Integration of features in comparing multifeature stimuli

ROBERTA L. KLATZKY and ANDREA THOMPSON
University of California, Santa Barbara, California 93106

For a series of trials, the subject indicated whether or not two face-like stimuli were identical with respect to the features they contained (a positive or negative response, respectively); reaction time (RT) was recorded. The stimuli were separated by .5 or 4.0 sec, and each contained either two or three features. On some trials, the features of the first stimulus were represented within a single head outline; on others, they were distributed over several spatially separated outlines (components). Positive-response RT increased with the number of stimulus features and with the number of first-stimulus components (*c*); these two factors were additive. Negative-response RT increased with *c*, decreased with the number of differing features, and these two factors interacted. These effects did not vary with interstimulus interval. The results indicate that the subject did not integrate the first-stimulus components, but instead compared the second stimulus to each component, holistically (for positive responses) or analytically (for negative responses).

The integration, or synthesis, of whole stimuli from partial information is a process that has often been hypothesized in theories of visual information processing. For example, a synthesis process has been suggested as a component of pattern recognition (e.g., Neisser, 1967; Selfridge, 1959), memory for visual forms (e.g., Franks & Bransford, 1971; Palmer, 1975), and generation of an anticipated visual stimulus (e.g., Klatzky & Stoy, 1974; Posner, 1969; Tversky, 1969).

The concept of integrated stimulus representations has also been relevant in tasks where subjects judge whether visual stimuli are the same or different. In one type of task, the subject indicates for a series of trials whether or not two stimuli, each determined by combining selected values of various features, are identical with respect to some designated features. The subject's reaction time (RT) is recorded on each trial. A question arises in such tasks as to whether the stimuli are compared feature by feature or as integral, holistic forms. Nickerson (1972), after reviewing the literature on multifeature judgments, suggests that the most functional model is a two-process one (as suggested by Bamber, 1969; Beller, 1970; Lockhead, 1972), according to which either type of stimulus comparison may occur. Holistic processing may underlie positive (same) responses, particularly if the stimuli are familiar. Principal evidence for this assumption is that RT for positive responses does not increase regularly with the number of relevant features. In contrast, analytic processing may underlie negative (different) responses. This is

indicated by the facts that negative-response RT decreases as the number of features differing between the two compared stimuli increases, and it increases with the number of features to be compared.

The purpose of the present study was to investigate stimulus integration by presenting subjects with fragmented visual forms and instructing them to mentally integrate them into whole stimuli. The stimuli were facial forms, constructed by combining values of two or three features. These stimuli were selected because subjects know the relative positions of facial features and are familiar with faces in an integrated form. The procedure of the experimental task was as follows (see Figure 1). On each trial, the subject first viewed a multifeature facial stimulus. The stimulus was on some trials a single face outline

Figure 1. Examples of first and second stimuli which could be presented on a trial. Top row: 3 features, 3 components, positive response; middle row: 3 features, 2 components, negative response; bottom row: 2 features, 1 component, positive response.

This research was supported by Grant MH 25090 from the National Institute of Mental Health to the first author. The authors wish to thank P. J. Geiwitz and D. Messick for their helpful comments on the manuscript.

containing information about all the features (for example, an oval having a pair of eyes, a nose, and a mouth). On other trials, the features of the stimulus were divided over several spatially separated components—partial stimuli—each depicting information about some subset of the features, with the limiting case one feature in each component. Next, the subject was presented with a second stimulus, which consisted of facial features within a single face outline. The subject was required to indicate whether or not the second stimulus was identical to the first, with respect to the particular features the two stimuli contained, and RT was recorded.

The principal independent variables in this study were the number of features in the facial stimuli, the number of components—face outlines—over which the features of the first stimulus were distributed, and the interstimulus interval (ISI—the time between the offset of the first stimulus and the onset of the second). The number-of-components variable is of particular importance in assessing whether integration of the stimuli occurred. If the subject integrates the first stimulus, in the sense of combining its features into a unitary representation regardless of the number of components into which it was initially divided, there should be no effect on RT of the number of components. The number-of-features variable can be used to determine whether integration, if it occurs, is performed on a visual or verbal basis. These two variables also serve to indicate whether stimulus comparisons are analytic or holistic. Finally, the variation of RT with ISI provides information about changes in the memorial code of the first stimulus over time.

