Perceptual encoding in comparative judgments of race*

KENNETH E. FRIEND

Carnegie Mellon University, Pittsburgh, Pennsylvania 15213

Ss made comparative judgments along the black-white racial dimension using opposite response sets ("Which is blacker?" or "Which is whiter?"). Three classes of pictures were used: B (black). W (white). and M (racially mixed). For the BM judgments. "Which is whiter?" took significantly longer. The MW and BW judgments showed no difference between response instructions. Analysis of the stimuli showed that skin tone was not the primary perceptual attribute used in racial encoding. The observation of a funnel effect in the judgment latencies is shown to be consistent with other research demonstrating the effect of response set in comparative judgments. Further. the results support sociological observations about the cultural use of the linguistic labels "black" and "white."

In the comparison of two stimuli. a growing body of research has shown an interaction between (a) the absolute positions of the stimuli along the judgmental dimension. and (b) the response set given by the E's instructions for such judgments as brightness (Audley & Wallace. 1964). pitch (Wallace & Audley, 1964), color preference (Shipley. Coffin. & Hadsell. 1945; Shipley, Norris. & Roberts. 1946). age (Ellis. 1972). and probability (Marks, 1972). In all these studies, the difference in response times for opposite instruction sets (e.g.. "choose brighter" vs "choose darker") varies with the location of the stimuli on the perceptual dimension; this result is labeled either the "crossover effect" (if the pattern of response latencies completely reverses) or the "funnel effect" (if there is no difference in latencies at one end of the dimension).

This study investigated the perceptual encoding of racial characteristics by examination of data on comparative judgments along the black-white racial dimension. A model recently proposed by Ellis (1972) is used in the interpretation of the data. This model assumes that the time between stimulus presentation and response execution is occupied by a series of processing stages which contribute additively to the total reaction time (cf. Sternberg. 1969). The response instructions are encoded in a preliminary stage (which does not enter into the reaction time). For the brightness dimension used by Audley and Wallace (1964). the response instructions might be symbolized as (brighter(?)) or (darker(?)).[1] In Processing Stage 1 of the Ellis model, the stimuli are encoded in a similar form: (brighter(right)) or (brighter(left)) if the stimuli are relatively bright. and (darker(right)) or (darker(left)) if the stimuli are relatively dark. Stage 2 consists of a comparison between the encoding terms (the outer strings) of the response instruction and stimulus encodings. If the result of this comparison is a mismatch, the directional indicator (the inner string) of the

stimulus encoding is reversed. This reversal adds a time increment to the reaction time. Stage 3 makes the response appropriate to the directional indicator (the inner string) of the stimulus encoding.

The model proposed by Ellis (1972) will predict the crossover effect if the result of Stage 1 is an encoding consistently in terms of the overall level of the stimulus pair. For example, in Experiment 4 of Audley and Wallace (1964) (which actually produced a crossover), the background illumination was intermediate to the levels of the stimulus pairs—this permitted the darker stimuli and the lighter stimuli to be consistently encoded in the appropriate linguistic terms. Similarly, in the relative age judgments (Ellis, 1972), the Ss' ages (college sophomores) were intermediate to the ages of the younger and older pairs of faces to be compared permitting consistent encoding. The funnel effect is obtained from the model when the result of Stage 1 is an inconsistent encoding of one of the stimulus pairs. On a given trial. according to the linguistic model, the stimulus pair will be encoded in just one linguistic form. However. it is clear that inconsistency in encoding (equiprobable use of either linguistic form) for pairs of stimuli at one end of the dimension will produce a funnel effect through an averaging artifact (Ellis, personal communication). Inconsistency in encoding might arise from one of several conceivable causes. In the Audley and Wallace (1964) study, opposite response instructions gave equal latencies when the stimuli were near the level of the background illumination; when stimulus pairs are close to the background level, there is no reason for the judgment encodings to be biased toward one linguistic form. Inconsistency in encoding might also arise if the stimuli were at opposite ends of the perceptual dimension. In such a case, there would be no bias in the stimulus pair which might lead to consistent encoding in one type of linguistic form.

