Focused attention in the perception and retrieval of multidimensional stimuli

ANNE TREISMAN

Department of Experimental Psychology, University of Oxford, Oxford, England

This paper reports some further experiments on successive matching of multidimensional stimuli in which the correct conjunctions of features must be specified; it also modifies and extends the model proposed earlier by Treisman, Sykes, and Gelade (1977). The results obtained in the previous experiment were replicated despite a change from fixed to varied targets, and from spatial to temporal separation of the targets, thus extending their generality. The modified conjunction-matching model proposes that the subject progressively narrows his focus of attention, first to deal with a single display stimulus at a time and finally, if necessary, to a single target. The model was tested quantitatively by simulating the separate stages hypothesized to underlie conjunction-matching in some simpler, single-attribute-matching conditions, and using the differences between latencies in these conditions to predict differences between latencies in the appropriate conjunction-matching conditions. The results were consistent with the general hypothesis that one important role of focused attention and serial processing is to integrate separable attributes or features into the correct conjunctions, which correspond to the objects actually presented or stored.

Most of the objects, people, and scenes we perceive produce complex, multidimensional, changing patterns of stimulation on the retina. At least some of their attributes appear to be registered by independent neural channels, specializing in different aspects, such as orientation, color, spatial frequency, brightness. Analysis of complex stimuli into more elementary features may be logically necessary in order to allow generalization across instances of the many categories we form and, at the same time, discrimination between them. But this immediately raises the question of how the component properties are resynthesized into the correct compounds, so that we correctly see the shirt as blue and the trousers as gray, for example, rather than the reverse. Treisman, Sykes, and Gelade (1977) offered some preliminary suggestions and findings relating to this problem, in the context of visual search and visual matching. This paper reports some further investigations of the matching task, which extend the generality of the findings; it then attempts to develop and test the model we proposed earlier.

There were two main areas of interest in the previous paper: (1) The first was the hypothesis that separable dimensions or features [i.e., those that are processed by independent analyzers (Garner, 1974)] are perceptually integrated into the correct compounds by means of focal attention and serial processing of information from different spatial locations. Features which are registered in the same location and temporal interval, i.e., within the same central focus of attention, can then be coded as belonging to the same object. This implies that

parallel processing and divided attention to multiple inputs should be possible only when either of two conditions is met: (a) the inputs can be distinguished by differences in a *single* feature, or (b) they vary along different dimensions, so that erroneous interchanges of features leading to illusory conjunctions cannot occur. The first experiment we reported tested the prediction that serial processing would be required in a visual search task if the target differed from nontargets only in its particular combination of features, while if it could be distinguished by a single feature processing would be parallel. The results were consistent with this hypothesis.

(2) The second problem we began to explore was how conjunctions, once perceived, are held in working memory and matched in a recognition task. We noted three possible methods. The first depends on a unitization process (cf. Laberge, 1973), in which higher level units are formed to encode directly the formerly separable features into new entities. The second involves some form of spatial tagging, allowing all features sharing the same location tag to be retrieved together, as when a knitting needle is used to retrieve a selected subset of punched cards. A final possibility is the formation of associative links either between the separable features themselves or between these features and a higher level unit, such as (but not necessarily) a verbal label. The last two forms of storage, in which the features retain their independent status, would allow the possibility of errors due to loss or confusion of either spatial tags or associations, resulting in the formation of temporary or permanent "illusory conjunctions." The experiment

used a successive "same-different" matching task with two targets (the same throughout the experiment) and two display stimuli, varying from trial to trial. In one condition, these were schematic faces, and in the other, they were colored letters. In each case, the targets and display stimuli varied along two attributes, the eyes and mouths of schematic faces or the shape and color of letters. The two targets differed from each other on both attributes, and so did the two display stimuli presented on any one trial. Subjects decided as quickly as possible whether either display stimulus exactly matched either target. The present paper reports a further series of experiments using the same paradigm and explores the theoretical implications of the results more fully. We therefore reproduce the mean response times in Figure 1 and summarize the conclusions we drew in some detail. (The lines fitted to the data in this figure differ from those in the previous paper; they will be explained in the final section.)

There were two main points of interest in the results. The first is the higher error rates (about 15% compared to 2% or 3%) that occurred when the color (or mouth) of one target item was conjoined with the shape (or eyes) of the other in the same display item, and the even higher rate (over 30%) when both target attributes were present in both display stimuli but in the wrong combinations (e.g., target pink "O" and blue "H" followed by display items pink "H" and blue "O"). These "conjunction errors" suggest that the attributes did retain some form of independent status in memory rather than fusing completely into unitary compounds or gestalts. They would be consistent either with spatial tagging or with associative links between attribute values which had formed particular conjunctions. The errors could then result either from loss of the labels or associations or from incomplete matching which terminated after the matching attributes were detected but before the spatial tags or associations were checked.

The second finding of interest concerned the response latencies. With both colored letters and faces, we found evidence that three additive components contributed to the total RTs: for each of the two attributes, a component was added whenever matching values were present; and a third component was added whenever a pair of these matching values were spatially conjoined in one display item. On the other hand, no significant addition to RT resulted when a pair of matching values were spatially joined in a target item but separated in the display.

