Perception, 1985, volume 14, pages 619-628

Detection versus discrimination of visual orientation

Dov Sagi, Bela Julesz
AT&T Bell Laboratories, Murray Hill, NJ 07974, USA
Received 21 June 1984, in revised form 3 November 1984

Abstract. The role of focused attention in vision is examined. Recent theories of attention hypothesize that serial search by focal attention is required for discrimination between different combinations of features. Experiments are reported which show that the mixture of a few (less than five) horizontal and vertical line segments embedded in an aggregate of diagonal line segments can be rapidly counted (also called 'subitizing') by a parallel (preattentive) process, while the discrimination between horizontal and vertical orientation requires serial search by shifting focal attention to each line segment. Thus detecting and counting targets that differ in orientation can be done in parallel by a preattentive process, whereas knowing 'what' the orientation of a target is (horizontal or vertical, ie of a *single* conspicuous feature) requires a serial search by focal attention.

1 Introduction

Since Aristotle, one way to define the role of vision is to know *what* is *where*. Indeed, there are several behavioral-anatomical reports on two visual systems, one for detecting and locating objects and the other for recognizing them (Held 1968; Trevarthen 1968). More recently, psychologists have suggested a rather different kind of two-visual-systems theory in which a preattentive parallel system discriminates easily between objects that differ in certain conspicuous features, while discriminating between objects that are defined by a combination of features requires serial search by focal attention (Bergen and Julesz 1983; Julesz 1981a, 1981b; Julesz and Bergen 1983; Neisser 1967; Treisman and Gelade 1980).

In these preattentive-attentive theories of vision the dichotomy is not because of the different tasks of detection versus discrimination, but rather because of detecting the difference between a single feature versus detecting differences between the combination of many features. Here we shall show that the preattentive-attentive dichotomy of vision holds for targets that differ only in a *single* feature. Thus the crucial point is the difference between knowing that a target is different from other targets in some feature versus knowing *what* this feature actually is.

To explain the rationale that motivated our studies we give a brief outline of the findings by Bergen and Julesz. Bergen and Julesz (1983) studied the detection of one +-shaped element in an array of many L-shaped elements, and found that this task could be rapidly performed even when eye movements were prevented and performance did not depend on the number of elements. One + among 35 Ls could be detected almost error-free in about 100 ms presentation time (ie a masking array has been presented 100 ms after the onset of the stimulus array that contained the + and Ls). In contrast to this parallel system, they observed that a T-shaped element in the array of Ls required time-consuming scrutiny, and for 35 Ls the task could not be performed at all, showing that the afterimage did not last long enough for this serial search process by focal attention. They inferred that detecting a T in Ls requires 30-50 ms additional time as the number of letters is increased.

It appears that in order to tell a T from an L one had to perceptually glue together the two orthogonal line segments in their exact relative positions. In preattentive vision this positional location between the line segments seems to be lost, and it is assumed that the positional glue (coupling) exists only in a small disk of focal attention. These findings support the texton theory of Julesz (1981a, 1981b) that assumes a preattentive parallel system and a serial attentive system. The preattentive system can detect only areas in which some local conspicuous features (textons) differ, while the way textons are glued together can be recognized only within a small aperture of focal attention. The role of the preattentive system is to direct this aperture of focal attention to these areas of texton differences.

Treisman and her collaborators (Treisman and Gelade 1980) have a similar model. They also assume the existence of conspicuous features and assume that differences (disjunctions) in these features can be detected in parallel, while detecting a certain combination (conjunction) of features requires serial search by focal attention.

Here we show that this dichotomy of vision still exists for a single feature. The essential task that requires focal attention is not only the need to discriminate between the relative position of line segments (in order to recognize a T or an L) or feature conjunctions, but also the task of knowing 'what' a certain feature is. In our study the single feature is the orientation of line segments. The experimental methods we employ do not rely on any of the theories mentioned above; thus the observed dichotomy of parallel versus serial processing has general implication for any theory of visual perception.

