not discuss at this time.) It is a difficult matter to determine whether processing is serial or parallel in any given task (e.g., Townsend, 1971). We certainly would not wish to conclude that the present results necessarily indicate that processing was serial in our tasks simply because performance tended to get worse as numerosity increased. However, based on our research there seems to be little reason to reject models that assume that numerosity judgement in the subitizing range contains a serial component (e.g. Klahr & Wallace, 1976; Ullman, 1984).

Experiment 10: Digit vs. Letter Categorization

We turn now to a series of three studies all of which presumably require a response based on a more abstract property than the orientation of a line segment. The question is whether any of them permit spatially parallel processing. In the first of these three studies subjects indicated whether characters were digits or letters.

Method

Subjects. Thrty-two students participated in this experiment.

Equipment. Stimuli were displayed on a green monochrome monitor controlled by an IBM XT computer with a Hercules Graphics with a Hercules Graphics adapter. Responses were made by pressing the space bar on the computer keyboard. Millisecond timing was accomplished with the computer's internal clock.

Stimuli and Procedure. On each trial one, two, or three letters or one, two, or three digits were displayed for 150 msec. A go-no go task was used; half of the subjects were to press a response key when the display contained letters and to refrain from responding when the display contained digits; the other half of the subjects reponded when the display contained digits. The set of digits used was 2,4,5,6,8; the set of letters used was B,G,K,S,Z. It is important to note that all of the characters on a given display were identical. Thus a display might consist, for example, of G or 44 or BBB, but not 245 or BGK or 2K5.

The centers of the characters were located on an imaginary circle 2.14 deg. in diameter. The locations used were at 12:00, 4:00, and 8:00. Individual characters were light on a dark background; they were 0.3 deg in height and 0.2 deg. in width.

Each subject received 20 randomly selected practice trials followed by 360 experimental trials that were divided equally between letter and digit presentations. For both letter and digit displays one-, two-, and three-character presentations were equally frequent. The various characters and possible display locations were also used equally often. Because of the difficulty of the required discriminations, subjects received error feedback.

Results

The data were analyzed separately for subjects who responded to letters and those who reposnded to digits. Mean RTs for the one-, two-, and three-letter displays were 553, 535, and 535 msec., respectively, and 517, 506, and 495 msec. for the corresponding digit displays. There was a significant effect of number of redundant targets, both Fs(2,30) > 13.00, ps <.001.

Errors tended to decrease as target numerosity increased; for both letters and digits, $p < .05$.

Inspection of the data for single-target conditions showed no evidence of a fixed favored position, and thus the analysis in terms of a fixed favored position artifact was not carried further.

The redundancy gain was tested against the random favored position hypothesis, as in the preceding experiment. The mean RTs for the subset analysis appear in Table 2. There was, of course, a significant effect of number of targets. More instructive is the lack of any interaction between subset and number of targets. For both letters and digits, Fs < 1.5. Thus, the random favored position hypothesis cannot account for the observed redundancy gain.

Our conclusion is that the most natural accounting of these results is in terms of unlimited capacity parallel processing.

Explicit in our introduction of the letter-digit classification task is the assumption that it is categorical and not solvable by detecting one or another specific feature. This assumption has been made before, but it is controversial. For example Krueger (1984) has claimed that the control for featural differences has been incomplete in most studies of digit-letter classification. When he matched letters and digits with respect to featural differences, the "category effect" disappeared. More recently, however, Dixon and Shedden (in press) have countered that similarity of items between categories may have biased observers against using category information in Krueger's task. When this problem was eliminated a category effect reemerged. This controversy is, we are sure, not over yet. The problem with digits and letters is that there are relatively few of them. An ingenious advocate can always find some small set of features that could conceivably be used to distinguish stimuli drawn from the two categories. However, if a similar effect could be demonstrated with words the possibility of such a featural analysis would be very unlikely.

Experiment 11: Semantic Categorization

Having found that parallel processing is possible in the categorization of characters as digits or letters, it seemed a plausible extension to see if it would also be possible in a task that requires semantic categorization of words.

Method

Subjects. Twelve students participated in this experiment.

Equipment. The equipment was the same as in the preceding experiment.

Stimuli and Procedure. A total of 40 different four-letter words was used. Individual words were 0.6 deg. in height and 1.2 deg. in width, and were located 0.6 deg. above or below fixation.

On each trial either one or two words were presented for 150 ms.; when one was presented it was equally likely to be above or below fixation. When two words were presented they were in fact two physically identical copies of the same word (e.g., LION and LION, one above the other).

Each subject received a total of 640 experimental trials, divided into four equal blocks. In each 160-trial block subjects made a different categorical decision. For example, in one block a subject might respond to animal names and refrain from responding to names of tools, while in another block that same subject might respond to names of articles of clothing and refrain from responding to vehicle names. There were four target categories to which subjects were to respond by pressing the space bar (animals, body parts, birds, clothing), and four nontarget categories for which subjects were to refrain from responding (tools, vehicles, furniture, fruits and vegetables). Each category contained five words. The specific combinations of the four positive and four negative categories was balanced in a Latin Square design.

Each 160-trial block consisted equal numbers of one-word and two-word displays, and equal numbers of positive and negative trials. Within each block the order of presentation of trials was random.

Results and Discussion

Mean positive RTs for one-target and two-target trials were 545 and 547 ms, respectively, $F(1,11)=0.45$. The error rates for the two conditions were 4.2% and 3.5%, respectively, which did not differ significantly. There is, obviously, no effect of redundancy in this experiment, and so no further analyses were required.

It should be noted first that there is no evidence that subjects were using an exhaustive processing rule in this task. If they had, then mean RT would have increased with display size. In fact, mean RT was uninfluenced by display size. This suggests the existence of a capacity limit for word categorization. Whether processing is serial or parallel cannot be determined.

The difference between Exps. 10 and 11 is intriguing. One possible reason for the discrepancy is simply that there were too many letters on the screen on Exp. 11. There were four or eight characters on the screen in that experiment as compared to a maximum of three in the previous experiment. In addition, the characters in Exp. 11 were packed together to form words; lateral interactions among adjacent letters may have impaired performance. Then again, there may be something about word categorization that imposes heavier demands on capacity than digit-letter categorization. It may even be the case that word analysis is inherently serial.

Experiment 12: Lexical Decision Task

To begin to explore the reasons for the difference between the results of Exps. 10 and 11 the stimuli in the present study included the same words as were used in Exp. 11, but they were used in a different task. Word trials were mixed with nonword trials; subjects had to respond to words and refrain from responding to nonwords. The equipment was also the same as in the preceding experiment.

Method

Subjects. Sixteen students participated in this experiment.

Stimuli. The stimuli were the same as in the preceding experiment except that the 20 words from the 4 nontarget categories were converted to pronounceable nonwords by changing one letter of each word; e.g., raft became reft. The four letter positions served about equally often as the locus of the change.

There were 640 trials divided into 8 equal blocks. However, here the purpose of the blocks was simply to provide subjects with opportunities to rest; the blocks did not correspond to a change in any substantive factor. Thus, the appearance of any particular kind of word was not restricted to a particular block (e.g., the animal names did not all appear in one block). Words and nonwords were selected randomly with the constraint that each item appeared equally often. As before, when two stimuli were presented on a trial they were identical.

Results

The overall positive mean RTs for the one and two target conditions were 513 and 502 msec., respectively, $F(1,15) = 11.65$, $p < 0.01$. The corresponding error rates were 5.4% and 2.9%, respectively, $F(1,15) = 14.9$, $p < .05$. (Note that this is the opposite of a speed-accuracy tradeoff.) The redundancy gain was reliable, but further analysis was necessary to determine whether the effect might be due to a favored position artifact.

Inspection of the data from the single-target trials showed no evidence of a fixed favored position, and so no further analysis along these lines was carried out.