METHOD

Subjects

The subjects were 16 students at the University of California, Santa Barbara. They took part in six experimental sessions each, receiving class credit in the introductory psychology course for three sessions and paid \$2 for each remaining session. There were 11 female subjects and 5 males.

Stimuli

The stimuli were facial forms, composed of facial features arranged in oval frames. (An oval represented the outline of a head.) The number of facial features (f) in a given stimulus was either two (in which case the features were eyes and mouth; referred to as the $f = 2$ condition) or three (eyes, nose, and mouth; the $f = 3$ condition).¹ Each feature could adopt any of three possible values (that is, there were three different mouths, three different noses, etc.).

In the case of the first stimulus presented on a trial, the facial features could be presented within a single oval frame, representing a complete face, or could be distributed over several spatially separated ovals. Each oval frame and the features it contained will be referred to as a *component*. (See Figure 1 for examples of components.) Thus, in the $f = 2$ condition, each first stimulus on a trial contained two features, but the features could be represented within a single oval frame (called the $c = 1$ condition, where c represents the number of components) or could be separated, with each feature in a distinct oval ($c = 2$). Similarly, in the $f = 3$ condition, the three features could be represented within one, two, or three ovals ($c = 1, 2$, or 3 , respectively). Regardless of the

number of components, each feature was always presented in the same location with respect to the oval frame containing it; with eyes at the top, nose in the center, and mouth at the bottom (see Figure 1).

For cases with more than one component, the distribution of features over components could be accomplished in several ways. When $f = c$, each component contained a single feature. In that case, any feature could occur in any left-to-right serial position. In the case where $f = 3$, $c = 2$, there were two features in one component and the third feature in the other. All possible combinations of features in components were used, and each type of component could occur on either the left or the right.

The set of possible first-stimulus configurations was increased by the fact that each feature could take on any of three different values. For the $c = 1$ case, all possible stimuli were constructed for use during the experiment for a total of 9 faces where $f = 2$ and for a total of 27 where $f = 3$. In the $f = 2$, $c = 2$ condition, all possible configurations (separating features into components and varying the serial order of components and particular feature values, for a total of 18) were also used. However, for the $f = 3$, $c > 1$ conditions, the number of possible configurations was too great to permit construction of all possible stimuli. In these cases, therefore, a subset of possible configurations was used in the experiment, containing 81 stimuli for $c = 2$ and 27 for $c = 3$. The particular stimuli to be constructed were selected within the constraints that all possible distributions of feature types over components were represented approximately equally often; components containing a given type of feature(s) occurred equally often in each possible serial position; and all combinations of particular feature values were used equally often.

In the second stimulus presented on a trial, the features were always combined in a single component. All possible combinations of feature values were used, for a total of 27 faces when $f = 3$ and 9 faces when $f = 2$. Thus, the second stimuli were equivalent to the first stimuli used in the $c = 1$ case.

Apparatus

The apparatus was a three-field Iconix tachistoscope. For use in the apparatus, the stimulus components were mounted side by side on 23 x 15 cm white cards so that the components were centered in the subject's viewing area. Each component measured 2.1 x 3.4 cm, and the separation between components (if any) was 1.0 cm. Stimulus presentations were controlled by the experimenter, who sat opposite the subject. The subject placed his head against a head mask and looked into the exposure box. The viewing distance was approximately 76 cm for each field.

A rectangular response panel measuring 13 x 18 cm was on a table before the subject. The panel contained two response switches, one in each of the upper corners, and a 7.6-cm-long start bar centered 3.2 cm from the lower edge. The subject was instructed to depress the start bar with his thumbs, simultaneously holding his right and left index fingers on the corresponding response switches. Each subject was allowed to select which of the response switches would correspond to a positive response—14 subjects selected the right-hand switch for positives and the left for negatives; the remainder used the reverse arrangement. (Previous research with this apparatus indicates no differential effects on RT of response arrangements in similar experiments, and subsequent data analysis showed no differential effects of response assignment in these experiments.)