2 Experimental design

We were interested in the question whether the detection of a difference in the orientation of line segments is mediated by a different process from knowing their *actual* orientation. Our question can be answered by using two different experimental paradigms: one of *detection*, the other of *identification*. Performance on these two tasks can be compared with each other while the number of targets is varied. While these paradigms are well documented in the literature (Atkinson et al 1976; Chase 1978; Kaufman et al 1949; Rock and Gutman 1981; Treisman and Gelade 1980; Watson and Robson 1981), our experimental design differs from the earlier ones in significant ways. Our targets were a mixture of a few horizontal or vertical line segments embedded in a *large aggregate* of densely packed diagonal line segments. This design of detecting a few targets among many requires feature (orientation) discrimination at the earliest stage.

In the detection task we were not interested in knowing the actual orientations (horizontal or vertical) of the targets, but only their number. We capitalized on the known finding (Atkinson et al 1976; Kaufman et al 1949) that rapid counting of less than five targets (also called subitizing by Kaufman et al) can be done without error. However, our stimuli were followed by a mask after a short time, thus limiting processing time and bringing the visual system to its processing limit even at the 'subitizing' range. Thus we can compare this counting threshold to discrimination threshold of the same targets under the same conditions. For increasing number of targets we expect increasing processing time for a serial system; but constant processing time for a parallel system. This expectation might prove wrong in cases such as a parallel system with local interference, since in these cases a parallel system will show increase in processing time when the number of targets increases; however, this can be controlled and be equated on both tasks. Moreover, we introduce here a second experiment, where we test the effect of different temporal presentation of the targets while keeping their number constant.

It should be remembered that comparison is done between two tasks, with the same stimulus arrangements and targets. Thus the differences in performance should reflect the differences in the tasks and not differences in the stimuli.

3 Methods

3.1 *Stimulus generation*

The stimuli were displayed on the face of a Hewlett Packard 1310B cathode ray tube with P4 phosphor. The display was controlled by a PDP 11/23 computer via special-purpose hardware designed by W J Kropfl (described in Julesz et al 1976). The same PDP 11/23 controlled the experiments and collected the observer responses from the keyboard. The keyboard was also used by the observers to initiate the stimulus display at their own pace. The stimuli consisted of white line segments (0.8 deg of visual angle length and 4 min width) on black background. The line arrangement can be seen in figure 1. Stimulus diameter was 8 deg. The observers were seated at a distance of 1 m from the display.

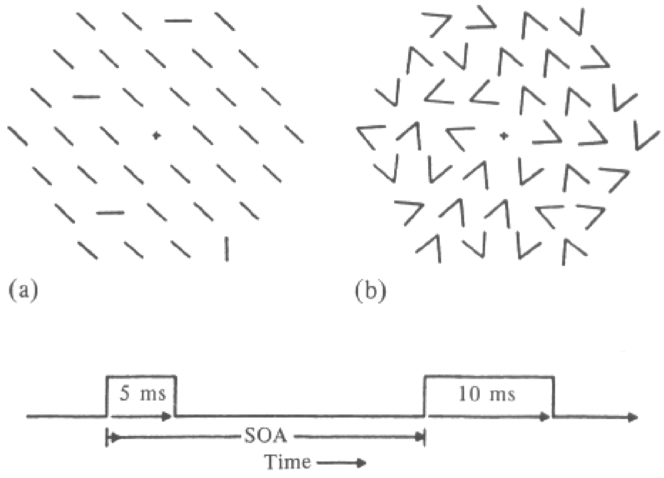

Figure 1. (a) One of the stimuli used in the experiments (in this case with four target lines); (b) the mask. The stimuli were presented for 5 ms and the mask for 10 ms. The SOA is defined as the time difference between the onset of the stimulus and the onset of the mask. The stimulus had a diameter of 8 deg, with white lines (0.8 deg length and 4 min width) on black background, and was viewed from a distance of 1 m.

3.2 *Stimulus presentation*

The stimuli were presented for 5 ms (see figure 1), preventing the possibility of more than one fixation during the exposure, but obviously visual persistence time was much longer. We limited the processing time available to the observer by masking each line segment with patterns of randomly oriented V-shapes, presented for 10 ms at variable times after the onset of the stimuli (this interval is called stimulus onset asynchrony, SOA). For short SOA the mask proved to be effective and performance was at chance level, but for longer SOA performance improved, reaching 100%.