In discussing the perceptual encoding of racial characteristics in terms of the relative judgments paradigm, it is natural to think of a black-white continuum (the task used here was presented to Ss in these terms). Race, as a biological concept, is not well

defined because difference between the races "... are typically quantitative and not qualitative, matters of degree and not kind [Pettigrew, 1964, p. 59]." Further, although we usually speak of race in pigmentation terms (e.g., black and white), "... an individual's skin color largely depends on the activity of a single enzyme system [which] is apparently influenced by as many as five different genetic components which are genetically independent of other characteristics which differentiate Negroes and Caucasians [Pettigrew, 1964, pp. 60-61]." In spite of these facts supporting the notion of a skin pigmentation dimension in particular or a black-white genetic dimension in general, people (of both races) in the American culture tend to use the terms "black" and "white" to denote nominal categories. As Pettigrew (1964, p. 69) notes, "No matter how Caucasian one's genes may be in origin, one known trace of Negro ancestry makes you 'Negro.'"

These considerations lead to three possible mechanisms for the linguistic encoding of racial characteristics. Each mechanism postulates that certain physical characteristics are most important in the encoding process. These mechanisms are: (a) genetic dimension encoding—judgments about blackness or whiteness are made along a racial dimension according to such features as skin shade, nose and lip shape, and hair form; (b) skin tone encoding—judgments about blackness or whiteness are made along a dimension defined purely in terms of skin shade; and (c) nominal encoding—discriminations are made by use of the same characteristics mentioned above, but the linguistic form of the encoding is determined by the nominal categories in which the stimuli appear.

Facial photographs were used in this study. For the purposes of the experiment, three classes of pictures were defined: Negroid or black (B), Caucasian or white (W), and a genetic mixture of the two races (M). The three possible pairings of these stimulus classes (BM, MW, BW) served as the stimulus pairs to be discriminated. In terms of a genetic dimension ranging from Negroid through Caucasian, the average positions of the BM and MW pairs are at opposite ends of the dimension. If Ss use genetic dimension encoding, then the overall level of the BM pairs is black, while the overall level of the MW pairs is white. A similar statement holds for skin tone encoding. However, in terms of a nominal encoding, the MW and BW pairs both *cross* racial boundaries, while judgments about the BM pairs would occur within the single racial category of "black." The prediction for any of the encoding mechanisms is that the BM pairs would be encoded in terms of blackness. Because the results of Ellis's (1972) Stage 1 would be an encoding of the form (blacker(right)) or (blacker(left)), the response instruction "Which is whiter?" should take more time to execute than the response instruction "Which is blacker?".

Predictions for the MW and BW pairs are not so easy.

It follows from the discussion above that these pairs will *not* be consistently encoded in terms of blackness. If Ss make use of the genetic dimension or skin tone encoding, one would predict an encoding in terms of whiteness for the MW pairs. This would lead to observation of a crossover between the BM and MW stimulus pairs. However, if the nominal categories of "black" and "white" predominate in the encoding of stimulus pairs, then the MW and BW pairs are both likely to show no difference in average judgment times for opposite response instruction questions, since these pairs both cross racial boundaries. That is, stimuli crossing a racial boundary would produce inconsistent linguistic encodings, and this would produce a funnel effect if nominal encoding occurs.

METHOD

Selection of Stimuli

An attempt was made to obtain pictures which were highly similar in all respects except their membership in one of the three experimental categories: B (black), W (white), or M (racially mixed). All pictures of nonwhite males (more exactly, males clearly belonging to the B or M class) were clipped from a local high school yearbook. The nearest white male picture to each of the above pictures was also clipped. All pictures were either full- or three-quarters-face black and white photographs posed with a suit and tie. The pictures were 6.35 cm high x 4.44 cm wide with the face being roughly 3.8 x 2.5 cm in size.