Procedure

The events of a single trial of the experiment were as follows: The subject first pressed a start bar, and was then presented with a stimulus containing information about two or three facial features, divided among one, two, or three components (three components only when there were three features). The subject could view the stimulus as long as he desired, but had to view it for at least 6 sec. The subject was instructed to try to imagine the stimulus as a single-component face, even if it was presented as more than one component. The subject then pressed the start bar for a second time, causing the first stimulus to vanish and a fixation point (a black dot centered in the visual field) to appear for the ISI. This was

followed by the appearance of a single-component face, containing the same number of features as the first stimulus. The subject indicated, by pressing a response key, whether or not the particular features represented in the first stimulus, regardless of their distribution over components, were identical to those in the second stimulus. The subject was informed if the RT was longer than 2 sec or if an incorrect response was made. After an interval of approximately 15 sec, during which the experimenter arranged the stimuli in the apparatus, the next trial began. Reaction time, defined as the interval between the onset of the second stimulus and the subject's response, was recorded on each trial; subjects were instructed to respond as quickly as possible without error.

Each subject participated in one training session and five subsequent test sessions of approximately 115 trials each. Each session lasted about 50 min. The variable, number of features in the stimuli ($f = 2$ vs. $f = 3$), was varied within each session but was held constant for a block of about 10 trials and alternated between blocks. No other variable was blocked. For a given value of f , positive and negative trials occurred randomly with equal probability and the value of c varied randomly with each value occurring equally often. In addition, within each value of f for negative trials, the number of differing features (d , which varied from 1 to f) varied randomly, with each number occurring equally often, and the identity of differing features was randomly determined with equal probabilities. Within these constraints, the particular stimuli used on a trial were randomly selected.

A further variable, the length of the ISI, was varied between subjects. One group of eight subjects was assigned an ISI value of .5 sec; the other group of eight subjects had an ISI of 4 sec. The assignment of subjects to groups was random.

RESULTS

The principal data are mean RTs in milliseconds. The data analysis includes correct responses only and the data are combined over test sessions. (When the data were examined by sessions, there were no apparent differences in the forms of the RT functions, although mean RT gradually declined with practice.) Since subjects were instructed to make as few errors as possible, error rates were generally low, with a range over subjects of 1.7%-10.7% and a mean of 5.9%. The data from the training sessions and from three practice trials which preceded each test session were not included in the data analysis. Aberrant RTs, defined as those which were over three standard deviations larger than the mean within their condition, were also excluded from the analysis.

When the data were examined for effects of the ISI variable, which was the sole between-subject manipulation, no ISI effects were obtained. That is, for both positive and negative responses, the two groups of subjects did not differ with respect to the effects of the independent variables. In fact, for the RT functions reported below, there were no differences between groups in the absolute values of the slopes or intercepts. For that reason, the data reported henceforth include both groups of subjects.

The RTs for positive responses in this experiment are shown in Figure 2. The figure depicts positive RT as a function of c for each value of f . One-way analyses of variance with repeated measures were performed to assess the effects of c , for each value of f . In both cases, the effects were significant [for $f = 2$, $F(1,15) = 26.2$, $p < .01$; for $f = 3$, $F(2,30) =$

Figure 2. Mean RT for positive responses, as a function of the number of components in the first stimulus (c), for each value of f . The best fitting linear functions are also shown.

20.3, $p < .01$]. In the $f = 3$ case, 96% of the variance was accounted for by a linear component [$F(1,30) = 38.9$, $p < .01$], leaving an insignificant residual variance. The effects of f were revealed by t tests, which showed no significant differences between the slopes of the functions relating RT to c for $f = 2$ vs. $f = 3$, and a significant intercept effect [$t(15) = 1.8$, $.05 < p < .10$, and $t(15) = 4.4$, $p < .01$, for slopes and intercepts, respectively].

The data from negative responses are shown in Figure 3, which depicts RT as a function of c for each value of d (number of differing features), for $f = 2$ and $f = 3$. For each subject, the mean RT over negative conditions was greater for $f = 3$ than for $f = 2$. A two-way analysis of variance with repeated measures was performed to assess the effects of c and d , within each value of f . In both cases, there were significant main effects of c and d [$F(2,30) = 41.0$, $p < .01$; $F(2,30) = 8.7$, $p < .01$; $F(1,15) = 87.6$, $p < .01$; and $F(1,15) = 31.1$, $p < .01$, for the effects of c and d within $f = 3$ and $f = 2$, respectively]. The interaction of c by d was significant for $f = 3$ [$F(4,60) = 3.1$, $p < .02$] and nearly so for $f = 2$ [$F(1,15) = 4.0$, $.05 < p < .07$]. Tests for the linearity of RT as a function of c for the $f = 3$ data were significant at each level of d [$F(1,30) = 12.5$, 10.0, and 5.3; $p < .01$, .01, and .05; and variance accounted for by linearity = 86%, 95%, and 93%, for $d = 1, 2$, and 3, respectively]. Thus, the observed c by d interaction can be attributed to decreases in the slopes of these RT functions with increasing values of d .