It is clear that many questions arise from these preliminary experiments, some of which we attempt to clarify in this paper. It is divided into two parts: the first reports two experiments that replicate the main findings under different conditions, to extend their generality and to explore further the processes in-
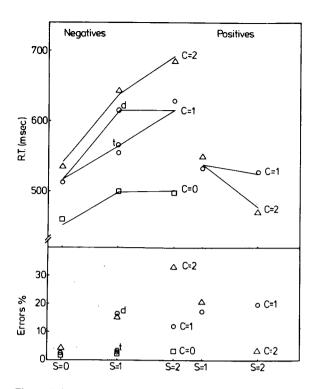

Figure 1. Mean RTs and errors in previously published experiment (Treisman, Sykes, & Gelade, 1977) on colored shapes. The abscissa shows the number of matching shapes ("s"), and the parameter is the number of matching colors ("c"). "d" gives the latency when one matching shape and one matching color are combined in the display but separate in the targets, and "t" gives the latency when they are combined in a target but separate in the display. The fitted lines are predicted from the model to be described in a later section.

volved. The second part attempts to modify, extend, and further test the detailed model we had proposed to account for conjunction-matching.

EXPERIMENT 1

One obvious question is how far the results depend upon keeping the target items fixed throughout the experiment, allowing a long-term memory code to be established for each of them. Would the asymmetry we found between the effect of wrongly paired matching attributes in the targets and in the display still hold when the targets as well as the display stimuli varied randomly from trial to trial? In the first experiment, we check this possibility.

Method

Stimuli. Eight different target cards were prepared, containing examples of each of the 16 color-shape combinations possible with the colors blue, pink, orange, and green and the letters "O," "H," "X," and "N." The letters were drawn with the same stencils and in the same colored inks as those in the previously published experiment. Each letter subtended 0.8° and was centered 0.7° from the fixation point. The two target letters were placed above and below the fixation point, and the two test letters on

each trial were to the left and right of the fixation point. Each color and each shape were placed equally often in the upper and lower positions on the target card. For each of these target cards, a set of 19 test cards was prepared, one example of each type of negative and two of each type of positive trial. The different trial types were the same as those described in the previous experiment; the description is repeated here to facilitate understanding.

There were 11 different types of negative trials and 4 types of positive trials. They are labeled by pairs of digits as follows: The first digit gives the number of colors, (0,1, or 2) that were shared by target and display, and the second digit gives the number of shared shapes. These are prefixed by the letters "N" for negative or "P" positive. The N(1,1) trials are further subdivided by using the two letters "t" and "d" to indicate whether the two matching values (one on each attribute) were spatially combined in the targets ("t") or in the displays ("d"). If they were spatially separate in both, no further symbol is used. Thus, for the targets blue "O" and pink "H," an example of the test trials labeled N(1,1) would be blue "N" green "H"; an example of N(1,1)t trials would be blue "N" and green "O"; and an example of N(1,1)d trials would be blue "H" green "N." The negative trials N(1,2), in which one color and two shapes matched, and N(2,1), in which two colors and one shape matched, necessarily contained a shape and a color which were combined in memory but separate in the display and a shape and color which were combined in the display but separate in memory [for example N(1,2) might be blue "H" and green "O"]. In trials N(2,2), both shapes and colors matched but were wrongly paired (blue "H" and pink "O"). In the positive trials, at least two matching values, one on each attribute, were correctly combined, although in P(2,1) trials there would be an additional matching color whose shape did not appear in the targets and in P(1,2) trials the reverse. P(2,2) trials had both target shapes and colors correctly combined.

For each trial type in which there were matching colors or shapes, these appeared equally often in each possible pair of target and display positions, (e.g., for P cards, the matching target was at the top and the matching display stimulus on the left in 2 of the 8 cards, the target was at the top and the display stimulus on the right in 2 cards, and so on). The number of times each particular color and each particular shape were presented was equated for both positives and negatives over the whole set of 152 test cards.

Subjects. The subjects were undergraduates or research staff at the university. They had not previously taken part in any experiment in this series, although four had done reaction time experiments before.

Apparatus. The cards were presented in a three-field tachistoscope. The background field contained a small black square in the center of the card, and was used to control fixation. The illumination of each field was the same. The onset of the display also triggered a timer, which was stopped by the subject's pressing one of two response keys.

Procedure. The instructions required the subject to press the positive (right-hand) key if one or both of the display stimuli exactly matched one or both of the targets, and the negative (left-hand) key otherwise. He was asked to respond as quickly and accurately as possible. Each subject was shown several examples of positives and of negatives to ensure that he understood the task.

Each trial consisted of the following sequence of events: The subject fixated the background card; then the experimenter said "Ready" and started the target presentation, which lasted for 1 sec. This was immediately followed by a test display, also presented for 1 sec. The RT was recorded, the subject was told whether the response was correct or an error, the target and display cards were changed, and the next trial was begun. The set of 8 target cards was repeated 19 times in a fixed order within a block. (The reason for the fixed order was simply that it allowed faster presentation of successive trials, using an automatic card-change for the target cards.) The display cards appropriate to each different target card were presented in random order. The order

of the target cards was different for each new block, and subjects, when questioned afterwards, claimed not to have learned, nor in most cases even to have noticed, the repeated order within each block. The set of 152 test displays was presented in a different random order to each subject, and in each block.

Each of five subjects was tested for three sessions consisting of about 20 practice trials on a random selection of cards, followed by the complete set of 152 experimental trials. A 5-min rest was allowed half-way through, plus any further breaks the subject requested. Each session took about 45 min.