These pictures were shuffled randomly and then assigned to the B, M, or W class by the E (who served as one judge). The pictures were then rejudged by another white judge and a black judge. All judges agreed on the members of the W class. Only pictures for which all three judges agreed were retained in the B and M classes. Pictures from the B, M, and W classes were then assigned randomly to the BM, MW, and BW pairs. Enough pictures survived the judgmental procedure so that four pairs might be formed for each of the three stimulus pair categories. The pictures were mounted side by side on white index cards for tachistoscopic presentation.

Subjects

Ten white students (seven male, three female) from Carnegie-Mellon University served as Ss in this study.

Procedure

The following instructions were read to each S in introducing the task: "In this task you will be viewing a series of cards—each card containing a pair of pictures. The pictures are facial photos of individuals who, genetically speaking, range from Caucasian or white through Negroid or black. Before viewing a set of cards, you will be given one of two questions: 'Which is blacker?' or 'Which is whiter?' Your job in this task is to decide as quickly as possible which is the correct answer to the question and then press either the right or left button on the response box according to which of the two pictures—right or left—answers the question. A group of black and white judges have viewed these pictures at their leisure earlier and their judgments will define the correct answer." These instructions were followed by the usual admonition about trying to be both quick and accurate. The details of the tachistoscope operation and response box procedure were then explained to the S.

Each block of trials consisted of a randomized presentation of the 12 stimulus pairs. Each S was exposed to one block of

practice trials followed by six blocks for which reaction times were recorded. For half of the Ss. the question "Which is blacker?" was given for the first six practice trials. followed by the question "Which is whiter?" on the remaining practice trials. The "blacker" and "whiter" questions were then alternated from block to block for the six blocks of actual trials. The order of questions was reversed for the other half of the Ss.

Stimulus pairs were centered in the viewing screen of a tachistoscope at a viewing distance of about 0.5 m (at this distance. the pair of pictures subtended a visual angle of 8 deg in the vertical and 11 deg in the horizontal). The response box had three buttons. The two outside buttons were used for making the actual response. while the center button was used by the S to initiate each trial. Stimulus onset occurred exactly 1 sec after the center button was depressed. Stimulus presentation was response terminated: average presentation time was about 590 msec (the shortest exposure time was 330 msec). Reaction times were recorded to the nearest 10 msec. Ss were informed when they made errors.

RESULTS

Each S made 12 responses for each of the six experimental conditions (four pairs in each stimulus category times three blocks for each question). The median latency for each of the six conditions for each S was calculated from the 12 or fewer obtained latencies (only times for correct responses were used). Medians rather than means were used for the individual data, since. typically. a few judgment times were quite longer than the others. Table 1 and Fig. 1 show the mean response latencies which were obtained by averaging the 10 individual medians.

A repeated measure analysis of variance demonstrated a significant effect for response instruction question $[F(1,9) = 6.74, p < .05]$ and an interaction between response question and stimulus pair category which just misses significance at .05 level $[F(2,18) = 3.49. p = .06]$. There was also a significant effect of stimulus pair category on reaction time $[F(2,18) = 25.11. p < .001]$. This latter effect is presumably due to simple differences in discriminability for the different stimulus pairs. As Table 1 shows. the question "Which is blacker?" is faster for the BM stimulus pairs $[t(9) = 2.96, p < .01,$ one-tailed]. The difference in response instruction questions did not differ significantly from zero for either the MW or BW pairs. The results indicate a funnel effect.