3.3 *Psychophysical procedure*

In all the experiments the observers had to discriminate between two kinds of stimuli, and had the choice between two response alternatives. In the counting task the choice was between two numbers n (when n targets were presented) and $n+1$ (when $n+1$ targets were presented); in the discrimination task the choice was between 'same' when all the targets were the same and 'different' when one of the targets was different from the others. The two alternatives were equiprobable. In case of uncertainty the observers had to guess; feedback was supplied on each error. Error rates were calculated as the average of the error rates on the two alternatives, in order to eliminate subjective preference toward one of the alternatives (that is, tendency to guess more frequently toward one of the alternatives, thus reducing error rate on that alternative and increasing error rate on the other).

The experiments were in blocks of 50 trials. In each of these blocks the temporal properties of the stimuli were kept constant (the SOA in experiments 1 and 2, the delay in experiment 2). For the discrimination task the number of targets in a trial was kept constant through each block; for the counting task there were only two possible numbers of targets presented in each block. On each of the trials in a block the targets could appear in any of 36 locations with equal probability. However we kept a minimum intertarget distance of the size of three interline distances, so the separation between any two targets was always more than 3.2 deg of visual angle.

Each session lasted 1-1.5 h and contained blocks of different temporal parameters and numbers of targets. The two tasks were not intermixed and were performed in different order by different observers (that performed on both tasks). Experiment 2 was done after experiment 1.

Error rates were averaged from 2 to 9 blocks of trials, 3.5 blocks per data point on the average. Standard error (between blocks) of the error rates ranged from 1.5% on the average for error rates smaller than 10% to 3.6% on the average for error rates greater than 20%.

3.4 *Observers*

Four observers participated in these studies, all of them with normal or corrected-to-normal vision. One was the first author (DS), but the others (MW, TK, AD) were not aware of the purpose of the experiments. All the observers were well practiced in the experiments reported here; they were given as many trials as needed to practice, before data collection started.

4 Experiment 1

We used a display of line elements as shown in figure 1a. The display consisted of 36 elements, most of them diagonal (45°), except for one, two, three, or four, that were vertical (V) or horizontal (H). The vertical and horizontal lines were called targets. We used this display as stimulus and performed two classes of experiments, which differed in the task the observer had to perform. In the first class of experiments the observer had to count the number of target elements, irrespective of their orientation (V or H). In the second class of experiments the observer had to judge whether all the target lines had the same orientation or not. In both cases all the targets could be of the same orientation, or one could be different (when the number of targets was greater than one). In the counting task, in each block of trials, the observer had to decide between two numbers (1 versus 2, 2 versus 3, 3 versus 4); thus increasing the number of targets did not increase the number of possible responses, and so avoided an increase in response uncertainty. In the orientation discrimination task the number of targets was kept constant ($n = 2, 3, 4$) through each block of trials.

The stimuli were presented briefly (see section 3.2), preventing the possibility of more than one fixation during exposure. Visual persistence was limited by the masking frame that appeared at different SOAs after stimuli. We assumed that SOA was a measure of the processing time needed for a particular task.

For tasks that could be done by the parallel preattentive system we expected performance to be independent of the number of targets, but we assumed that processing time of the serial attentive system would depend on the number of targets.

In figure 2 we present psychometric functions obtained for both classes of experiments (for different observers). We can see an increase in the probability of correct response with increasing SOA. On the left side of each of the figures 2a, 2b, and 2c we plotted three curves for the counting task, corresponding to the number of counted targets (2, 3, 4). The three curves overlap, showing that counting from three