Results and Discussion

The mean RTs are shown in Figure 2. The data points are joined by lines to bring out the relationships between trial types, and to facilitate comparison with the other experiments. The pattern is very similar to that found in the earlier experiment with targets kept constant throughout the experiment. A between-groups analysis of variance on the two experiments showed no difference between them and no interaction of experimental groups with trial types, although the differences between trial types were highly significant, $[F(14,168) = 28.30, p < .001]$. Thus there is no evidence that using targets which varied from trial to trial produced any differences in strategy or in performance. The only trial types which gave a mean difference of more than 20 msec in the two experiments were N(0,1) and N(0,2), which were both about 35 msec slower in the present experiment. However, this difference was due to only two of the subjects and was not statistically significant. The trial types of most interest to us, i.e., those that contained one or more matches on both attributes,
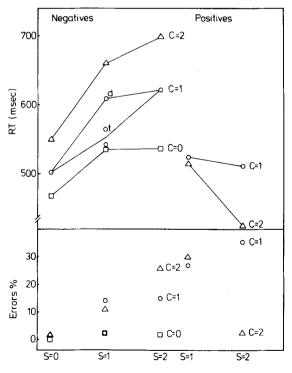

Figure 2. Mean RTs and errors in Experiment 1.

show a very similar pattern of variation across the two experiments. Error rates did not differ significantly across the two experiments either. Thus there is no evidence that either any idiosyncratic features of the two fixed targets in the previous experiment or the opportunity for long term learning of the target stimuli determined the asymmetry between conjunctions in the display and in the memorized targets. The fact that latencies could be as short as 470 msec in condition N(0,0) with varied targets appears incompatible with verbal coding and comparisons, since four names (two colors and two shapes) would have to be found and compared to four others before the response could be made.

EXPERIMENT 2

We turn now to a second variant of the paradigm so far studied. We have suggested that, in perceiving complex stimuli which are physically present, we process spatial locations serially, thus providing the "glue" which ties together different attribute values in the correct conjunctions. It is possible that stored information about spatial location is also important in the retrieval of memorized conjunctions. On the other hand, it may be that spatial tagging plays its role only during perception, to control the temporal sequencing of focal attention to different stimuli, and that some other mechanism takes over the task of holding together the features of each identified stimulus. One way to test these alternative hypotheses would be to present the target stimuli successively in the same location instead of simultaneously in separate locations. If performance changes or deteriorates, this would suggest that information about spatial location plays a role in specifying the target in memory as well as in ordering their initial encoding. If so, its role could be further clarified by testing whether it exclusively affects conditions involving mismatched conjunctions, or whether its absence impairs all conditions equally. In effect, the successive presentation conditions would mimic what we assume the subject achieves with simultaneous, spatially separate targets by processing the two targets sequentially. Our hypothesis would not therefore predict any difference between the two conditions. Experiment 2, then, compares the efficacy of time of presentation and of spatial location in providing the "glue" to retain conjunctions of attributes.

Method

Stimuli. The same varied target items with their appropriate display stimuli were used as in Experiment 1. However, there were two different presentation conditions: The two target items were either shown simultaneously at the top and bottom of a single test card, as in Experiment 1, or they were shown successively in the same position, either top or bottom. The two

display stimuli were always, as before, spatially separate and presented simultaneously.

Subjects. The subjects were again undergraduates or research staff at the university. Two had previously taken part in Experiment 1.

Procedure. This was the same as in Experiment 1, with the following exceptions. There were four separate presentation conditions for the target items. In condition SimL (simultaneous, late presentation), the subjects saw a blank card for 1 sec, followed by two targets on the same card (one at the top and one at the bottom) for 1 sec, immediately followed by the test display, also for 1 sec. Condition SimE (simultaneous, early presentation) was the same except that the blank card followed instead of preceding the target card. In condition SeqT (sequential, top position), the two target items were presented successively for 1 sec each in the same spatial position at the top of the card. In condition SeqB (sequential, bottom position), the two targets were presented successively at the bottom of the card. The two conditions with simultaneous presentation, either in the first or the second 1-sec interval, were necessary in order to equate the mean time delay between target and display stimuli to that which occurred in the successive presentation conditions. The two conditions with successive presentations were necessary in order to test the upper and lower spatial positions equally often, so as to match the spatial difficulty with the simultaneous presentation condition in which both positions were used on each trial.

Eight subjects each did one block of 152 trials in each of the four conditions, the order being counterbalanced by a repeated Latin square. Each subject was given 76 trials for practice with successive presentation and 76 with simultaneous presentation of the targets before beginning the experiment proper.

Results

The mean RTs are shown in Figure 3. Analysis of variance showed no significant effect of the presentation method on either latencies or errors, and no interaction of this factor with trial types, although the effect of trial types was again highly significant [F(13,91) = 16.62, p < .001]. A separate analysis carried out on conditions SimE and SimL showed no effect of the 0- vs. 1-sec delay between targets and display stimuli, and no interaction of delay with trial types. An analysis of conditions SeqT and SeqB failed to show any effect or interaction with position of the targets. The pattern of latencies for different trial types is very similar to that in the previous experiments. The overall means were analyzed together with those of Experiment 1, and no significant difference between the experiments or interaction with experiments emerged. Thus, temporal coding of conjunctions of attributes appears to be neither more nor less efficient for memory storage than spatial coding.

