SAME AND DIFFERENT JUDGMENTS FOR WORD-COLOR PAIRS
WITH "IRRELEVANT" WORDS OR COLORS:
EVIDENCE FOR WORD-CODE COMPARISONS

FREDERICK N. DYER

U.S. Army Medical Research Laboratory, Fort Knox, Kentucky

Latencies for same and different judgments resulting from comparisons between words and colors were determined with one of the pair of stimuli being both a word and a color. The irrelevant aspect of this dual stimulus bore each of five possible relationships to the stimuli that were relevant to the match. These five conditions plus two control conditions produced large differences in latencies for making comparisons. Same responses were generally faster than different responses. Correspondence between the irrelevant stimulus and the combined relevant stimulus facilitated same responses but showed no facilitation of different responses. These findings indicate that basic differences exist between same and different decision processes. In some conditions, irrelevant words delayed matches between words and colors more than irrelevant colors delayed such matches. This suggests that central comparisons between the pair of stimuli were in a form more closely related to words than to colors.

Treisman and Fearnley (1969) showed that judgments of the equivalence and non-equivalence of two stimuli took more time when each stimulus was characterized by a different stimulus attribute than when the two stimuli were characterized by the same attribute. Colors and color names were compared in a card-sorting task. An example of one of the cards for "cross-attribute" matching was the word blue printed in black above a series of green Xs. Since the word and color do not correspond, this particular card would be sorted into the different pile. An example of "within-attribute" matching for the attribute of color was a color name in blue printed over a series of green Xs. An example of within-attribute matching for color names was the word red printed twice, one above the other, with the upper word in colored ink. The last two examples would be sorted different and same, respectively. A condition with simpler stimuli for within-attribute matching (both stimuli Xs for color matches and both words in black for word matches) was not included in the Treisman and Fearnley study; but the fast times for within-attribute matching with their complex stimuli relative to cross-attribute matching with the simple stimuli described above indicated that little or no interference to within-attribute matching resulted from the competing value on the other attribute.

In addition to the shorter time for within-attribute matching relative to cross-attribute matching, Treisman and Fearnley (1969) found that cross-attribute matching was considerably delayed when the upper stimulus involved in the comparison was both a word and a color, i.e., a color name printed in colored ink. The additional delay, produced by the irrelevant aspect of these combination stimuli, occurred about equally in the condition where a color was compared to a word (CW matching) and the condition where a word was compared to a color (WC matching). The difference between the stimulus cards for CW and WC matching was in the choice of a black word or a series of colored Xs, respectively, for the lower noncombination stimulus. The irrelevant aspect of the upper combination stimulus was thus a word for CW matching and a color for WC matching.

Since anomalous situations can provide

\[1\] The author wishes to acknowledge the assistance of Laurence J. Severance who helped develop procedures and collected the data.

\[2\] Requests for reprints should be sent to Frederick N. Dyer, Experimental Psychology Division, U.S. Army Medical Research Laboratory, Fort Knox, Kentucky 40121.
unique information about the normal processes that are disrupted (e.g., optical illusions), the present study was designed to further explore the delaying effect of these combination color and word stimuli on cross-attribute matching. Such combination stimuli constitute the basic elements of the Stroop (1935) test, where the time for color naming with a series of colored color names (each of which is printed in a color that does not correspond to the color name) is delayed in comparison to the time for color naming with a series of rectangular color patches. The Stroop test itself provides a means for studying the interaction of central processes associated with color and word perception (Dyer, in press). The Treisman and Fearnley (1969) procedure also lends itself to such investigations and, in fact, has an advantage in that only the two responses—same and different—are required regardless of the number of levels of the perceptual dimension.

In addition, the Treisman and Fearnley (1969) matching paradigm is excellently suited for the study of the processing of the same and different judgments themselves. Their results indicated that these judgments were disrupted for cross-attribute matching with irrelevant stimuli present. If different processes are involved in same and different judgments as has been suggested (Bindra, Donderi, & Nishisato, 1968), these judgments may differentially reflect effects of redundant and/or conflicting relationships between the irrelevant and relevant stimuli.

In the present study, the times for individual verbal same and different responses were analyzed instead of the total card-sorting times of Treisman and Fearnley (1969). This allowed separate comparisons of matching latencies for the following seven categories of relationship among the two relevant stimuli and the irrelevant stimulus:

Relevant values matching (same judgments)

**Control–Same**—no irrelevant stimulus, upper stimulus is a word (for WC matches) or color (for CW matches) but not both;

**Congruent–Same**—upper irrelevant stimulus, color or word, corresponds to the relevant stimuli;

**Incongruent–Same**—upper irrelevant stimulus is different from the relevant stimuli.

Relevant values nonmatching (different judgments)

**Control–Different**—no irrelevant stimulus, upper stimulus is either a word or color but not both;

**Congruent–Different**—upper irrelevant stimulus differs from lower stimulus but agrees with the relevant stimulus it is combined with;

**Incongruent–Different A**—upper irrelevant stimulus differs from both relevant stimuli;

**Incongruent–Different B**—upper irrelevant stimulus agrees with lower relevant stimulus (which implies a disagreement with the combined upper relevant stimulus).