The test for the random favored position hypothesis was analogous to the one in Exp. 6. The relevant data for fast and slow random subsets are shown in Table 2. The difference between one- and two-target trials is not significantly larger for the slow than the fast subset; for the interaction $F(1,15) < 1.0$. Thus a random favored position artifact does not seem to account for the data.

Discussion

The processing mode suggested by the results of this experiment is unlimited capacity parallel processing. This result indicates that the difference between the digit-letter classification and semantic categorization studies is probably not due to such factors as the absolute number of letters in the display or lateral interactions among letters that form words. In the absence of further research it is not possible to say with any confidence just what it is about the semantic categorization task that leads to results that are so different from that of digit-letter classification and lexical access. This is obviously an interesting area for further research.

Conclusions

One of the major issues addressed in this chapter concerns the ability to process in parallel individual stimuli that differ with respect to a single dimension. This was the topic of Exps. 1-9. Our conclusion is that parallel processing is indeed possible even in tasks that appear to require "identification" of the levels of each target on the critical dimension and do not permit a response simply upon detection of a feature gradient. This conclusion is broadly consistent with the thinking of a variety of researchers who have written about preattentive processing (e.g., Beck, 1982; Julesz, 1986, Treisman, 1985). We see the contributions of our research as twofold. First, we feel we have in large part clarified the challenge to current theorizing posed by the work of Sagi and Julesz (1985a,b). In particular, we conclude that their finding that performance deteriorated linearly with number of targets does not require the assumption that processing is serial. Second, by the use of a broad range of tasks and experimental conditions (e.g., textured background and no background conditions; blocked and mixed presentation modes; search and redundant-target paradigms; clear and degraded stimuli) we have substantially broadened the empirical basis for theories of preattentive processing. Of course, even with simple, unidimensionally varying stimuli, processing is not necessarily parallel. This is suggested by the results of our study of subitizing, which are compatible with a serial processing model.

The other major issue addressed in this chapter concerns the ability to process in parallel stimuli that differ with respect to their meaning. This was the topic of Exps. 10 and 11 and, depending on how one thinks the lexical decision task is accomplished, possibly Exp. 12 as well.

The results of the letter-digit classification task (Exp. 10) confirm the earlier finding of Egeth, Jonides, and Wall (1972) that processing in that task can be characterized as parallel with unlimited capacity. However, when the classification task was based on word meaning (Exp. 11), limited capacity processing was found; whether it was parallel or serial is not possible to determine at this time.

Perhaps the single most striking finding of the present set of nine experiments is that the lexical decision task (Exp. 12) permitted unlimited capacity parallel processing. On the one hand, this result should perhaps not be considered completely unexpected, as McClelland and Mozer (1986) have previously argued that two word strings can simultaneously achieve access to structural knowledge about word form. Note, however, that (a) their experiments did not provide a direct test of this hypothesis, and (b) the hypothesis does not require that processing capacity be unlimited. On the other hand, if one's point of departure is something like feature integration theory (according to which something as seemingly simple as finding a red circle in a field of red squares and blue circles requires serial processing) then the present result is very surprising indeed.

The redundancy-gain paradigm was introduced in the hope that it would provide a clear test of processing mode. However, it does suffer from one of the problems discussed in the Background section. Because a response can be made as soon a single target is identified it does not necessarily require full processing of all target items (Pashler & Badgio, 1985). It is possible, for example, that the redundancy gain is due to a "race" between low-level perceptual processes. A process of stimulus categorization, whether semantic or otherwise, may then be applied just to the stimulus whose low-level processing finishes first. This argument may be countered by noting that a race among low-level processes should be just as beneficial in a sema ic categorization task as a lexical-decision task. That there was no redundancy gain in the semantic task thus stands as at least some evidence against the application of Pashler and Badgio's (1985) argument to our lexical decision task.

Several interesting implications arise from the results of Exps. 10, 11, and 12. First, these results seem to imply, contrary to early selection theories, that complex structural information can be analyzed in a spatially parallel manner on independent channels. It is difficult to see how the lexical decision task could be performed without a detailed analysis of form. Second, these results seem to imply, contrary to late selection theories, that word meaning cannot be derived in a spatially parallel manner on independent channels.

Clearly there is no easy resolution to these issues, and equally clearly more research is required to determine whether these preliminary explorations are accurate portrayals of the processing architecture. However, it would appear that a compromise of some sort may be necessary between theories that assume that all interpretive analyses of visual stimuli can be conducted in parallel without capacity limitations and theories that assume only rudimentary analyses can be performed in this manner. At this point, it is somewhat unclear how detailed the structural analysis of visual stimuli can be before limitations are encountered. Our research suggests that simple features can be processed in parallel and independently, and in some instances stimuli such as alphanumeric characters and even words can be processed in that way. An interesting enterprise would be to attempt to understand what it is about the processing system involved in semantic analyses of words that results in the observation of capacity limits.

References

Allport, D.A., Tipper, S.P., Chmiel, N.R.J. (1985). Perceptual integration and postcategorical filtering. In M.I. Posner and O.S.M. Marin (Eds.), Attention and Performance XI. Hillsdale, NJ: Erlbaum.

Beck, J. (1982). Textural segmentation. In J. Beck (ed.), Organization and Representation in Perception. Hillsdale, NJ: Erlbaum.

Beck, J., & Ambler, B. (1972). Discriminability of differences in line slope and in line arrangement as a function of mask delay. Perception & Psychophysics, 12, 33-38.

Biederman, I., & Checkosky, S. F. (1970). Processing redundant information. Journal of Experimental Psychology, 83, 486-490.

Bjork, E. L., & Murray, J. T. (1977). On the nature of input channels in visual processing. Psychological Review, 84, 472-484.

Broadbent, D. E. J. (1971). Decision and Stress. London: Academic Press.

Connor, J. M. (1972). Effects of increased processing load on parallel processing of visual displays. Perception & Psychophysics, 12, 121-128.

Deutsch, J. A., & Deutsch, D. (1963). Attention: some theoretical considerations, Psychological Review, 70, 80-90.

Dixon, P., & Shedden, J. M. (in press). Conceptual and physical differences in the category effect. Perception & Psychophysics.

Donderi, D., & Zelnicker, D. (1969). Parallel processing in visual same-different decisions. Perception & Psychophysics, 5, 197.

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

Egeth, H., Jonides, J., & Wall, S. (1972). Parallel processing of multielement displays. Cognitive Psychology, 3, 647-698.

Egeth, H., & Santee, J. L. (1981). Conceptual and perceptual components of interletter inhibition. Journal of Experimental Psychology: Human Perception and Performance, 7, 506-517.

Egeth, H., Virzi, R. A., & Garbart, H. (1984). Searching for conjuntively defined targets. Journal of Experimental Psychology: Human Perception and Performance, 10, 32-39.

Farell, B. & Julesz, B. (1986). Perception of static directional flow fields. Technical Memorandum, AT&T Bell Laboratories.

Folk, C. L. (1987). Preattentive representation and processing of visual feature information. Unpublished doctoral dissertation, Johns Hopkins University, Baltimore.

Garner, W. R. (1978). Aspects of a stimulus: features, dimensions and configurations. In E. Rosch and B. B. Lloyd (Eds.), Cognition and Categorization (pp. 99-133). Hillsdale, NJ: Lawrence Erlbaum Associates.

Hinton, G. E., McClelland, J. L., & Rumelhart, D. E. (1986). Distributed Representations. In D. E. Rumelhart & J. L. McClelland (Eds.), Parallel Distributed Processing (pp.77-109). Cambridge, MA: The MIT Press.

Houck, M. R., & Hoffman, J. E. (1986). Conjunction of color and form without attention: Evidence from an orientation-contingent color aftereffect. Journal of Experimental Psychology: Human Perception and Performance, 12, 186-199.