Discussion

This experiment has extended the generality of the pattern of results to a new experimental condition, showing that time of presentation can be used as efficiently as spatial location to code the conjunctions into memory. This strengthens the hypothesis that spatial position is used only during presentation to direct a temporal scan of stimuli, and
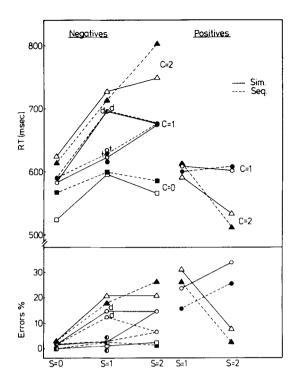
Figure 3. Mean RTs and errors in the two conditions of Experiment 2.

that it is this temporal scan which is responsible for correctly conjoining the separable attributes or features present in the display.

A MODEL FOR CONJUNCTION-MATCHING

We now explore in more detail the model we proposed for conjunction matching. We suggested, in our previous paper, that three stages of processing could be distinguished: stage 1 tested for the presence of a match on one attribute; if one was found, stage 2 tested for the presence of a match on the other attribute for that particular display item; finally, if a match was also found on that attribute, stage 3 checked whether that particular conjunction of display values was also conjoined in the memorized targets. These three stages were carried out sequentially and in a self-terminating fashion, so that the subject responded as soon as he had sufficient evidence to decide which key to press. Schematic faces differed from colored shapes in that each attribute (eyes and mouth) of each display item appeared to be checked serially for the schematic faces, while for colored shapes some parallel processing appeared to occur. This pattern for the colored shapes, in which a match on a second display item added less to the latency than a match on the first, was replicated in the present experiments, and cannot therefore be attributed to overlearning of the target items or to

their simultaneous and spatially separate presentation. In the previous paper, we tentatively suggested that the first stage, at which subjects searched for a matching value on either attribute, might be spatially parallel, at least for color. Subjects would then narrow their attention to focus serially on each display item that did contain a matching attribute in order to check whether it also matched a target item on the other attribute.

We now wish to develop further this idea that comparing conjunctions requires a narrowing from divided to focused attention, and to use it to specify in more detail the nature of the conjunction check we ' ad hypothesized as the third stage of processing. We suggest that the final conjunction check is, in fact, accomplished by a further narrowing of attention, to focus on a specific target as well as on a specific display item. Thus, at Stage 1, one attribute is processed for both targets and both display items; if a match is found, processing narrows down at Stage 2 to deal with one display item at a time, and the other attribute of that display item is checked against both targets; finally, if Stage 2 also proves positive, Stage 3 further narrows the focus to deal with one target at a time, and the first attribute of the selected target is *re*checked against the selected relevant display item. Stages 2 and 3 are repeated for the second display item if a matching value on the same attribute was also found in that location at Stage 1. This order of progressive focusing, first on display items and then on targets, would explain the asymmetry between the effects of false conjunctions in the display and those in the targets. The sequence would be self-terminating, and a response would be made as soon as a negative decision was reached for both display items at any of three stages, or as soon as a positive decision for either display item was reached at Stage 3.

An example may help to illustrate how this sequence of operations might be used to decide whether a matching conjunction was present. Suppose the targets were pink "O" and blue "H," and the display stimuli on an N(1,1)d trial were pink "H" and green "X." One of the sequences the subject might follow would be (1) to detect that one color (pink) matched a target and the other did not; (2) to locate the pink item in the display and to check whether the H shape which was conjoined with it there was also present in the memorized targets; (3) having detected its presence, to locate it in the targets and to decide whether the color blue associated with the "H" target in memory was also conjoined with the "H" stimulus in the display. The first two checks here would be positive, but the third negative. On an N(1,1)t trial, on the other hand, only two stages would be required. If, for example, the display stimuli were pink "N" and green "O," the sequence

might be (1) as above, (2) as above except that the test for "N" in the targets would prove negative, and would determine the response "Different" without the Stage 3 check. Thus the subject would alternate between the two attributes with a progressively narrower focus of attention, stopping as soon as a mismatch was found, or when a positive outcome was reached at Stage 3.

In order to test this modified model more precisely, we would need to estimate a large number of parameters, since we do not know a priori the relations between the latencies with these different degrees of focused vs. divided attention. Our main interest was in the function and order of the processing stages required in comparing conjunctions rather than in the detailed nature of the processing of attributes within each stage (e.g., whether processing of shape and color is serial or parallel, with or without interference, either for targets or for display stimuli). The simplest way of testing the model might therefore be to determine the differences in the times taken by each of the hypothesized stages when performed by itself, and to use these to predict differences between latencies in the different conditions of the conjunction task. In Experiment 3, we therefore measured latencies to compare one or two targets with one or two display items when only one attribute was relevant at a time. The condition with two targets and two display items (2t2d) would simulate Stage 1 of the model; the condition with two targets and one display item (2t1d) would simulate Stage 2; and the condition with one target and one display item (1t1d) would simulate Stage 3. We could then use differences between the latencies in Experiment 3 with different numbers of comparisons, and with color or shape relevant, to reduce to four the number of free parameters to be estimated in testing the conjunction-matching model. We give the details below.