The error data are shown in Table 2. Over the six cells, the rank order correlation between error rate and mean reaction time is .83. indicating that there was no speed-accuracy tradeoff which might bias the reaction time data. The overall error rate was 5.0%.
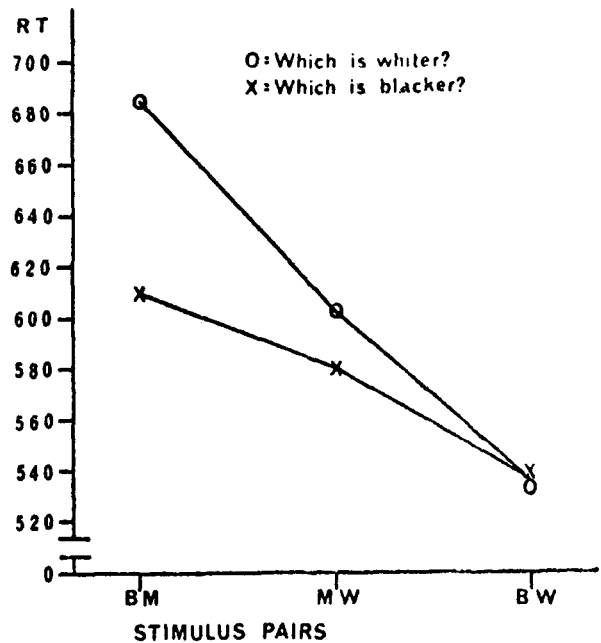

Fig. 1. Mean reaction times for racial judgments.

Table 2
Error Data for Racial Judgments

Stimuli	Whiter?	Blacker?	Marginals
BM	10	6	16
MW	9	8	17
BW	1	2	3
Marginals	20	16	

ANALYSIS OF THE STIMULI

The existence of a funnel effect seems to indicate that nominal encoding determines the linguistic form in which the comparative judgments are cast. In the judgment process itself, one would like to know whether a simple dimension such as skin tone was used in forming the racial judgments or whether the discriminations were based on more complex attributes of the faces. Indeed, the introspective reports of the Ss sometimes indicated that they felt they were merely using the overall shading of the photographs for making their responses. Although no data were obtained on the perceptual attributes that the Ss actually used, an analysis of skin tone was made to determine if this physical property was sufficient to account for the observed data on comparisons.

A Welch Densichron (having an orifice about 4 mm in diam) was used to sample reflective densities of photographed skin tone from six standard locations on the pictures (forehead, nose, chin, neck, and two cheeks). The Densichron was zeroed to the white index cards. Readings from this instrument are base-10 logarithms of the decrease in reflected light (thus. a

Table 1
Mean Reaction Times for Racial Judgments in Milliseconds

Stimuli	Whiter?	Blacker?	Difference	t Statistic
BM	685.0	610.5	74.5	2.96
MW	601.5	580.5	21.0	0.88
BW	535.5	536.5	−1.0	0.10

Table 3
Reflective Densities from the Three Stimulus Classes

	Forehead	Nose	Chin	Neck	Cheeks	X̄
B	.11	.17	.39	.58	.22	.29
M	.08	.12	.25	.52	.15	.22
W	.11	.16	.16	.40	.20	.20
X̄	.10	.15	.27	.50	.19	

reading of 1.00 would indicate 1/10 of the intensity of reflected light). Table 3 shows average readings for the three stimulus classes and five sample locations (readings from the two cheeks were averaged). An analysis of variance of these data was performed with stimulus class as a between-Ss variable and sample location as a within-Ss variable. As can be seen from the marginals in Table 3, there was quite a large main effect due to sample location [$F(4,84) = 75.50$, $p < .001$]. The interaction term was also quite significant [$F(8,84) = 3.63$, $p < .005$] compared to the relatively weaker main effect of stimulus class [$F(2,21) = 3.93$, $p < .05$].

The variation in photographed skin tone within a given face is larger than the variation in skin tone between the stimulus classes used in this experiment. In fact, the main effect of stimulus class seems to lie mostly in the greater darkness of the B stimuli compared to either the M or the W stimuli. The difference in average facial darkness between the M and W pictures is very small. However, the judges had no difficulty in distinguishing the W pictures from either the B or M pictures. Further, estimating from the facial darkness data, the ordering of difficulty in judgments should be MW as most difficult, while BW and BM should be about equal and noticeably less difficult. The actual ordering of error rates does show MW as most difficult (7.1%), but BW and BM are not equal (1.3% and 6.7%, respectively). The lack of correspondence between the data on photographed skin tone and the difficulties in comparative judgments indicate that skin tone is not the most important perceptual discriminator between stimulus classes. The MW discriminations (at least) must have been based largely on more complex attributes such as facial form or properties of the hair.