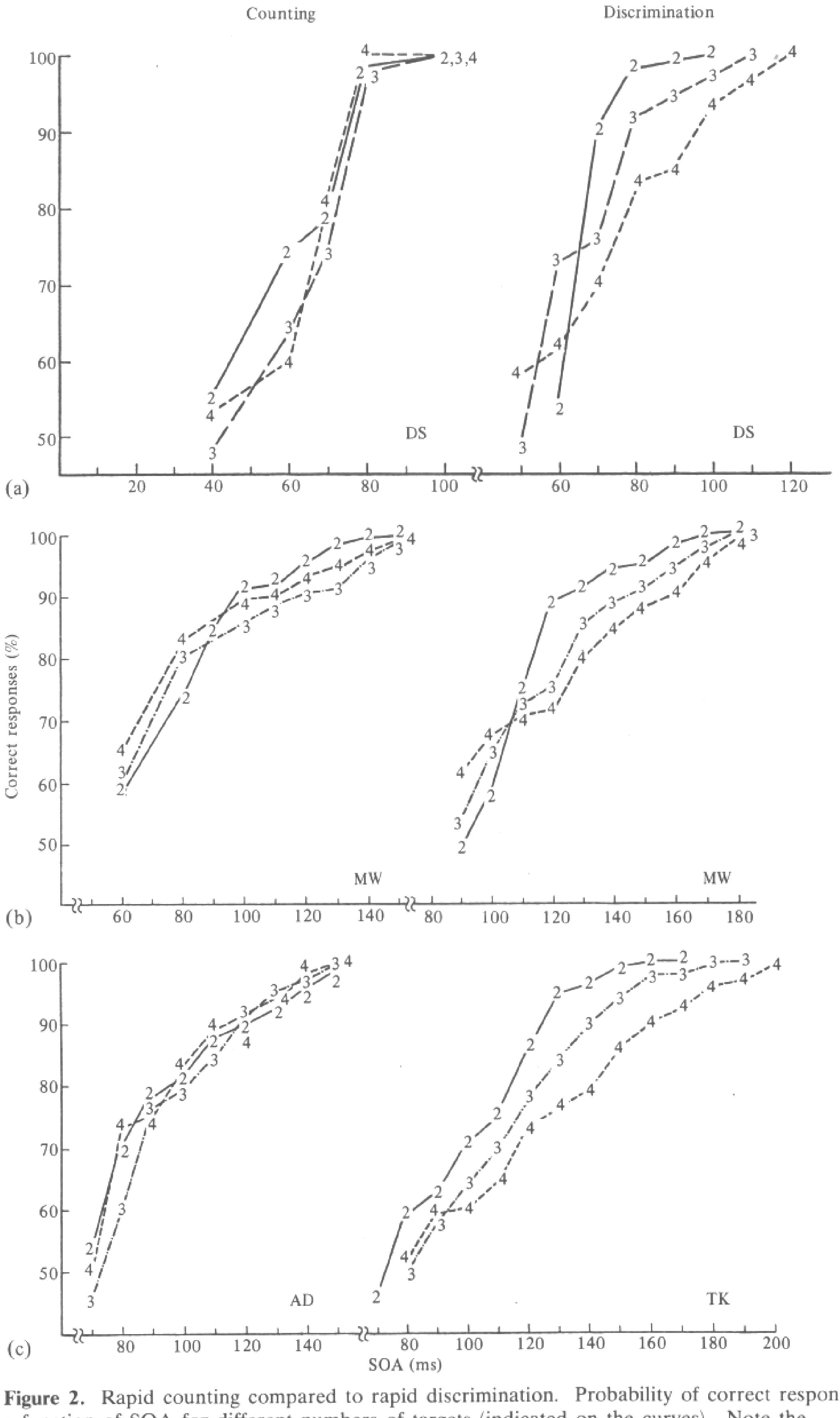

Figure 2. Rapid counting compared to rapid discrimination. Probability of correct response as a function of SOA for different numbers of targets (indicated on the curves). Note the overlapping curves for rapid counting (on the left), and the shift of the curves to the right with increasing target number for rapid discrimination (on the right). In addition to this shift, the slope of the curves seems to decrease when the number of targets increases. Observers: (a) DS, (b) MW, (c) AD on the counting task and TK on the discrimination task. Note that the different figures have different scales.

to four takes as much time as counting from one to two. The probability of correct responses was calculated as the average of correct response from the two alternatives (see section 3.3). Inspection of the data shows no significant differences between error rates on trials with n targets and $n + 1$ targets. (The average percent differences were 0.05 ± 1.11, -0.02 ± 1.35, 1.61 ± 0.83; standard error from 113, 62, 96 blocks for observers AD, DS, MW, respectively.)