Although the main aim of the experiment was to throw light on the processing of conjunctions, the results should also be of interest in relation to the question of how the separate attributes of shape and color are processed. This is a task in which our hypothesis would predict parallel processing of both targets and display stimuli at least for the colors, which are clearly unidimensional stimuli. For the shapes, the prediction is less certain, since shapes may not always be distinguished by single features. Recent theories have suggested that most shapes can be decomposed into sets of such more elementary features as vertical line, angle, curve, etc. In our previous experiment (Treisman et al., 1977) we found that when a single shape feature (curvature) could mediate discrimination, parallel processing was possible for shapes as well as for colors. However, for shapes which differ only in the arrangement

of shared features, the evidence suggests that focal attention is necessary (Beck & Ambler, 1973). The present experiment was also used to throw light on the question of whether shape and color are processed in parallel with each other, using our particular stimuli. Previous research suggests that they can be (Hawkins, 1969; Saraga & Shallice, 1973).

EXPERIMENT 3

Method

Stimuli. The stimuli were again the letters "O," "H," "X," and "N" in pink, blue, green, or orange ink. The targets were always "O" and/or "H," or pink and/or blue. The size and spatial arrangements of the stimuli were the same as in the previous experiments, as was the apparatus. We tested the four possible combinations of one or two targets with one or two test stimuli. 2t and 1t refer to the number of targets in memory; 2d and 1d to the number of stimuli in the display. When two targets and two display stimuli were presented, the targets were shown above and below the fixation point and the display stimuli to the left and right of it. When only one item was presented, its position was equally often top or bottom for the targets, left or right for the display stimuli. When there were two display stimuli, they always differed from each other on both attributes. On positive (2t2d) trials, either both display stimuli matched targets on the relevant attribute or only one of the two did. The number of matches between targets and display stimuli on the irrelevant attributes was always zero or one, and these conditions were analyzed separately for all combinations except those with two display stimuli and two targets.

Subjects. The subjects were drawn from the same pool as before; four had previously taken part in either Experiment 1 or Experiment 2, and two were new to this type of research. There proved to be no systematic differences in performance due to the presence or absence of this prior experience.

Procedure. Each trial consisted of the following sequence of events: The subject fixated the background card; then the experimenter said "Ready" and started the target presentation (the same throughout a block of trials), which lasted for 1 sec. This was immediately followed by a display card, also presented for 1 sec. The subject responded by pressing the right-hand key for a positive and the left-hand key for a negative test card. The RT was recorded, the display card changed, and the next trial was begun. The subjects were tested in four sessions, with the order of conditions counterbalanced. Two consecutive sessions tested color matching and two tested shape matching. Within each session, there were three separate blocks: one with both targets presented and one with each of the single targets presented. The order of testing two vs. one target conditions was reversed from one session to the next and also counterbalanced across subjects. Trials with one and with two display stimuli were randomly mixed within each block, as were negative and positive trials. The instructions were to press the right-hand key as quickly as possible without making errors if there was any match on the relevant attribute between test and target stimuli, and to press the left-hand key if there was no match. Six subjects were tested.

Results

The results are given in Tables 1 and 2 and in Figure 4. The irrelevant attribute was said to be "consistent" with the relevant attribute if, on positive trials, it also matched the target and, on negative trials, it also differed; otherwise, it was said to

Table 1
RTs as a Function of Number of Targets (t) and
Number of Display Stimuli (d)

	Shape				Color			
	Positives		Negatives		Positives		Negatives	
	2t	1t	2t	1t	2t	1t	2t	1t
2d	429 (367)	375	456	418	394 (344)	368	410	394
1d	380	339	433	391	353	335	405	386

Note—Trials in parentheses are those in which both stimuli were positive.

Table 2
Differences in Mean RT Between Consistent and
Competing Irrelevant Attribute

	Shape			Color		
	2t1d	1t2d	1t1d	2t1d	1t2d	1t1d
Positive	24	40	21	15	22	19
Negative	15	20	24	27	11	23

Note—Not analyzed for 2t2d trials.

"compete" with the correct decision. We discuss first the data in Table 1, which show the effects of number of targets and display stimuli in relation to the relevant attribute only, averaging across consistent and competing irrelevant attribute values.

Analysis of variance on the positive shape-matching trials showed significant effects of both the number of targets [F(1,5) = 11.7, p < .025] and the number of display stimuli [F(1,5) = 19.4, p < .01], but no interaction. The positive color-matching trials showed a significant effect of the number of display stimuli [F(1,5) = 38.2, p < .01] but not of the number of targets, and no interaction. In these analyses, only trials with a single matching stimulus were used for the 2t2d conditions, since with 1t2d trials only one of the two display stimuli *could* be positive. Analyses of variance on the negative shape-matching trials showed significant effects of the number of targets [F(1,5) = 18.3, p < .01] and of the number of display stimuli [F(1,5) = 8.2, p < .05], but no significant interaction. For color negatives, no effects were significant. We then analyzed separately the 1t shape trials on which "H" and "O" were the targets. There was a significant difference between "H" and "O" trials and a significant interaction between "H" vs. "O" and the number of display stimuli, their effect being much greater for "H" than for "O." In fact, for the "O" negatives, the difference between one and two display stimuli was only 7 msec, which is about the same as the 9 msec found with the color negatives; with the "H" negatives, it was 39 msec.