DISCUSSION

As hypothesized, there was a difference in judgment times between the response instructions "Which is whiter?" and "Which is blacker?" for the BM stimulus pairs. This result might be explained by saying that these stimulus pairs are at the blacker end of the skin tone dimension and that the result therefore directly parallels the effect obtained by Audley and Wallace (1964) for brightness-darkness. However, this explanation can be ruled out because skin tone alone does not seem to be sufficient to differentiate the stimuli in the manner in which Ss and judges made their comparisons. Similarly, this result might be explained by saying that these

stimuli are at the (racially) blacker end of a genetic dimension. This explanation implies that the "whiter" question should be faster for the MW pairs that are at the whiter end of the genetic dimension—this result was not observed. The data therefore indicate that the encoding of the BM pairs in terms of blackness is due to the culturally determined perception that both faces in these pairs are "black."

Response times for the MW and BM stimulus pairs showed no differences between the "blacker" and "whiter" questions, even though the MW pairs showed no difference in skin tone while the BW pairs showed noticeable differences. This result supports the notion that the black-white nominal categories are the primary terms used in the encoding of racial stimuli; the M stimuli are included in the "black" category. Thus, the MW and BW pairs both crossed the racial boundary. This hypothetically leads to inconsistent encodings of the judgments about these pairs and hence to no latency differences for opposite response questions.

The observation of a funnel effect in this study adds to a growing body of research literature which shows the importance of response set in comparative judgments. The cultural use of the term "black" to refer to all individuals with any apparent Negro ancestry was shown to influence the encoding of pictures in this task. This influence was consistent with the Ellis (1972) model for the crossover effect, and it demonstrated the validity of sociological observations about the use of racial terms (Pettigrew, 1964, Chapter 3).

REFERENCES

Audley, R. J., & Wallace, C. P. Response instructions in the speed of relative judgments. I. Some experiments in brightness discrimination. British Journal of Psychology, 1964, 55, 59-73.

Clark, H. H. Linguistic processes in deductive reasoning. Psychological Review, 1969, 76, 387-404.

Clark, H. H., & Chase, W. G. On the process of comparing sentences against pictures. Cognitive Psychology, 1972, 3, 472-517.

Ellis, S. H. Interaction of encoding and retrieval in relative age judgments: An extension of the "crossover" effect. Journal of Experimental Psychology, 1972, 94, 291-294.

Marks, D. F. Relative judgment: A phenomenon and a theory. Perception & Psychophysics, 1972, 11, 156-160.

Pettigrew, T. F. A profile of the Negro American. Princeton: Van Nostrand, 1964.

Shipley, W. C., Coffin, J. I., & Hadsell, K. C. Reaction time in judgment of color preference. Journal of Experimental Psychology, 1945, 35, 206-215.

Shipley, W. C., Norris, E. D., & Roberts, M. L. The effect of changed polarity of set on decision time of affective judgments. Journal of Experimental Psychology, 1946, 36, 237-243.

Sternberg, S. The discovery of processing stages: Extensions of Donders' method. In W. G. Koster (Ed.), Attention and performance II. Acta Psychologia, 1969, 30, 276-315.

Wallace, C. P., & Audley, R. J. Response instructions and the speed of relative judgements. II. Pitch discrimination. British Journal of Psychology, 1964, 55, 121-132.

NOTE

1. The Ellis model assumes a linguistic-like encoding of the response instructions and of the stimulus judgments. Here, these encodings are explicitly represented in propositional form. This representation is consistent with the deep structure propositional representations used by Clark (1969) and by Clark and Chase (1972).