Julesz, B. (1986). Texton gradients: the texton theory revisited. Biological Cybernetics, 54, 464-469.

Kahneman, D., & Chajczyk, D. (1983). Tests of the automaticity of reading: Dilution of stroop effects by color-irrelevant stimuli. Journal of Experimental Psychology:Human Perception and Performance, 9, 497-508.

Kaufman, E.L., Lord, M.W., Reese, T.W., & Volkmann, J. (1949). The discrimination of visual number. American Journal of Psychology, 62, 498-525.

Keele, S. W. (1973). Attention and Human Performance. Pacific Palisades, CA: Goodyear.

Klahr, D., & Wallace, J.G. (1976). Cognitive development: An information-Processing View. Hillsdale, N.J.: Erlbaum.

Krueger, L. E. (1984). The category effect in visual search depends on physical rather than conceptual differences. Perception & Psychophysics, 35, 558-564.

La Heij, W., & Van der Heijden, A. H. C. (1983). Feature-specific interference in letter identification. Acta Psychologica, 53, 37-60.

Levitt, H. (1970). Transformed up-down methods in psychoacoustics. The Journal of the Acoustical Society of America, 49, 467-476.

Mandler, G., & Shebo, B.J. (1982). Subitizing: an analysis of its component processes. Journal of Experimental Psychology: General, 111, 1-22.

Marr, D. (1982). Vision. San Fransisco: W.H. Freeman and company.

McClelland, J.L. & Mozer, M.C. (1986). Perceptual interactions in two-word displays: familiarity and similarity effects. Journal of Experimental Psychology: Human Perception and Performance, 12, 18-35.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. Psychological Review, 63, 81-97.

Minsky, M. & Papert, S. (1969). Perceptrons: An introduction to computational geometry. Cambridge, MASS: MIT Press.

Mozer, M. C. (1983). Letter migration in word perception. Journal of Experimental Psychology: Human Perception and Performance, 9, 531-546.

Nakayama, K. & Silverman, G. H. (1986). Serial and parallel processing of visual feature conjunctions. Nature, 309, 264-265.

Neisser, U. (1963). Decision time without reaction-time: Experiments in visual scanning. American Journal of Psychology, 76, 376-385.

Neisser, U. (1967). Cognitive Psychology. New York: Appleton-Century-Crofts.

Norman, D. A. (1968). Towards a theory of memory and attention. Psychological Review, 75, 522-536.

Olson, R., & Attneave, F. (1970). What variables produce similarity grouping? American Journal of Psychology, 83, 1-21.

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

Pashler, H., & Badgio, P. (1985). Visual attention and stimulus identification. Journal of Experimental Psychology: Human Perception and Performance, 11, 105-121.

Pomerantz, J.R., & Sager, L.C. (1976). Line-slope versus line-arrangement discrimination: a comment on Ambler and Finklea's paper. Perception & Psychophysics, 20, 220.

Poulton, E. C. (1982). Influential Companions: Effects of one strategy on another in the within-subjects designs of cognitive psychology. Psychological Bulletin, 91, 673-690.

Quinlan, P. T., & Humphreys, G. W. (1987). Visual search for targets defined by combination of color, shape and size: an examination of the task constraints on feature and conjunction searches. Perception & Psychophysics, 41, 455-472.

Sagi, D., & Julesz, B. (1985a). "Where" and "what" in vision. Science, 228, 1217-1219.

Sagi, D., & Julesz, B. (1985b). Detection versus discrimination of visual orientation. Perception, 14, 619-628.

Sagi, D., & Julesz, B. (1987). Short-range limitation on detection of feature differences. Spatial Vision, 2, 39-49.

Santee, J. L. & Egeth, H. E. (1982). Independence versus interference in the perceptual processing of letters. Perception & Psychophysics, 31, 101-116.

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

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

Townsend, J.T. (1971). A note on the identifiability of parallel and serial processes. Perception & Psychophysics, 10, 161-163.

Townsend, J. T. (1972). Some results on the identifiability of parallel and serial processes. British Journal of Psychology, 25, 168-199.

Treisman, A. M. (1964). Verbal cues, language, and meaning in selective attention. American Journal of Psychology, 77, 533-546.

Treisman, A. (1985). Preattentive processing in vision. Computer Vision, Graphics and Image Processing, 31, 156-177.

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

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

Treisman, A., Sykes, M., & Gelade, G. (1977). Selective attention and stimulus integration. Attention and Performance VI. Hillsdale, New Jersey: Lawrence Erlbaum, 333-361.

Ullman, S. (1984). Visual routines. Cognition, 18, 97-159.

Van der Heijden, A. H. C. (1975). Some evidence for a limited capacity parallel self-terminating process in simple visual search tasks. Acta Psychologica, 39, 21-41.

Van der Heijden, A. H. C. (1981). Short-term Visual Information Forgetting. London: Routledge & Kegan Paul.

Van der Heijden, A. H. C., La Heij, W., & Boer, J.P.A. (1983). Parallel processing of redundant targets in simple visual search tasks. Psychological Research, 45, 235-254.

Table 1

Summary Statistics for Functions Relating Reaction Time to Target Numerosity in Each Condition of Experiments 1 - 4.

Exp.	Background	Response	Slope	Intercept	Significance of Linear Component	Percentage Linear
1	No Texture	Same	4.8	513	n.s.	---
		Different	4.7	539	n.s.	---
	Texture	Same	-9.9	676	n.s.	---
		Different	13.3	631	n.s.	---
2	No Texture	Same	9.3	419	<.05	67.9
	Texture	Same	-16.1	630	n.s.	---
3	No Texture	Same	1.0	521	n.s.	---
		Different	6.5	531	n.s.	---
4	Texture	Same	24.8	554	n.s.	---
		Different	37.1	583	<.05	98.8

Table 2

Mean Reaction Times of Subsets for the Random Favored Position Analysis in Experiments 6, 10, and 12

Exp.	Number of Targets	Subsets: Fast	Intermediate	Slow
6	1	384	441	532
	3	354	399	474
10	Digits			
	1	434	501	618
	3	419	480	588
	Letters			
	1	469	543	649
	3	461	525	621
12	1	447	---	580
	2	439	---	556

Note. Reaction times were ordered for each subject from fastest to slowest and then partitioned into as many subsets as there were target locations.

Figure Captions

Figure 1. Sample stimuli with (B and D) and without (A and C) background texture.

Figure 2. Mean RT and error rate as a function of target numerosity for same and different trials with and without texture in Experiment 1.

Figure 3. Mean RT and error rate as a function of target numerosity for texture and no texture trials in Experiment 2. Only same responses were required.

Figure 4. Mean RT and error rate as a function of target numerosity for same and different trials in Experiment 3. There were no texture elements and target numerosity was blocked.

Figure 5. Mean RT and error rate as a function of target numerosity for same and different trials in Experiment 4. Stimuli contained textured backgrounds; target numerosity was blocked.

Figure 6. Mean RT and error rate as a function of target numerosity for same and different trials in Experiment 5. Data points from the low visual quality condition (superimposed noise mask) are connected by dashed lines, that from the high visual quality condition (no mask) are connected by solid lines.

Figure 7. Mean reaction time and error rates for same and different trials as a function of target number in Experiment 7.

Figure 8. Mean reaction time and error rates for same and different trials as a function of target number in Experiment 8.

Figure 9. Results of Experiment 8 (solid lines) superimposed on those of Experiment 7 (dashed lines).

Figure 10. Accuracy (percentage correct) as a function of number set for each of six daily sessions in Experiment 9.

Figure 11. Accuracy (percentage correct) as a function of number set for each of three levels of overall accuracy in Experiment 9. See text for details.