We look next at the effects of the irrelevant attribute. Analysis of variance of 1t trials showed significant effects of "consistent" vs. "competing" irrelevant attribute [F(1,5) = 95.91, p < .001], of positive vs. negative response [F(1,5) = 13.11, p < .025], of the number of display stimuli [F(1,5) = 32.81, p < .01], and of the interaction between the number of display stimuli and positive vs. negative responses [F(1,5, = 10.05, p < .025]: the difference between 1d and 2d was again larger for positives. No other interaction reached significance.
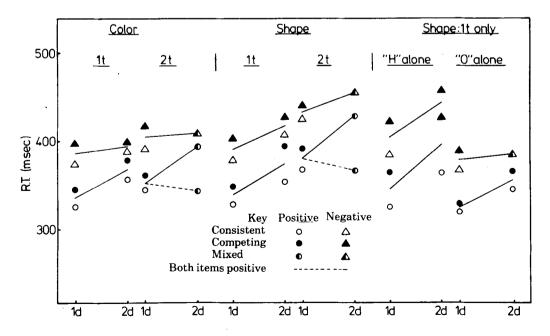
Looking at the effects of display stimulus posi-

Figure 4. Mean RTs in Experiment 3.

tion, we found no evidence for consistent left-to-right serial processing on the 2d positive trials. The RTs were, in fact, slightly faster when the matching stimulus was on the right, but this was probably an effect of S-R compatibility, since positive responses were always made with the right hand. The advantage of right over left was as great when only one stimulus was presented, and was reversed for negative 1d trials.

DISCUSSION

The aims of this experiment were (1) to obtain empirical estimates of the times required to compare single values or pairs of values on each attribute, so that these could be used to test our model for the more complex conjunction-matching task; and (2) to discover whether the colored shapes could be processed in parallel, either as memory targets or as display stimuli, and with respect both to spatially separate stimuli and to different attributes of the same stimulus. We will consider first the subsidiary question of whether separate targets and display stimuli were processed serially or in parallel. Here the negative trials given the clearest evidence, since there could be no conflict between decisions about the two stimuli on the relevant attribute. When two targets or display stimuli were presented, both would be negative, whereas on positive trials with one negative and one positive item, a mismatch with the negative stimulus might interfere with the positive decision (cf. Williams, 1974).

The negative RTs for colors suggest that two targets and two display stimuli can all be processed in parallel: latencies were affected neither by the number of targets nor by the number of display stimuli. The negative RTs for shapes, on the other hand, showed significant effects both of the number of display stimuli and of the number of targets. The more detailed analysis of 1t trials allows further clarification: when the target was "O," the two negative display stimuli ("N" and "X") were rejected in parallel. However, when the target was "H" there was a difference in the times to reject one and two display stimuli. This confirms our previous conclusion that shapes which can be discriminated by a single feature can be processed in parallel, while those which share features cannot. But it leaves the effect of the number of targets and the absence of any interaction between number of targets and number of display stimuli to be explained. If both targets and display stimuli were processed serially there should have been an interaction in the latencies; thus at least one of the two must have been parallel at the comparison stage. One possibility is that the effect of number of targets

occurred at the retrieval stage, the target shapes being retrieved serially, then compared in parallel to each successive display stimulus. Another possibility is that subjects handled *either* the display stimuli *or* the targets serially on any one trial, but not both. For example, suppose the "H" to "N" comparison could not be made in parallel with any other, but all the other required comparisons (i.e., "O" to "X," "O" to "N," and "H" to "X") could occur in parallel, the number of comparisons to be made would average one on negative 1t1d trials, 1½ on 1t2d and on 2t1d negative trials, and 2 on 2t2d negative trials, which would predict our results quite well. Thus the results on negative trials appear to be consistent with our expectation about the conditions under which parallel processing should be possible.

However, the positive trials appear to differ: with both color and shape, the effect of adding a negative target or display stimulus on a positive trial was greater than the effect of adding one on a negative trial. This suggests that some additional factor contributes to the 2t or 2d latencies on positive trials when only one matching value is present. One possibility is that the opposite outcomes of the two comparisons (one match and one mismatch) produces conflict either at the decision or the response stage. Those 2t2d trials on which both comparisons were positive were slightly faster than 2t1d trials; this could be due either to facilitation or to self-terminating parallel processing.

We look now for evidence on the relations between the attributes of shape and color. Since both attributes affected latencies even though only one attribute was relevant, we can infer some processing of color when shape was relevant and vice versa. Either the two were processed in parallel or subjects sometimes started with or switched to the wrong attribute. Conflict can apparently arise from a match or mismatch on an irrelevant as well as a relevant attribute (as Williams, 1974, also found). In the present experiment, it was in fact almost as great as that produced by a second competing value on the relevant attribute (i.e., 2t vs. 1t and 2d vs. 1d trials).

What can we conclude? If the arguments are acceptable, the most likely synthesis is that, with these particular shapes and colors, processing of two colors is parallel in both memory and display; processing of two shapes is also parallel when there is no overlap in features. If the shapes share features, however, serial processing is necessary. Decisions are slowed both by the presence of a conflicting mismatch on the relevant dimension of a second stimulus, and, almost as much, by the presence of a conflicting match or mismatch on the irrelevant attribute of the matching stimulus.