The data for negative responses can be further broken down into serial position curves, showing RT for the case where the first stimulus contains one feature per component ($c = f$) and there is one differing feature ($d = 1$), as a function of the

Figure 3. Mean RT for negative responses as a function of the number of components in the first stimulus (c), for each value of d , for the $f = 3$ condition (left panel) and the $f = 2$ condition (right panel). The best fitting linear functions are shown for each value of d .

left-to-right serial position of the component containing that differing feature. The left panel of Figure 4 shows these data for each f value. In addition, feature identity curves can be plotted, showing RT for $d = 1$ as a function of the specific feature that differs. The right panel of Figure 4 shows these data, combined over all values of c , for $f = 2$ and $f = 3$. The data in the left panel of Figure 4 show no effect of left-to-right serial position for either value of f . (Sign tests for decreasing trend yielded $p > .5$ and $p > .25$ for $f = 2$ and $f = 3$, respectively.) In contrast, as shown in the right panel, RT is least when eyes differ and greatest when mouth differs. (Sign tests for an increasing trend from eyes to mouth yield values of $p < .01$ for both $f = 2$ and $f = 3$.)

DISCUSSION

The processing stages required for the subject to perform correctly in this task can be generally specified as follows for each trial. First, the subject encodes the first stimulus and stores it in memory; next, the subject encodes the second stimulus when it appears and compares it to the first-stimulus code; finally, the subject responds. (The RT includes the time for all stages following the second-stimulus presentation.) The principal focus of this study is the nature of the first-stimulus code, in particular, whether the subject encodes information presented in several components into an integral whole. Also of interest is the nature of the subject's comparison process.

We shall first consider the data from positive-response conditions, shown in Figure 2. The effect of c on RT clearly precludes the possibility that the subject integrates the first-stimulus components into some unitary visual representation, despite the

fact that such integration would be functional in the task. Nor does the subject encode and compare the stimulus features verbally, for this should also effectively combine the information from various components. (The magnitude of the effect of f —the intercept difference of 60 msec in Figure 2—is also inconsistent with the verbal hypothesis. For, with one fewer label to verbalize in the $f = 2$ condition than when $f = 3$, the subject should be faster by at least the time required to verbalize a single syllable—about 160 msec, according to Landauer, 1962.) One must conclude that the code of the first stimulus underlying positive responses is a visual-spatial one that does not integrate the stimulus components. Moreover, this code must be relatively abstract, since it is not affected by the ISI as a sensory trace would be.

Although the subject does not integrate the separated first-stimulus components into a whole, it can be inferred from the effects of f and c that he does treat each component as an integral whole in the comparison process. First, note that f and c act independently, which suggests, according to an additive-factors analysis (Sternberg, 1969), that they affect different processing stages. Since the c factor describes spatial characteristics of the first stimulus, it seems indicated that RT increases with c because the second stimulus must be compared to each component of the first. The slope of the function relating RT to c [henceforth referred to as $RT(c)$] thus represents the time required for each comparison. The effects of f , in contrast, are assumed to occur at a stage other than comparison, presumably second-stimulus encoding.

Next, note that the linearity of $RT(c)$ indicates that

Figure 4. Left panel: Mean RT for the case where $d = 1$ and $c = f$, as a function of the left-to-right serial position of the first-stimulus component that contains the differing feature, for each value of f . Right panel: Mean RT in the case where $d = 1$, as a function of the identity of the differing feature, for each value of f (averaged over all values of c).

the effect of adding a component to the first stimulus [the slope of $RT(c)$] is independent of the number of features that component contains (since as c increases, the number of features per component decreases). The slope of $RT(c)$ also does not vary with f . Since the slope presumably represents the time for comparing a single component of the first stimulus with the second, these facts indicate that the comparison time is unaffected by the number of features in either the component or the second stimulus. And this suggests that the second stimulus and component are compared holistically, rather than feature by feature.

In short, it appears that when performing positive responses, the subject encodes the second stimulus and compares it in gestalt fashion to each component of the first stimulus in memory. This process is similar to that proposed by Sternberg (1966, 1969) for high-speed scanning of short-term memory.