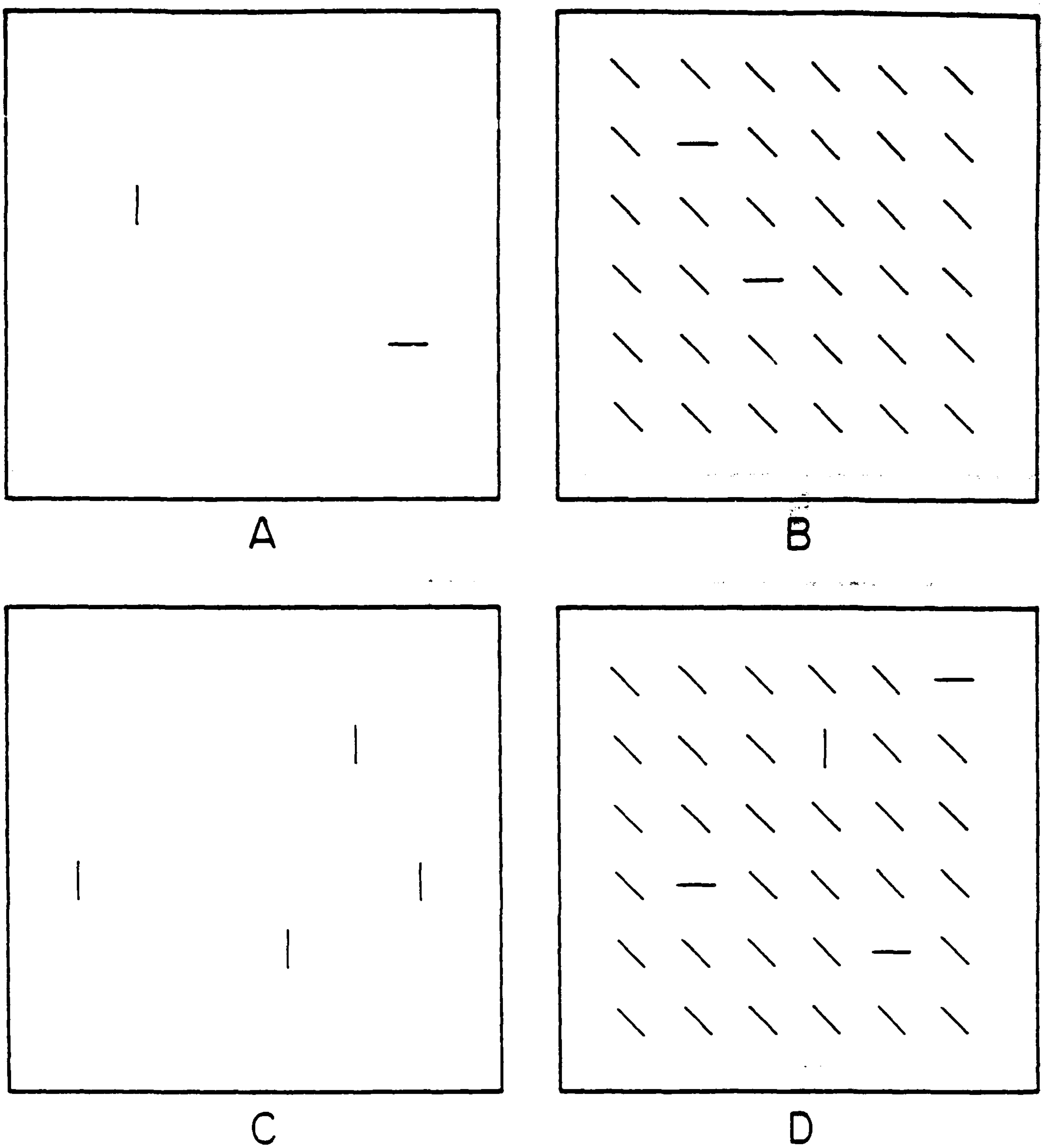


Figure 1

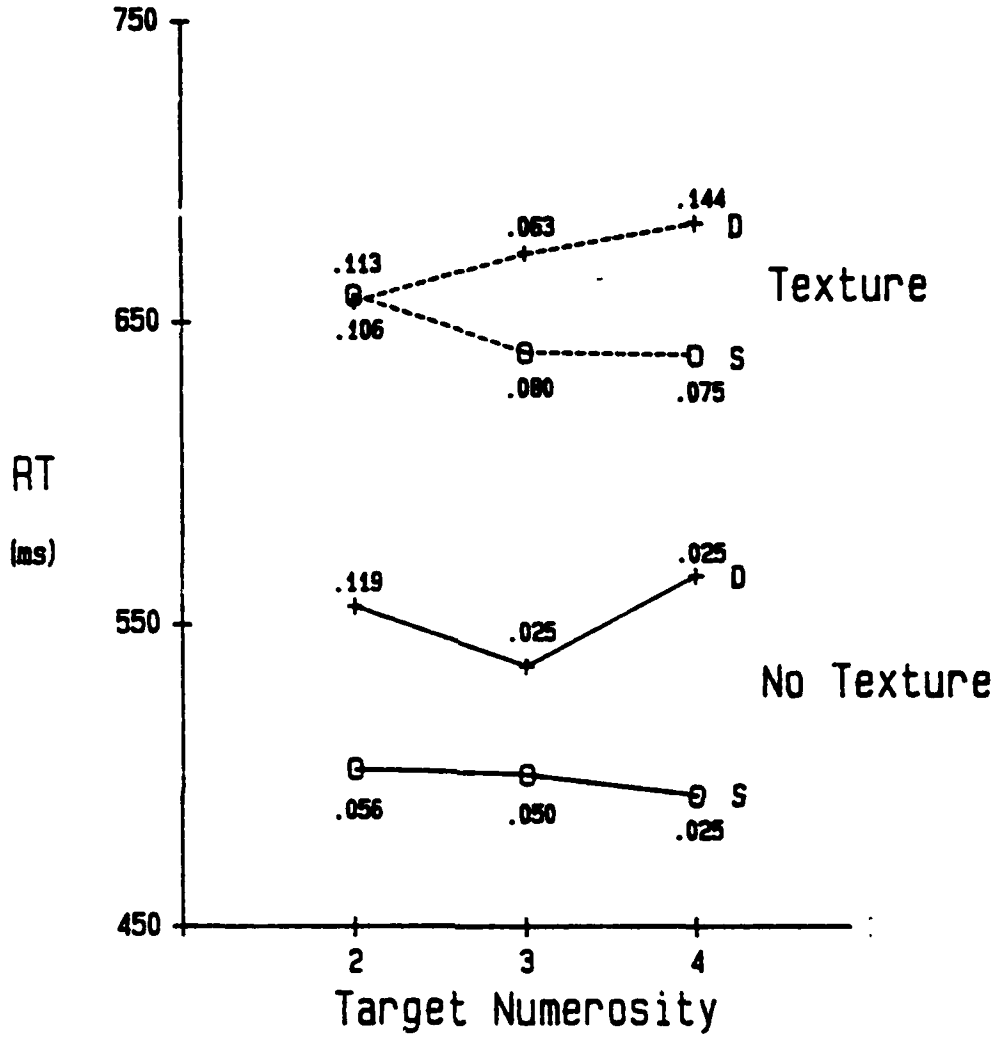

750
650
550
450
RT
(ms)
.113
.106
.063
.144
D
Texture
.080
.075
S
.119
.025
.025
D
No Texture
.056
.050
.025
S
2
3
4
Target Numerosity