We now attempt to use these data to test the conjunction-matching model, which proposed three

stages of processing: Stage 1 consisted of a search with divided attention for any matching value between the two targets and the two test stimuli, this search being parallel wherever possible. Decisions at this stage were based on color on half the trials (either within each subject or between subjects) and shape on the other half. (This assumption was an arbitrary one; it affects only the estimates of the absolute times taken for shape and for color comparisons, not the sequence of operations hypothesized. In fact, as discussed below, the present results seem consistent with approximately equal frequencies.) The search then progressively narrowed down, first at Stage 2, specifying a single display location, and finally at Stage 3, specifying one particular target. In order to test this modified model quantitatively against the data, we attempted to simulate the three stages by matching them to the different conditions of Experiment 3, with one or two display stimuli and one or two targets presented. The comparisons required at Stage 1 of the conjunction model would correspond to the comparisons made in the 2t2d condition of Experiment 3. Similarly the comparisons made in Stage 2 testing of two targets against one display stimulus correspond to the 2t1d condition with two targets and one display stimulus. Finally, Stage 3 of the conjunction model, where testing is focused both on a single display stimulus and on a single target, would correspond to the 1t1d condition with only one target and one display stimulus presented. These assumptions allow us to bypass the problem of deciding the exact roles of competition between positive and negative decisions and of parallel vs. serial processing in determining the differences in latencies. All we have to assume is that these effects are the same in both experiments. One further important assumption, which may not be justified, is that focusing is perfectly efficient, that a single comparison of one display stimulus to a target when two are presented takes the same time as a comparison of a single display stimulus presented alone. By making this assumption, we drastically reduce the number of parameters we would otherwise need to estimate. If it results in a good fit between data and predictions, this will support both the model and the assumption, although if it fails the implications will be ambiguous.

Adopting this hypothesis, then, we estimated just four parameters from the data in the conjunction-matching task: K_P and K_N (those components of the total latency for positive and negative trials, respectively, which were not affected by differences between the conditions we tested), and C_P and C_N (the times needed to make a positive and a negative decision, respectively, about a single color comparison between one target and one display stimulus). C_P and C_N were assumed to be the central compari-

Table 3
Estimated Comparison Times (msec) at Each Stage of Conjunction Model*

Conjunction Model: Equivalent Condition of Experiment 3		Stage 1 2t2d	Stage 2 2t1d	Stage 3 1t1d
Color	Both Positive	†C_P + 9 or 23		
	One Positive	C_P + 59	C_P + 18	C_P
	Negative	C_N + 24	C_N + 19	C_N
Shape	Both Positive	†C_P + 32 or 68		
	One Positive	C_P + 94	C_P + 45	C_P + 4
	Negative	C_N + 70	C_N + 70	C_N + 5

*Taken from the differences in latencies obtained in those conditions of Experiment 3 which we assumed to involve equivalent comparisons.
†Unfortunately, Experiment 3 did not include a condition appropriate to estimate these two comparison times, since no condition required a response contingent on the presence of two matching shapes or colors. We tried two alternative hypotheses about the appropriate values to use for these two stages: (1) We assumed that the differences between the double positives ($C_{P2,2}$ and $S_{P2,2}$) and the single positives ($C_{P1,1}$ and $S_{P1,1}$) and the single negatives ($C_{N1,1}$ and $S_{N1,1}$), respectively, since in neither case was there any conflict between the two decisions in the two stimulus conditions; (2) we used the actual 2t2d double positive times from Experiment 3 to provide estimates of the differences, assuming that these times reflected parallel exhaustive processing. The latter assumption gave a slightly better fit to the data and was used in the test reported below.

son times in the 1t1d color trials of Experiment 3. The remaining parameters were all anchored to C_P or C_N and were taken directly from the differences between the appropriate conditions in the single attribute-matching experiment. All the comparisons at Stages 1, 2, and 3 of the conjunction model were matched to the corresponding conditions of Experiment 3, by equating the attribute in question, the number of targets and of display stimuli being compared, and whether the outcome was positive or negative. The appropriate differences between the other conditions of Experiment 3 and the 1t1d color conditions were added to C_P or C_N to give estimates of all the other RT components in the conjunction-matching task. The details are spelled out in Table 3.

A successive iteration program was used to estimate the four parameters in this modified model from the mean RTs in the fixed target conjunction experiment (Treisman et al., 1977). The data from this particular version of the conjunction-matching paradigm were used to test the model because they were based on the same fixed targets, pink "O" and blue "N," as were used in Experiment 3. The fit of the model to the data (shown in Figure 1) was very good: a regression of observed on predicted means gave a correlation of 0.993, accounting for 98.6% of the variance due to trial types. The estimates for C_N and C_P were 49 and 50 msec, respectively, so that S_N and S_P would both be 54 msec; for K_N and K_P,

the estimates were 355 and 279 msec, respectively.

A further prediction can be made to check the consistency of our hypothesis about the relation between the two experiments. If we are correct, it should be the case that $K_N + C_N$ gives the 1t1d negative color latency of Experiment 3, and $K_P + C_P$ should give the 1t1d positive color latency. These parameters were estimated independently from the conjunction-matching data and, in fact, fit the mean data from the single attribute-matching experiment quite well, the mean difference being only 12 msec, despite the fact that different subjects were involved.

Two further points should be mentioned. First, the model is fit to the data of the group as a whole. It involves two alternative routes to reach a response, one starting with color comparisons and one with shape. Individual subjects may divide their trials equally between these two strategies, or concentrate on one, or select any mixture of the two. Secondly the assumption that for the group as a whole the strategies were equally mixed was initially an arbitrary one. The slope of latencies against number of matching color or shapes could reflect many combinations of particular comparison times for that attribute and particular proportions of trials starting with that attribute. However, the fact that the comparison times from Experiment 3 fitted neatly into the hypothesized stages of the conjunction model suggests that the assumption of an equal mixture of strategies was approximately correct for the group as a whole.