In the data of negative responses shown in Figure 3, the effect of c on RT again rules out the hypothesis that the subject integrates the components of the first stimulus, either visually or verbally, when he stores it in memory. Instead, as in the positive-response case, the first stimulus seems to be held as a visual-spatial code which maintains the components separately, and the second stimulus must be compared with each component. However, another aspect of the negative data contrasts with the positive case. Specifically, the decrease in RT with increases in d suggests that negative comparisons involve a self-terminating, feature-analytic process, as noted by other researchers (Nickerson, 1972). That is, it appears that given the second stimulus of the trial, the subject encodes each feature, compares it against each component of the first stimulus, and terminates processing when he finds a difference. The same feature-analytic processing cannot be proposed to underlie positive responses, however, for it predicts that the comparison time, and thus the slope of $RT(c)$, should vary with f , which is incorrect.

The data of Figure 4 support the idea that a self-terminating, feature-comparison process underlies negative responses. The right panel of the figure shows that when one feature differs between the two stimuli, RT is affected by the identity of that feature. This result suggests that the stimulus features are processed in order, from the topmost (eyes) to the bottommost (mouth) and that processing terminates when a difference is obtained. (This result was also obtained by Smith and Nielsen, 1970.) In contrast, the data in the left panel show no effect on RT of the left-to-right position of the component that contains the differing feature. This suggests that processing of the first-stimulus components does not proceed in a fixed serial order.

In summary, for both positive and negative responses in this study, no support was obtained for the hypothesis that subjects integrate fragmented stimuli into wholes. A natural question that arises is under what conditions integration might be induced. Possibly, the present use of stimuli whose components can themselves be viewed as wholes precluded an integrating strategy; the definition of situations in which such a strategy occurs remains a focus for future research.

REFERENCES

- BAMBER, D. Reaction times and error rates for "same"-different judgments of multidimensional stimuli. *Perception & Psychophysics*, 1969, **6**, 169-174.
- BELLER, H. K. Parallel and serial stages in matching. *Journal of Experimental Psychology*, 1970, **84**, 213-219.
- FRANKS, J. J., & BRANSFORD, J. D. Abstraction of visual patterns. *Journal of Experimental Psychology*, 1971, **90**, 65-74.
- KLATZKY, R. L., & STOL, A. M. Using visual codes for comparisons of pictures. *Memory & Cognition*, 1974, **2**, 727-736.
- LANDAUER, T. K. Rate of implicit speech. *Perceptual and Motor Skills*, 1962, **15**, 646.
- LOCKHEAD, G. R. Processing dimensional stimuli: A note. *Psychological Review*, 1972, **79**, 410-419.
- NEISSER, U. *Cognitive psychology*. New York: Appleton-Century-Crofts, 1967.
- NICKERSON, R. S. Binary-classification reaction time: A review of some studies of human information-processing capabilities. *Psychonomic Monograph Supplements*, 1972, **4**, 275-318.
- PALMER, S. E. Visual perception and world knowledge: Notes on a model of sensory-cognitive interaction. In D. A. Norman & D. E. Rumelhart, *Explorations in cognition*. San Francisco: Freeman, 1975.
- POSNER, M. I. Abstraction and the process of recognition. In J. T. Spence & G. H. Bower (Eds.), *Advances in learning and motivation* (Vol. III). New York: Academic Press, 1969.
- SELFRIDGE, O. G. Pandemonium: A paradigm for learning. In *The mechanization of thought processes*. London: H. M. Stationary Office, 1959.
- SMITH, E. E., & NIELSEN, G. D. Representations and retrieval processes in short-term memory: Recognition and recall of faces. *Journal of Experimental Psychology*, 1970, **85**, 397-405.
- STERNBERG, S. High-speed scanning in human memory. *Science*, 1966, **153**, 652-654.
- STERNBERG, S. The discovery of processing stages: Extensions of Donders' method. *Acta Psychologica*, 1969, **30**, 276-315.
- TVERSKY, B. Pictorial and verbal encoding in a short-term memory task. *Perception & Psychophysics*, 1969, **6**, 225-233.

NOTE

1. The use of eyes and mouth in the two-feature case was to maximize the face-like quality of the stimulus, although it produces a confounding between the number of features and the particular features in the stimulus. However, previous research (Smith & Nielsen, 1970) indicates that this confounding is not critical to the results. Moreover, the present data (Figure 4) indicate that the nose feature was not processed most slowly, and thus that effects of such factors as differential familiarity of features were minimal.

(Received for publication April 14, 1975;
revision received August 25, 1975.)