On the right side of each of the figures 2a, 2b, and 2c we plotted the three curves for the discrimination task, again for different numbers of targets to be tested. This time the three curves are clearly not overlapping: as the number of targets increases more time is needed to reach high performance, although at the low-performance end the discrimination curves tend to overlap. Also, note that an increase of the number of targets in the discrimination task causes a decrease in the slope of the curves. This stems from the nature of the task. One can have above-chance performance in the four-target experiment at a SOA where two targets can be identified. Since the response is whether there is a different target or not, there is a 50% chance of having a different target between the two that can be identified. Thus one can have 75% correct response for the four targets, when correctly identifying only two, and 87.5% when correctly identifying three targets out of four. Hence the processing time required to identify the fourth target can be inferred only from the high-performance end of the psychometric curves; the low-performance end might reflect partial identification and is more sensitive to response uncertainty and strategy differences.

Figure 3 summarizes the results for the different observers (AD, DS, MW, TK). This time we plotted the SOA required to reach 95% correct response as a function of the number of targets. On the left, in the counting task for three observers, no dependence of processing time on n can be noted; the average slope for the three observers is 1.9 ms/target (6.3 standard deviation). On the right, in the discrimination task, there is a clear dependence on n. Surprisingly, it seems that the inspection rate (slope of the curves on the right) is similar for all observers, 16.6 ms/target on the average ($\pm$ 3.2 ms target, standard deviation of three observers), in spite of the relative big difference in the average SOA. The latter seems to be in the same order of magnitude as the counting time (for observers MW and DS).

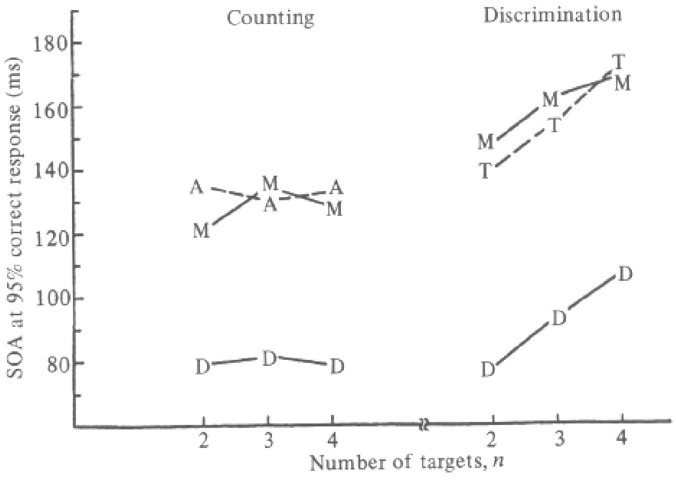

Figure 3. SOA required to reach 95% correct response as a function of the number of targets, for rapid counting (on the left) and rapid discrimination (on the right). Data collected for different observers (indicated by letters on curves: M for MW, D for DS, A for AD, and T for TK), taken from psychometric curves. Note that the curves for different observers are parallel for rapid counting, ie there is no dependence on the number of targets, while for discrimination each target adds 17 ms to the SOA.

The results presented above show that orientation discrimination can be described as a serial process, since adding targets increases the processing time. It is not likely that this additional time is required to overcome noise, since even when two targets can be discriminated perfectly there is considerable error rate in the four-target discrimination. The targets were kept apart by keeping at least two background elements between any two targets (that is a distance of more than 3.2 deg between any two targets), so local interference, if at all present, would be kept down to minimum. However, an alternative explanation can be given, namely that the increase in processing time is caused by the limited capacity of the system (Townsend 1971). Here the system divides its resources among all the targets simultaneously, thus giving less power to each analyzed target when there are more targets. With this latter explanation in mind we devised the second experiment.

5 Experiment 2

Here the targets were not presented at the same time, but some of them were delayed. The logic behind this experiment was that a parallel limited-capacity system might need all the processing time available—that is, all the targets would have to be processed in all the available time (here limited by masking)—while in a serial system only one target might be processed at a time. Thus in a serial system some of the targets could be taken out for a limited time without increasing the error rate, but not so in a parallel system.