Figure 2

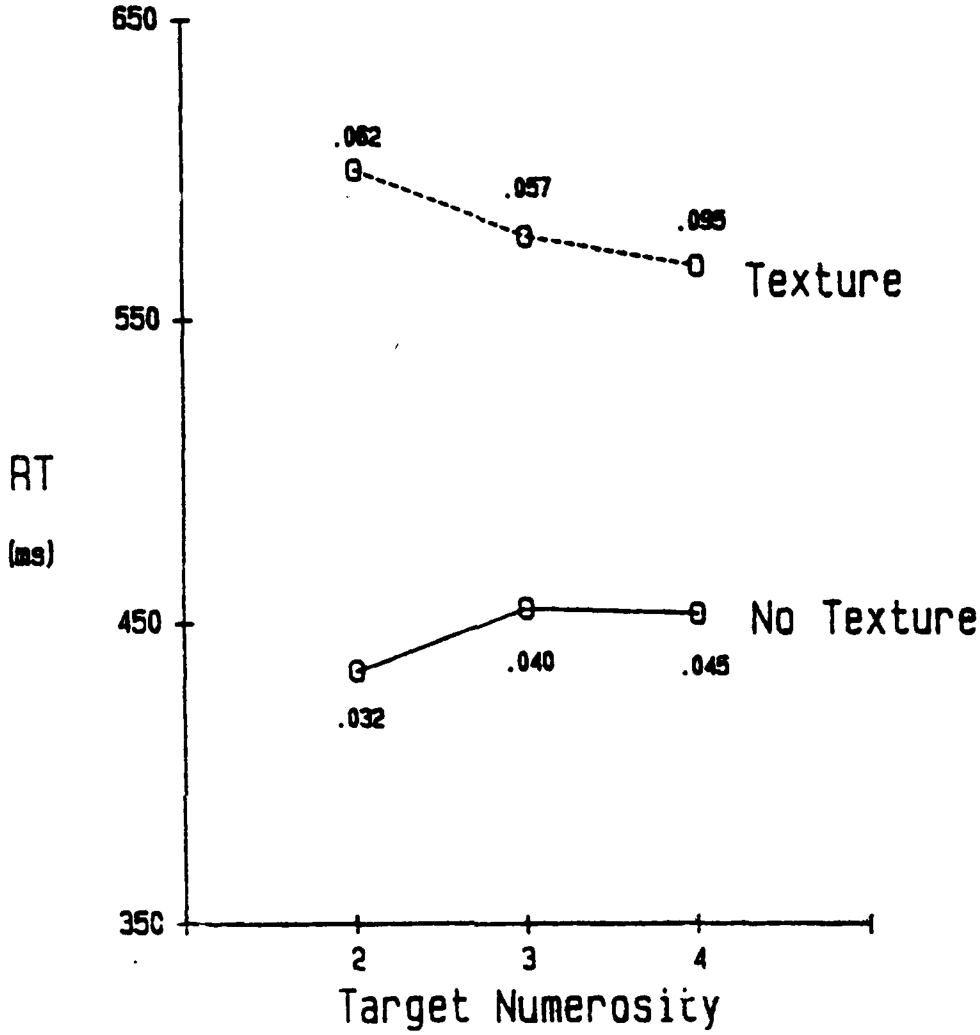

650
550
450
350
.062
.057
.095
Texture
RT
(ms)
No Texture
.040
.045
.032
2
3
4
Target Numerosity

Figure 3

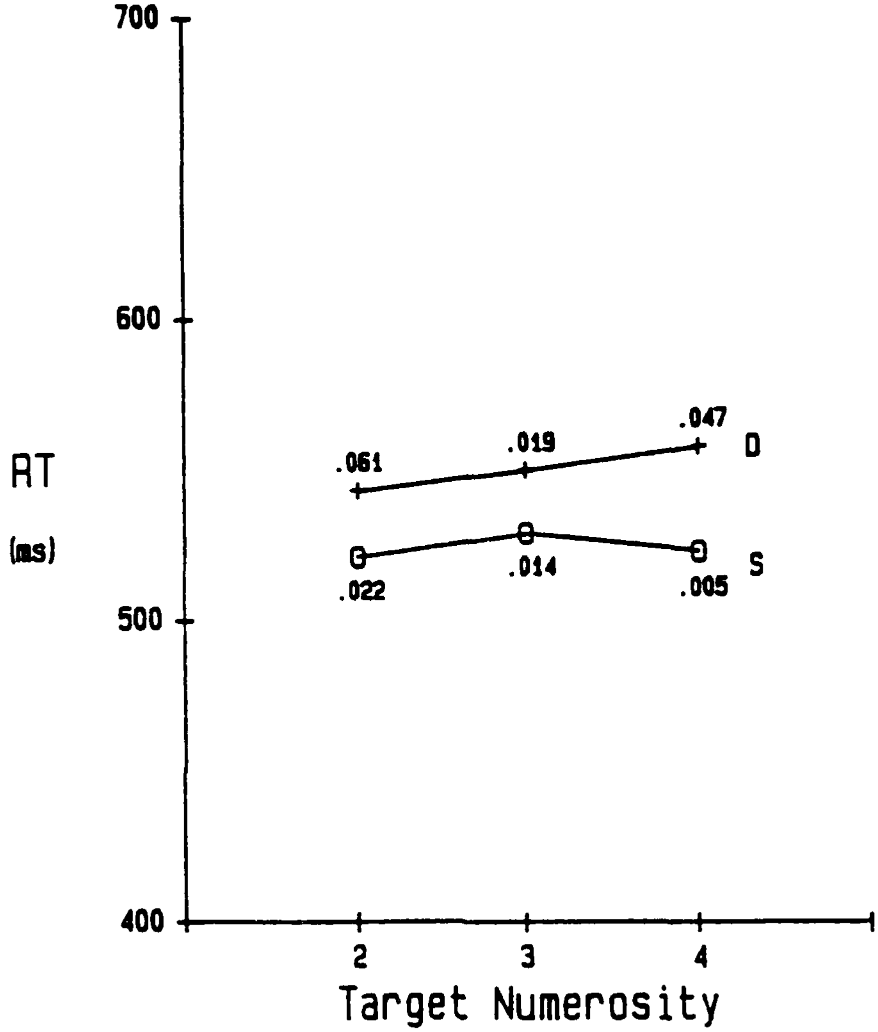
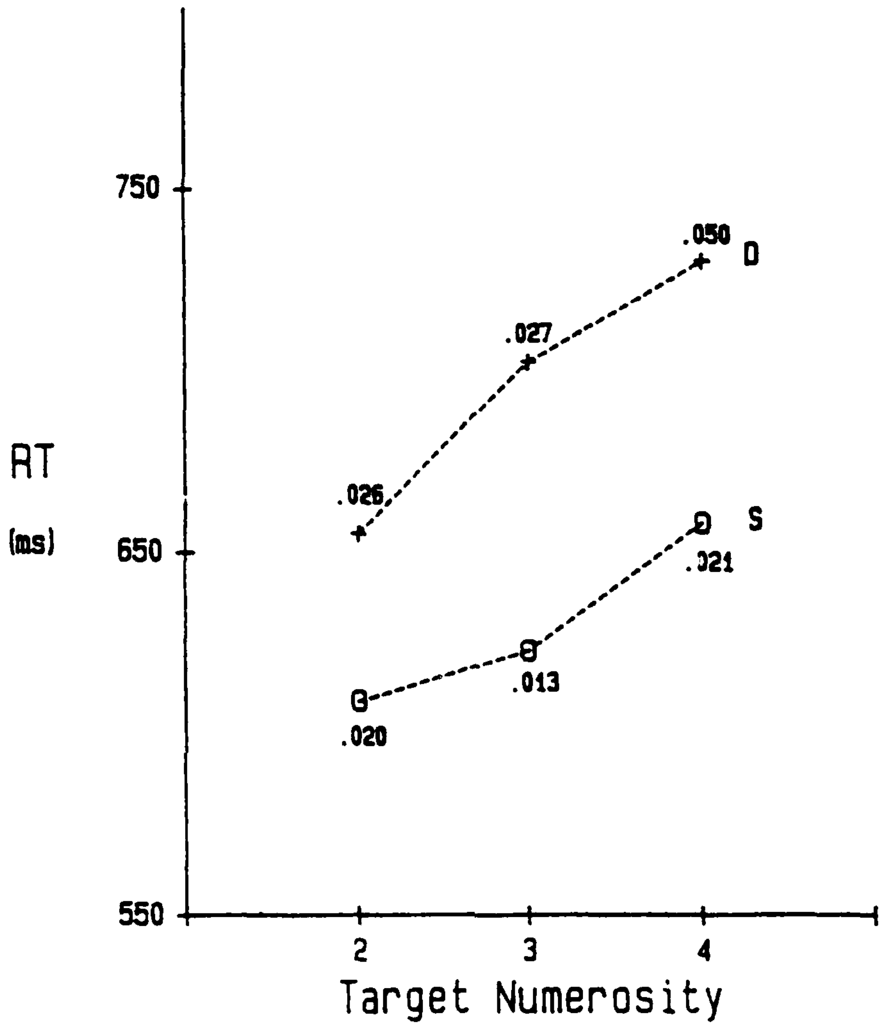