Figure 1 presents a pair of stimulus configurations for each of these seven categories. The first example of each pair is for CW matching and the other is for WC matching. It is suggested that the outline letters in Figure 1 be colored to clarify these easily confusable relationships. The lowercase letters r, b, g, and w stand for the colors red, blue, green, and white. White is neutral for these stimuli as black was for the stimuli of Treisman and Fearnley (1969).

The various redundant and conflicting relationships between the irrelevant stimulus and the relevant stimuli suggest some predictions of differences in time for same and different judgments of the seven categories. For the Congruent–Same, Congruent–Different, and Incongruent–Different A categories, the relationship of the irrelevant aspect of the upper combination stimulus to the lower stimulus is the same as the critical relationship of the relevant aspect of the upper stimulus to the lower
different possible

stimulus. This redundancy might be expected to permit fast judgments of same (Congruent–Same) or different (Congruent–Different and Incongruent–Different A) for these categories. On the other hand, this relationship conflicts for the Incongruent–Same and Incongruent–Different B categories, and classification of these stimuli might be expected to take longer.

**METHOD**

*Subjects.* Sixteen males ranging in age 20–24 yr. who had recently completed Army basic training served as Ss in the experiment. All had adequate visual acuity, with correction if necessary, and color vision within normal limits as measured by the Bausch and Lomb Orthorater.

*Apparatus.* Twelve 35-mm. slides were constructed for each of the seven categories. One half were colors combined with irrelevant words above white words, and the other half were words combined with irrelevant colors above colored Xs for CW and WC matches, respectively (see Figure 1). The stimuli represented all possible combinations of the three colors (red, blue, and green) and the three words (red, blue, and green). For the Incongruent–Same and all of the Different categories, the six slides for a category and match direction were all different. For the Control–Same and Congruent–Same categories, only three different slides are possible for each match direction. To make equal numbers of stimuli in all categories, each of these slides was used twice.

Slides were projected at a distance of 150 cm. The background of the words was dark. The height of the projected capital letters subtended a visual angle of about 2°, and the separation between the two words (or the word and the string of Xs) was equal to the letter height.

Verbal response latencies were measured by a voice-operated switch which stopped a Hunter Klockounter that had started with presentation of the stimulus.

*Procedure.* The Ss were tested at two sessions with about a week between them. Within each session, two random sequences of 42 stimuli were presented. Stimuli for WC matching were in one sequence and stimuli for CW matching in the other. Six stimuli from each of the seven categories were included in a sequence. Half of the Ss began with WC matching and half with CW matching, and at the second session this order was reversed.

Every random sequence of 42 stimuli in the experiment was different. Prior to each sequence, examples of the stimuli were presented and instructions for the required type of matching were given. At the first session, practice trials were given prior to each sequence. The interval between stimuli in a sequence was 10 sec. The S rested for 2 min. between sequences. Instructions stressed both speed and accuracy. Errors were noted and such trials were repeated in a short sequence of additional stimuli at the end of each regular sequence.

*RESULTS*

Response latencies for seven general categories of relevant and irrelevant stimulus combinations were analyzed as one factor (category) in a repeated-measures analysis of variance that also included a two-level factor of match directions, another two-level factor of sessions, and a six-level factor (combination) that represented the different particular combinations of the three colors (red, blue, and green) and the three words (red, blue, and green) used as attribute values. The standard deviation of the responses latencies, calculated from the nearly equal error terms for three- and four-way interactions, was 204 msec.

Highly significant differences appeared for the seven categories, $F (6, 90) = 39.4, \ p < .001$. The mean response latency for the Control–Same category was 914 msec. The Congruent–Same category, in which the irrelevant and relevant stimuli coincided, produced an average response latency of 859 msec. A Newman-Keuls test of

<table>
<thead>
<tr>
<th>CATEGORY</th>
<th>Color-to-Word Matching</th>
<th>Word-to-Color Matching</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control-Same</td>
<td>XXXXg</td>
<td>XXXXw</td>
</tr>
<tr>
<td>Congruent-Same</td>
<td>XXXXg</td>
<td>XXXXr</td>
</tr>
<tr>
<td>Incongruent-Same</td>
<td>XXXXg</td>
<td>XXXXb</td>
</tr>
<tr>
<td>Control-Different</td>
<td>XXXXb</td>
<td>XXXXw</td>
</tr>
<tr>
<td>Congruent-Different</td>
<td>XXXXb</td>
<td>XXXXr</td>
</tr>
<tr>
<td>Incongruent-Different A</td>
<td>XXXXb</td>
<td>XXXXw</td>
</tr>
<tr>
<td>Incongruent-Different B</td>
<td>XXXXb</td>
<td>XXXXg</td>
</tr>
</tbody>
</table>