A final question should be asked: How does this modified model fit the data with schematic faces? Here a difficulty does arise, since there was no evidence for parallel processing at any stage with the faces. Why then should a final recheck with focused attention be required at Stage 3 if processing of the features has been serial throughout? One possibility is that our original hypothesis was preferable, i.e., that the "conjunction check" at Stage 3 is quite a different operation from the feature checks at Stages 1 and 2. However, another account, more consistent with the revised model for colored shapes, may be possible. Serial processing, we have argued, *allows* conjunctions to be identified. However, it does not ensure that they will be. It is logically quite possible to envisage a serial check of one feature in a number of display locations or in a number of memorized targets without the other feature or features in these locations or targets being processed at all. Conjunctions would not then be registered despite the serial processing of one feature. There may be some cost involved in switching between features when one must specify the particular location or target whose other features are required. Thus the difficulty may be in the conjoint retrieval of two or more features relating to the same specified stimulus or target. The

Figure 5. Model previously proposed for same-different matching of schematic faces.

requirement to focus attention should be redefined then not as the requirement to switch from parallel to serial procesing for any one attribute, but as the requirement to locate or retrieve the several features of each particular memorized or physically present stimulus in turn. Stage 3 would then be necessary with schematic faces, whose features are serially processed, as well as with colored shapes. Thus, if it is in fact easier to process each feature separately than to focus on a particular stored conjunction, the progressive focusing model can also apply to the case where individual features are processed serially instead of in parallel or overlapping fashion.

Figure 5 shows the model we had previously proposed for the schematic faces. The only modifications now needed are to specify (1) that the first two feature checks for any one display stimulus each involve serial decisions about both targets; and (2) that the final conjunction check consists of a single target to display comparison on one feature. Thus the previous parameters E_N and M_N (the times to reach a negative decision about the eyes or mouth of one display face) actually consist of $2e_N$ and $2m_N$ (where the lowercase letters represent the time for each target to display comparison); the previous parameters E_P and M_P (the times to reach a positive decision about the eyes or mouth of one display face) actually consist of $e_N + e_P$ and $m_N + m_P$; and the previous parameters t_N and t_P (the times to reach a negative or positive decision about the conjunction) now consist of e_N or m_N, and e_P or m_P, respectively. We used the successive iteration program to estimate

best fitting times for these parameters and found that the fit was again very good, accounting for 98.8% of the variance due to trial types. The estimates for the parameters were as follows: $e_N = 47$, $m_N = 38$, $e_P = 65$, $m_P = 75$, $K_N = 486$, $K_P = 426$.

It is of interest that the estimated comparison times are faster here for negatives than for positives, while the reverse is true for the constant components, K_N and K_P. The normal finding for highly discriminable stimuli is that positives are faster than negatives. However, most experiments do not allow one to separate the central comparison times from the response times. The only paradigm we could think of which allows this separation is a visual search task in which one compares the slope of latency against number of items searched for in two different conditions—search for presence and search for absence of a target (Neisser, 1963). In the search for presence of a target, nontargets should require negative decisions at the comparison stage, while in the search for absence of a target, the nontargets would all require positive decisions. The comparison times are estimated from the slopes of the functions and therefore do not include the final response time. It is interesting that in this paradigm, too, the positive decisions apparently take longer than the negative ones. Thus the reversal our model produces for the features of eyes and mouth (though not for the colored shapes) is perhaps less implausible than it initially appears.

What conclusions can we draw from this series of experiments? If the model we propose is correct, it suggests that focusing attention requires effort. When speed of response is emphasized, subjects narrow down their analysis only to the extent that is needed to specify how potentially relevant features are conjoined. This narrowing appears to occur in two discrete stages, first for stimuli physically present in the display, and last, and only if necessary, for stimuli stored in memory. The asymmetry between physically present stimuli and memorized targets suggests that the cost of focusing is less when the focus is a physical spatial location than when it is an item in memory. Experiment 2 supports the notion that the code which integrates conjunctions in memory differs from the physical code of spatial location normally used to identify conjunctions during perception. This difference in the nature of the "glue" which unifies targets and display stimuli may explain the difference in the ease and efficiency of focusing attention. Display stimuli appear to be spatially addressable, while targets must be retrieved by some other means, perhaps by their temporal labels or by their content.

REFERENCES

BECK, J., & AMBLER, B. The effects of concentrated and distributed attention on peripheral acuity. *Perception & Psychophysics*, 1973, **14**, 225-230.

GARNER, W. R. *The processing of information and structure.* New York: Wiley, 1974.

HAWKINS, H. L. Parallel processing in complex visual discrimination. *Perception & Psychophysics*, 1969, **5**, 56-64.

LABERGE, D. Attention and the measurement of perceptual learning. *Memory & Cognition*, 1973, **1**, 268-276.

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

SARAGA, E., & SHALLICE, T. Parallel processing of the attributes of single stimuli. *Perception & Psychophysics*, 1973, **13**, 261-270.

TREISMAN, A. M., SYKES, M., & GELADE, G. Selective attention and stimulus integration. In S. Dornic (Ed.), *Attention and performance VI.* Potomac, Md: Erlbaum, 1977, in press.

WILLIAMS, C. The effect of an irrelevant dimension on 'same-different' judgments of multi-dimensional stimuli. *Quarterly Journal of Experimental Psychology*, 1974, **26**, 26-31.