Figure 4 shows the stimulus arrangement for the delay experiment. The line targets were presented in two frames, each frame containing in addition only half of the diagonal background lines selected at random to reduce target flicker. (A background element of the second frame had 0.5 probability of appearing on top of a background line of the first frame, and 0.5 probability of appearing on an empty position of the first frame; targets of the second frame appeared always on empty positions of the first frame.) The third frame was the same as in the previous experiments, serving as an erasure. The SOA, defined as the time interval between the onset of the first frame and the onset of the erasure, was kept constant. We measured the probability of correct discrimination (see section 3.3) as a function of the time delay between the first two frames, for different number of delayed targets. The total number of targets was always four in discrimination tasks, but in rapid counting experiments the observers were presented with three or four targets. In each block of trials the SOA, delay, and the number of delayed targets were kept

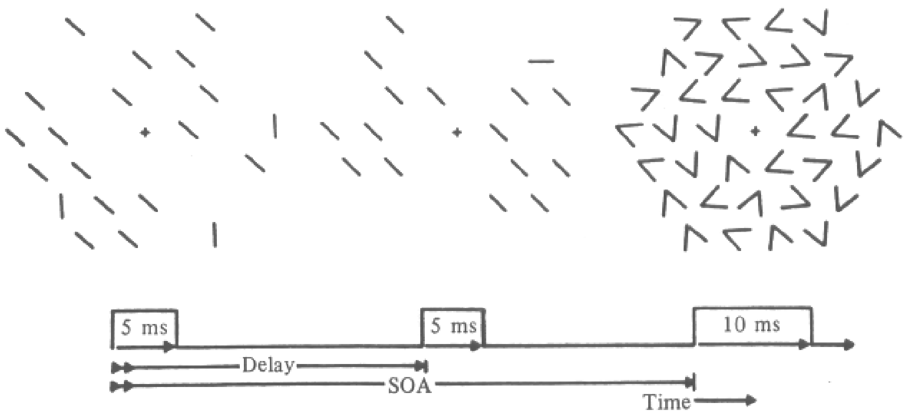

Figure 4. Stimulus arrangement used in the delay experiment. The display consisted of three frames, the first two contained the targets and the third was the mask. Delay time (variable) was measured as the time difference between the onset of the first two frames; SOA (constant) was defined as the time difference between the onsets of first and third frames.

constant. The SOA was chosen to be such that the probabilities of correct response would be around 90%, since shorter SOA might not reflect inspection of all the targets in the case of a serial system. As has been pointed out before (see section 4), one can have above-chance performance on the discrimination task by inspecting only some of the targets.

The results are shown in figures 5a and 5b (for observers DS and MW, respectively). On the left are data for the rapid counting experiment, where we can