700
600
500
400
RT
(ms)
.061
.019
.047
D
.022
.014
.005
S
2
3
4
Target Numerosity

Figure 4



750
650
550
RT
(ms)
.050
D
.027
.026
S
.021
.013
.020
2
3
4
Target Numerosity

Figure 5

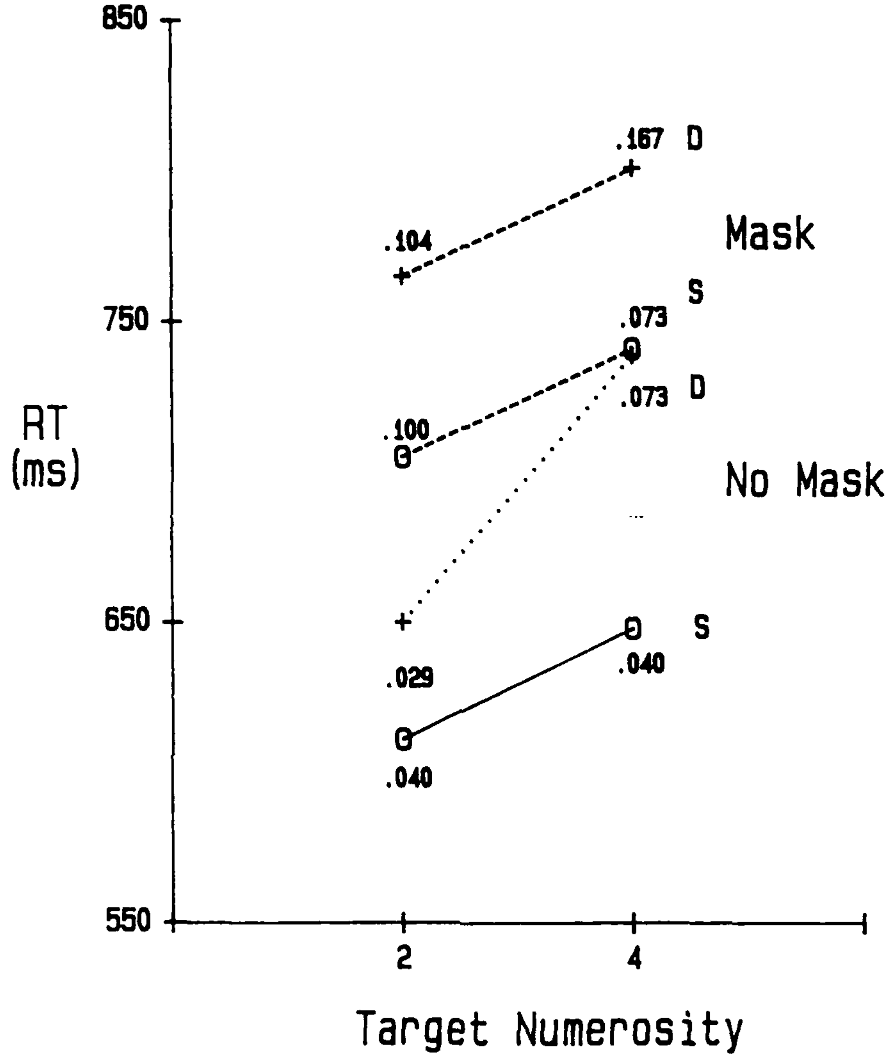

850
750
650
550
RT
(ms)
.167 D
Mask
.104
S
.073
.073 D
.100
No Mask
.029
.040
S
.040
2
4
Target Numerosity