Fig. 1. Examples of stimulus pairs for different categories of relationship between irrelevant and relevant aspects of the stimuli. (Lowercase subscripts denote color—red, blue, green, or white.)
ordered means indicated this to be significantly faster \((p < .01)\) than the Control–Same category. The Incongruent–Same category, where the irrelevant stimulus conflicted with the matching relevant stimuli, produced an average response latency of 1,093 msec. This response latency was significantly longer \((p < .01)\) than that of the six other categories.

Among the four categories requiring different responses, the Control–Different category with no irrelevant stimulus produced an average response latency of 971 msec. This was significantly faster than the other Different categories and significantly slower than the Control–Same and Congruent–Same categories \((p < .01\) for all comparisons). The Congruent–Different, Incongruent–Different A, and Incongruent–Different B categories did not differ from each other at the .05 level of significance. Means for these conditions were 1,047 msec., 1,034 msec., and 1,025 msec., respectively. Categories producing long response latencies also produced high error rates (see Figure 2). The rank-order correlation between response latency and error rate for the seven categories was .81.

The match direction main effect did not approach significance and the only interaction with this variable that approached significance was that involving category, \(F(6, 90) = 1.83, p < .10\). This interaction was significant for a separate analysis of the four categories requiring different responses, \(F(3, 45) = 3.06, p < .05\). Average response times for the two match directions are plotted for all of the categories in Figure 2. The only appreciable difference for the two directions of match occurred for the Congruent–Different category—CW matching was 87 msec. longer,
on the average, than WC matching for this category. The parenthesized number following each plotted point was the total number of errors in the experiment for that condition. The longer response latency for CW matching for the Congruent–Different category was accompanied by a higher error rate with nine incorrect same responses compared to only three for WC matching.

These differences in response latency and errors between the two match directions for the Congruent–Different category provide a strong clue to the nature of central processing of colors and words, at least as it occurred in the present study, and are discussed later in considerable detail.

The six levels of the combination factor corresponding to the particular different combinations of the colors and words represented somewhat different things for three subgroupings of the seven categories. Despite this fact, a significant main effect for combinations appeared for the overall analysis, $F (5, 75) = 13.0, p < .001$. All four Different categories utilized the same criterion to designate these levels, and in the separate analysis of these categories the main effect was also significant, $F (5, 75) = 9.0, p < .001$. The abscissa of Figure 3 designates the relevant word value and relevant color value that combine to differentiate these different levels. For example, at the first point on the abscissa the relevant word is red and the relevant color is blue. Main effect means for these six levels of the combination factor fall at the midpoints between the two curves presented in Figure 3. Of the 15 possible individual comparisons between these six means, the Newman-Keuls procedure indicated nine to be significant at at least the .05 level. Designating the points on the abscissa from left to right as Level 1 to Level 6 of the combination factor, Level 1 differed from all other levels with the exception of Level 5. Level 4 differed from Level 5, and Level 6 differed from all other levels.

The interaction appearing in Figure 3 between combination and match direction was not significant, $F (5, 75) = 1.55, p < .20$. It is plotted primarily to illustrate the main effect, but it also indicates possible differences between CW and WC matching for the stimuli containing the words red and blue. Parenthesized numbers again represent the number of errors in the experiment for the plotted condition.

The main effect of sessions was highly significant, $F (1, 15) = 13.9, p < .01$. Average response time for the first session was 1,037 msec. and for the second, 947 msec. No interaction of other variables with this factor was significant, indicating that this improvement over sessions was fairly constant for different categories, directions of match, and particular word–color combinations. No other main effects or interactions in the analyses approached significance.

**DISCUSSION**

Basic differences between same and different judgments are illustrated in the present results. The faster time for same judgments in Control–Same and Congruent–Same categories than for different judgments parallels the findings of Bindra et al. (1968) for within-attribute comparisons of colors; Posner, Boies, Eichelman, and Taylor (1969) for within-attribute comparisons of letters; and Seymour (1970) for
Cross-attribute comparisons between words and shapes. Such faster production of same judgments led Bindra et al. (1968) to conclude that different decision processes operated for same and different judgments. New evidence for separate decision processes comes from the even faster response times for same responses when the irrelevant stimulus agrees with the relevant stimuli (Congruent–Same) than in the Control–Same condition. Congruent combinations of words and colors can actually increase the speed of color naming over control conditions in Stroop paradigms using individual stimuli (Dyer, 1971b), and a possible explanation of the fast same responses in the Congruent–Same category would be that the congruent "irrelevant" word speeds encoding of the color stimulus. However, this would also predict fast different responses for the Congruent–Different category which definitely did not appear. In fact, every category of different judgments was significantly slower than the Control–Different condition. This differential utilization of redundancy in same judgments than in different judgments argues strongly for different decision processes for the two types of response.