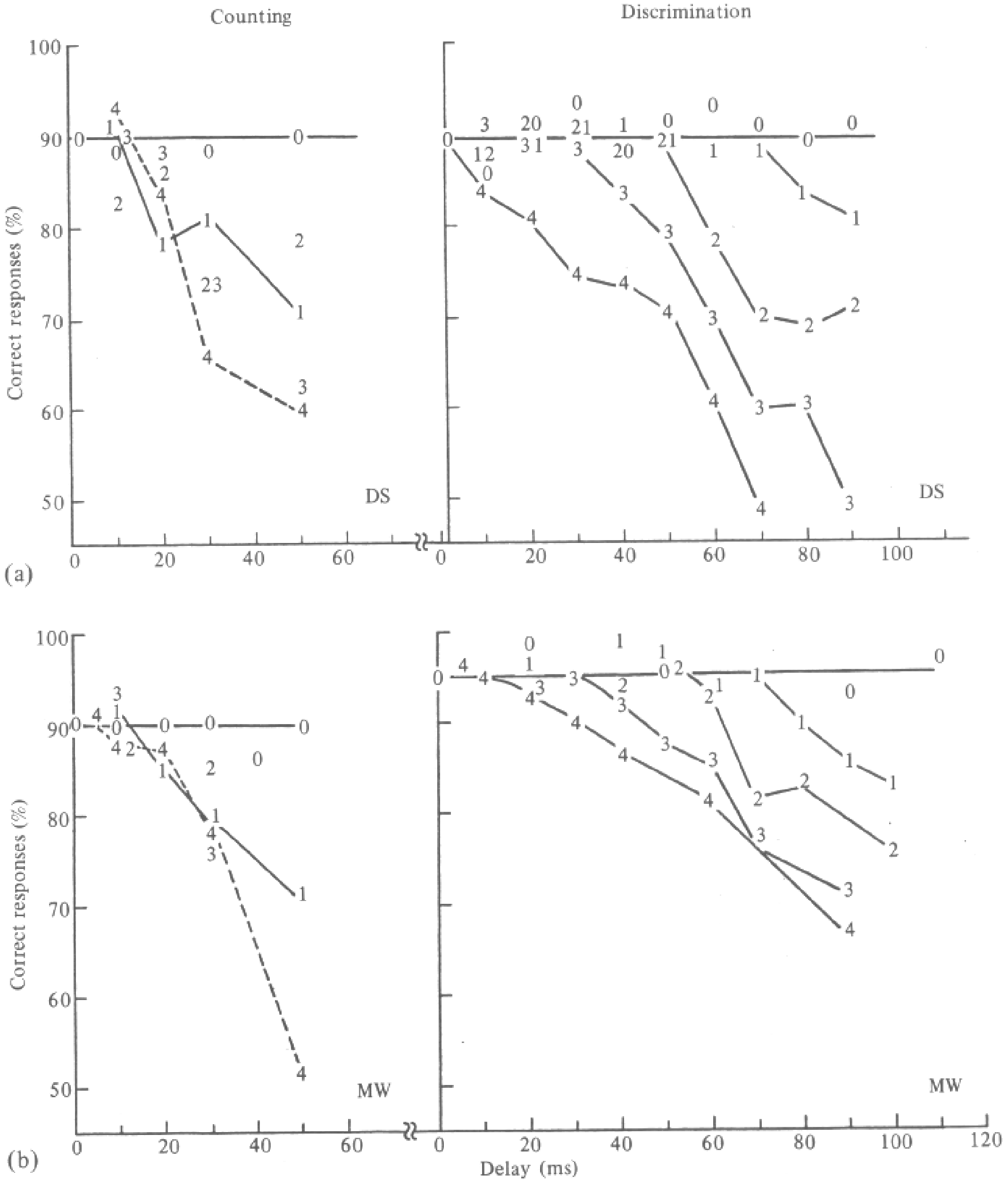

Figure 5. Effect of delaying a certain number of targets (indicated on each curve). The total number of targets was three or four for counting and four for discrimination. SOA was constant: 80 ms for counting and 110 ms for discrimination for observer DS (a), 130 ms for counting and 180 ms for discrimination for observer MW (b). The different curves show the percentage of correct responses as a function of the delay, each curve representing a different number of delayed targets. Note that delaying one or four targets for rapid counting has the same effect, while for discrimination one target can be delayed for a significantly longer time than four targets. The 'zero' data points are for cases where the second background frame was delayed without any target in it, and show that this in itself had no effect

see the probability of correct response as a function of delay when delaying four (whenever four targets were presented), three, two, one, or zero targets (the last one as control to ensure that delaying the background in the second frame has no effect). The two curves (1 and 4) are roughly the same: both show a rapid increase in error rate with increasing delay. However, on the right side of figures 5a and 5b there is a different behavior. Here, data for orientation discrimination (and delaying 4, 3, 2, 1, and 0 targets) are presented, and one target can be delayed for as much as 70 ms without damaging the performance. (This time is even longer than that predicted from the slope of the discrimination curves in figure 3, that is 3 × 16.6 or 50 ms.) Furthermore, when more targets are delayed this critical time is shortened by 20–30 ms per each additional delayed target.

Thus, according to the logic of the delay experiment, rapid counting is done in parallel while the task of orientation discrimination is consistent with a serial model.

6 Conclusions

The results of our experiments show a clear difference between counting (or detection) tasks and tasks of orientation discrimination of line elements. Our study, and thus the conclusions, are confined to a small number of targets ($n \leqslant 4$). In the literature (see references in Chase 1978) counting in this range is known to be faster than in the range $n > 4$, and was coined 'subitizing' by Kaufman et al (1949). However, Chase and Kaufman et al report an increase in reaction time of 25–100 ms per item counted even for $n \leqslant 4$, compared to 300 ms per target for $n > 4$. Those authors used experimental paradigms very different from ours, measuring reaction time in the order of seconds, with the number of possible responses increasing with the number of targets to be counted. We kept the responses down to two by presenting n versus $n+1$ targets. One might wonder whether this task requires really counting, or only discrimination of more versus less; however, in both cases the observers have to possess information from all the stimulus targets. Furthermore, our interest was in very fast processes that might reflect the behavior of the visual system at the very early stages of form perception.