Figure 6

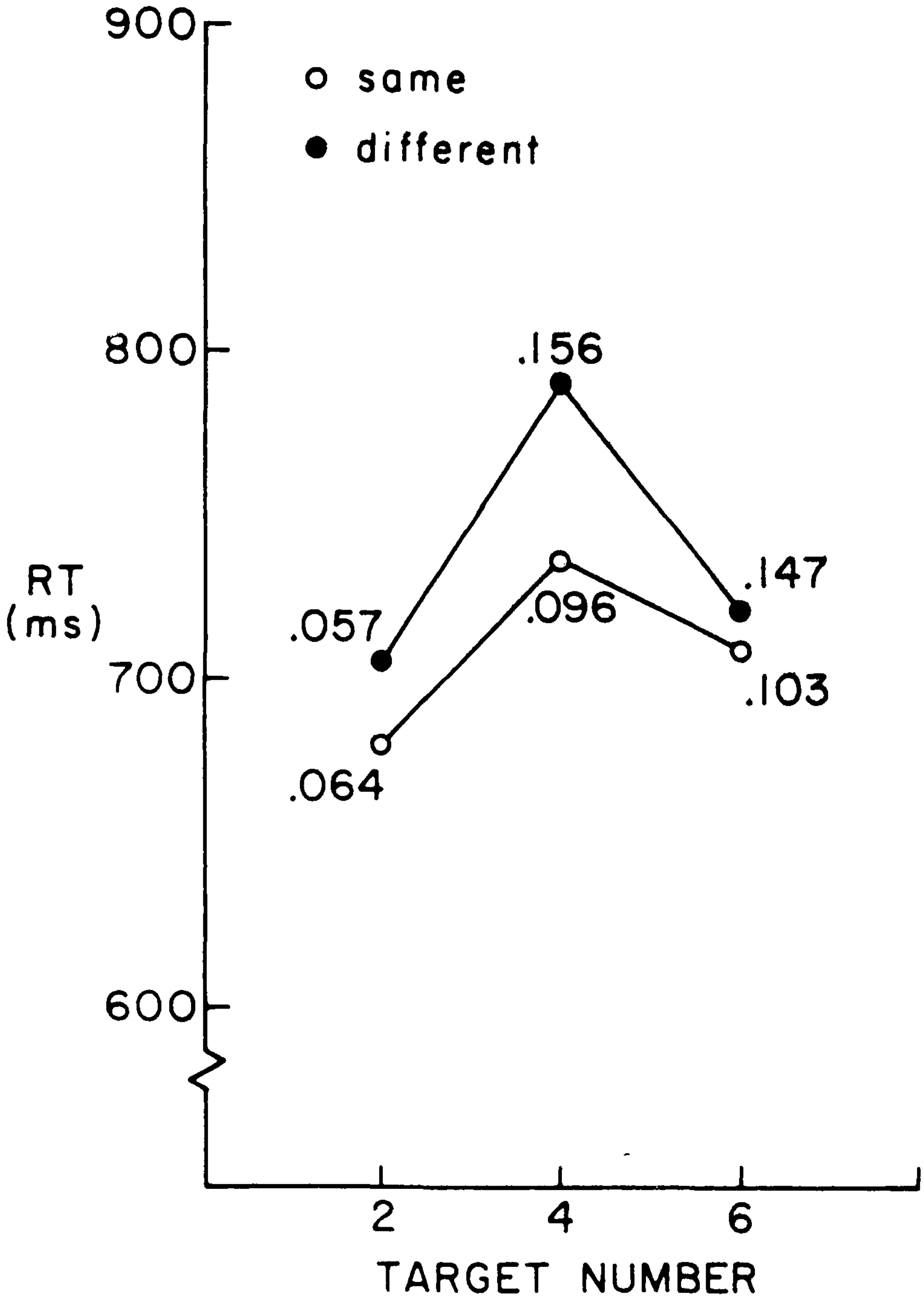
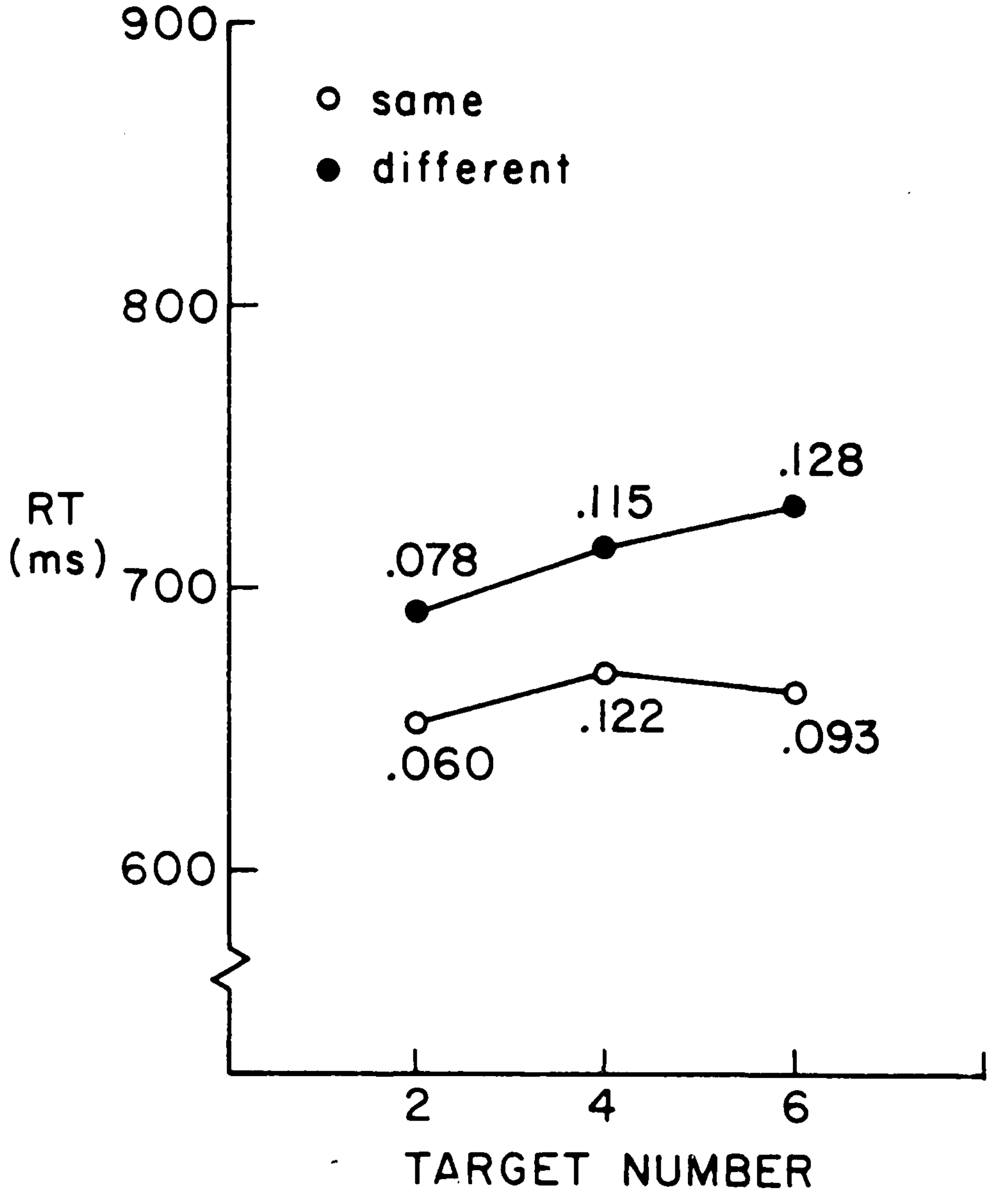

900
800
700
600
RT
(ms)
○ same
● different
.156
.147
.057
.096
.103
.064
2
4
6
TARGET NUMBER

Figure 7



900
800
700
600
RT
(ms)
same
different
.078
.115
.128
.060
.122
.093
2
4
6
TARGET NUMBER