Along with this illustration of basic differences between same and different judgments, considerable insight into the processes underlying cross-attribute matching can also be gained from the present data. The primary clue to these processes is the large difference in response time between the two match directions for the Congruent–Different category. The irrelevant stimulus and the relevant stimulus with which it was combined were equivalent in this category. When the match direction was color to word, with the irrelevant stimulus being a word, different responses were greatly delayed relative to the other match direction, in which the irrelevant stimulus was a color (see Figure 2). In addition, error rates were much higher for the color-to-word match direction.

Cross-attribute comparisons require transformation of one or both stimulus attributes to make them comparable at some central location. The difference in the Congruent–Different category for the different match directions can be explained by assuming that less transformation is required for one attribute than the other, and thus the form of the untransformed irrelevant stimulus for one match direction is already close to that form in which the central comparison is made. This assumption implies that for the other match direction the irrelevant stimulus would require transformation to be in this "central comparison" form. The situation with the irrelevant stimulus near to the form of the central comparison would be expected to provide more difficulty for the making of comparisons for at least two reasons. First, the relevant stimulus with which the irrelevant stimulus is combined would require transformation, and while this additional process is carried out, attention would be directed to the exact position of the irrelevant stimulus. Second, comparisons of the transformed relevant stimulus could easily be erroneously made to the irrelevant stimulus since it is in the same form as the other relevant stimulus involved in the comparison.

Since it was in matches of colors to words, where the irrelevant stimulus was a word, that long response times and high error rates occurred, the implication of the above reasoning is that in comparisons between words and colors the relevant color is transformed into a word or, at least, into a form that is more similar to words than to colors. In addition, this transformation of colors to words is implied for both directions of match.

Given that this analysis of the differences for the Congruent–Different category is correct, some predictions also follow for the other Different categories. For the Incongruent–Different B category, where the irrelevant stimulus is the same as the lower relevant stimulus, an accidental comparison during CW matching of a transformed color to the irrelevant word instead of to the relevant word would not expend as much time as it did in the Incongruent–Different A category because the irrelevant and relevant words are the same. When this similarity was noted—and the result of Treisman and Fearnley (1969) showing fast within-attribute matching suggests this would not take long—the erroneous comparison with the irrelevant word would not have to be repeated. The average response time for the Incongruent–Different B category was 23 msec. faster than for the Incongruent–Different A category (CW matches), and although this was not significant, it is in the direction that supports this interpretation. In addition, the Incongruent–Different B category was the only Different category that allowed faster comparisons for CW matches than for WC matches.

Unlike the conclusion of the present study that the attribute of color is transformed to a word code, Seymour (1970) concluded that
in comparisons of words and shapes, words are converted to shape codes for the central match. His conclusion was reached on the basis of faster comparisons when the word was fixated first than when the shape was fixated first. Fixation was not controlled in the present study, but if one assumes that the top of the two stimuli was fixated first, then the different results for the two match directions for the Control–Same category, although not significant, are again contrary to those of Seymour for words and shapes. A basic difference may exist between the processing of color and shape. Further research is planned where comparisons between colors and words are made with the order of fixation controlled.

The differences that appeared for particular combinations of words and colors are not readily explained. Differences in time for processing of particular words and colors may constitute a partial explanation. But the data presented in Figure 3 indicate that color effects depend on the word and that word effects depend on the color. Further explorations utilizing additional colors and words may permit better understanding of these reliable combination differences.

In providing evidence about the form in which the central comparison of colors and words was made, the Congruent–Different category stimuli, in conjunction with stimuli from other categories, demonstrate potential for similar analyses of other stimulus dimensions. Dyer (1971a) explored a form of the Stroop test that utilized achromatic shades and names of achromatic shades in combination instead of colors and color names. Shor (1970) combined arrows with words designating directions and also obtained a small amount of interference similar to that of the Stroop test. Dyer (1972) found a small but significant amount of interference to direction naming with incongruent direction names as the movement stimuli. It should be possible to construct stimulus pairs similar to those of the present study but combining values of other visual stimulus dimensions such as achromatic shade, spatial position, movement direction, size, etc., with words designating values of the respective dimensions. Analysis of the time required for comparisons between the words and stimulus values for different categories and match directions should permit determination of the central transformation that takes place to allow these comparisons.

REFERENCES


DYER, F. N. The duration of word meaning responses: Stroop interference for different pre-exposures of the word. Psychonomic Science, 1971, 25, 229–231. (b)


(Received June 9, 1972)