It is important to note that our observers counted local changes in feature densities (eg how many lines differ in orientation from the background), and not local changes in light intensity, where the latter can be explained simply by intensity integration over the visual field. This restriction in our experiment might help to rule out some of the hypothetical explanations for subitizing raised by Atkinson et al (1976). For example, an integration by low-spatial-frequency channels over a wide field might be sensitive to the number of light spots, but insensitive to local orientation changes. Atkinson et al also raised the possibility, in agreement with us, that counting units are the same units as those operating in tasks of perceptual segregation where an object must be located in visual space. However, our results with *many* narrow line segments show that rapid counting can be performed by using high-spatial-frequency channels. Furthermore, since we used a masking technique, we were able to separate attentive processes from preattentive processes, showing that locating targets can be done preattentively, while it was shown by Bergen and Julesz (1983) that detecting positional relationship between lines requires serial search by focal attention. Thus locating object in space might precede detection of the positional relationship among its parts. However, the parallel preattentive system is limited by its ability to detect only feature gradients. Hence locating objects that do not differ in any feature from their background might require knowing what these features actually are (identification)—a task performed by serial search.

The finding that local changes in feature densities (eg differences of orientation between spatially adjacent line elements) can be detected simultaneously might reflect their very importance at those first stages of vision. Interestingly, the direction of this feature gradient is not available at this early stage, but requires some time-consuming processing which is local and serial in its nature. It is tempting to assume that this serial scan is directed by the earlier parallel process.

In summary, our findings imply that detection of conspicuous targets and their rapid counting can be done in parallel, while their identification (discrimination) can be done only in the aperture of focal attention by serial scanning. What is surprising is the observation that detecting a line segment with different orientation from its surround can be done rapidly, regardless of the number of the line segments, while knowing its actual orientation requires step-by-step scrutiny.

References

Atkinson J, Campbell F W, Francis M R, 1976 "The magic number 4 ± 0: A new look at visual numerosity judgements" *Perception* **5** 327-334

Bergen J R, Julesz B, 1983 "Parallel versus serial processing in rapid pattern discrimination" *Nature (London)* **303** 696-698

Chase W G, 1978 "Elementary information processes" in *Handbook of Learning and Cognitive Processes* Ed. W K Estes (Hillsdale, NJ: Lawrence Erlbaum Associates) volume 5, pp 19-90

Held R, 1968 "Locating and identifying: Two modes of visual processing" *Psychologische Forschung* **31** 338-348

Julesz B, 1981a "Textons, the elements of texture perception and their interactions" *Nature (London)* **290** 91-97

Julesz B, 1981b "A theory of preattentive texture discrimination based on first-order statistics of textons" *Biological Cybernetics* **41** 131-138

Julesz B, Bergen J R, 1983 "Textons, the fundamental elements in preattentive vision and perception of textures" *Bell System Technical Journal* **62** 1619-1645

Julesz B, Breitmeyer B, Kropfl W, 1976 "Binocular-disparity-dependent upper-lower hemifield anisotropy and left-right hemifield isotropy as revealed by dynamic random-dot stereograms" *Perception* **5** 129-141

Kaufman E L, Lord M W, Volkman J, 1949 "The discrimination of visual number" *American Journal of Psychology* **62** 498-528

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

Rock I, Gutman D, 1981 "The effect of inattention on form perception" *Journal of Experimental Psychology* **7** 275-285

Townsend J T, 1971 "A note on identifiability of parallel and serial processes" *Perception & Psychophysics* **10** 161-163

Treisman A, Gelade G, 1980 "A feature integration theory of attention" *Cognitive Psychology* **12** 97-136

Trevarthen C B, 1968 "Two mechanisms of vision in primates" *Psychologische Forschung* **31** 299-337

Watson A B, Robson J G, 1981 "Discrimination at threshold: labelled detectors in human vision" *Vision Research* **21** 1115-1122