Figure 8

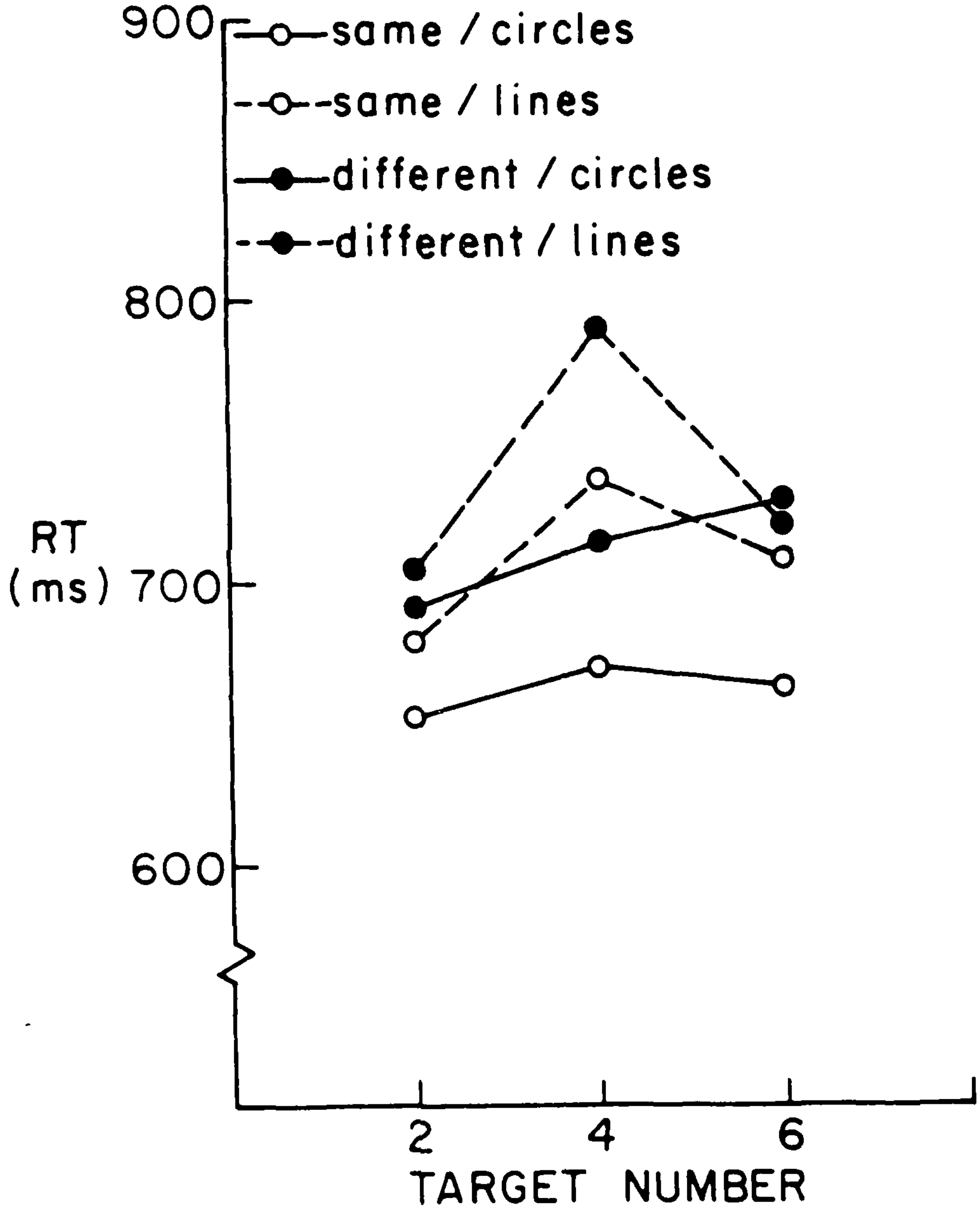

900
800
700
600
RT
(ms)
same / circles
same / lines
different / circles
different / lines
2
4
6
TARGET NUMBER

Figure 9

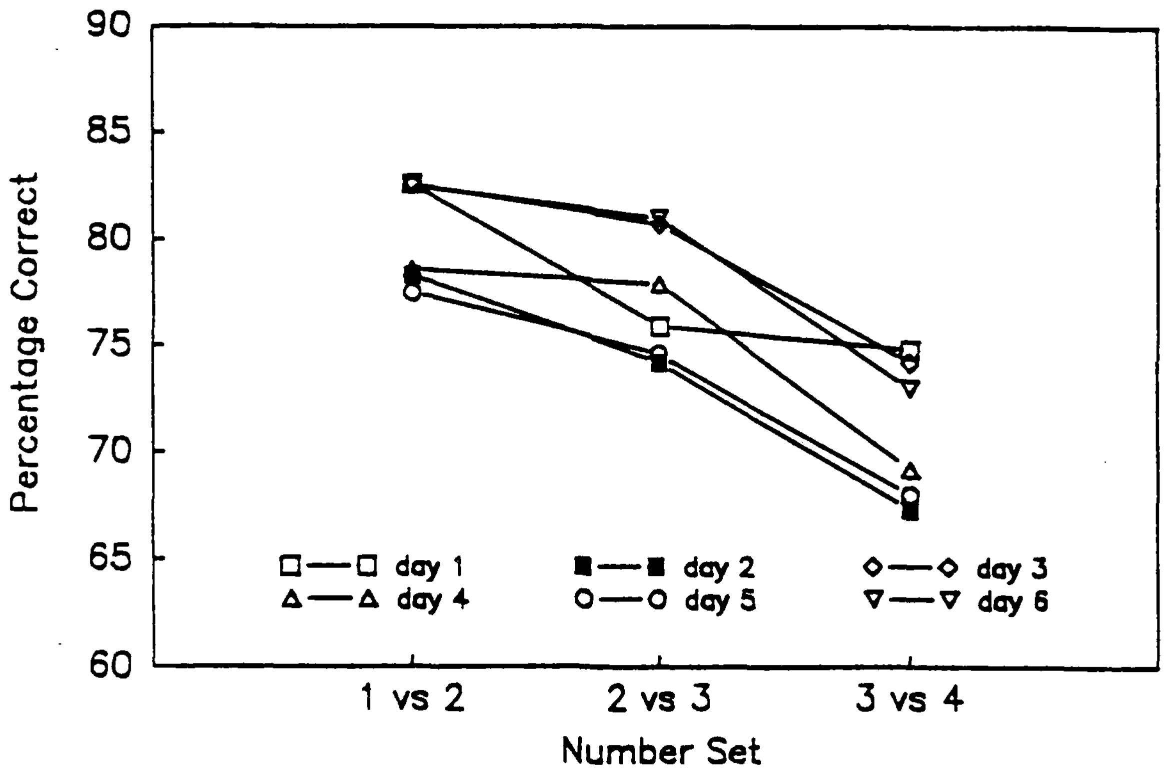

90
85
80
75
70
65
60
Percentage Correct
day 1
day 2
day 3
day 4
day 5
day 6
1 vs 2
2 vs 3
3 vs 4
Number Set

Figure 10

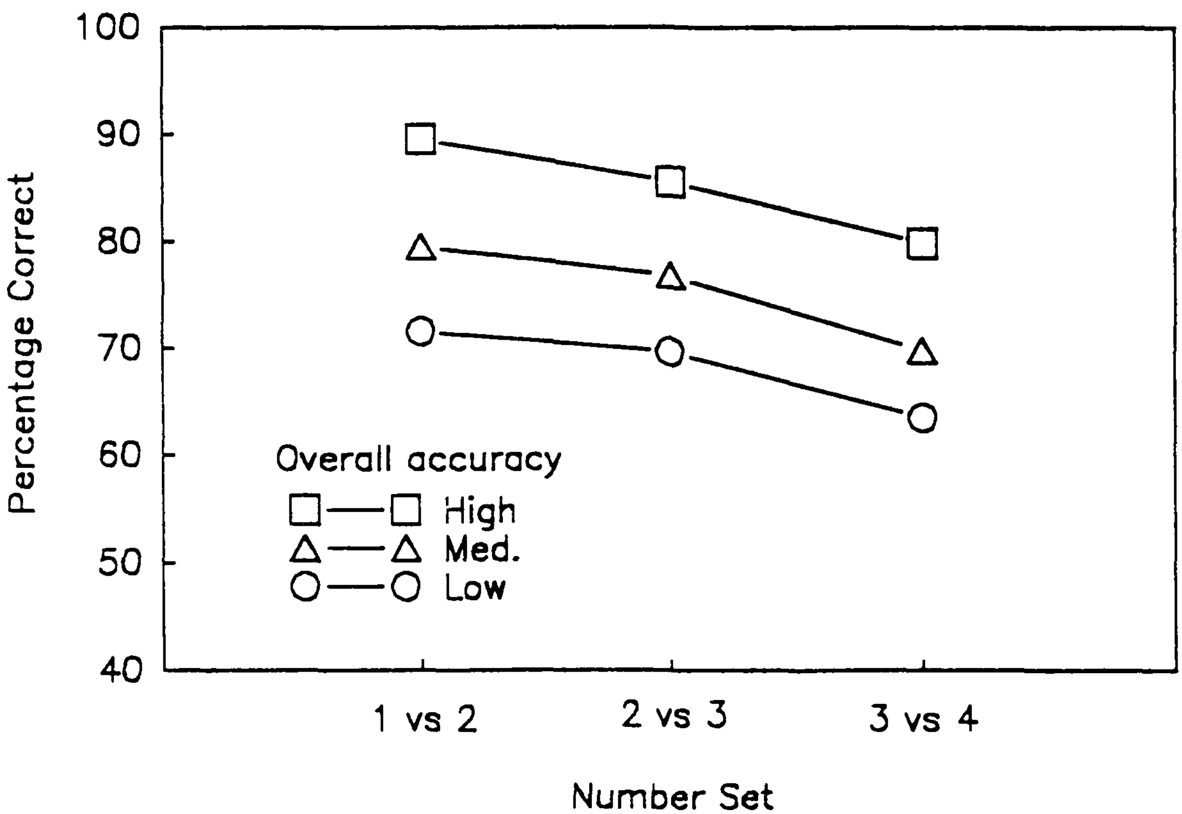

100
90
80
70
60
50
40
Percentage Correct
Overall accuracy
High
Med.
Low
1 vs 2
2 vs 3
3 vs 4
Number Set

Figure 11

END
DATE
9-88
DTIC