

Recognition memory for faces following nine different judgments

EUGENE WINOGRAD

Emory University, Atlanta, Georgia 30322

Recognition memory for male faces was tested following nine different kinds of judgments made when the face was studied. Memory was poorest following questions about a particular physical feature, e.g., size of nose or straightness of hair, than when other kinds of questions were asked. In general, it made no difference for recognition whether the face had been positively or negatively categorized when it was studied.

It is by now well documented that different activities engaged in at the time that a verbal event is encoded lead to large differences in memory whether measured by recall (Hyde & Jenkins, 1973) or recognition (Craik & Tulving, 1975; Schulman, 1971). The present research investigated the effects of nine different orienting tasks on subsequent recognition memory for adult faces and is basically an extension of research reported by Bower and Karlin (1974) and Warrington and Ackroyd (1975). The extension consists of examining the effects of a greater number of orienting tasks and, further, of inquiring whether or not recognition is better for faces classified positively, that is, as possessing the characteristic in question (e.g., a big nose), or for those classified negatively.

The three judgments made by the subjects in Bower and Karlin's (1974) experiments were about the sex, honesty, or likableness of the face. Bower and Karlin reasoned that sex classifications (male or female) represent a shallower level of processing, in terms of the depth of processing dimension proposed by Craik and Lockhart (1972), than trait judgments and should therefore produce poorer recognition. This expectation was confirmed with both incidental and intentional learning instructions. Warrington and Ackroyd (1975) compared memory for faces following three kinds of instructions: (1) One group was not given an explicit orienting task to perform but, instead, was read standard intentional learning instructions; (2) another group classified the faces as pleasant or unpleasant; (3) a third group classified the faces as tall or short. It is interesting to note that Warrington and Ackroyd regard judgments of both pleasantness and height as semantic in nature but as differing in what they call "relevance." The problems encountered in identifying particular orienting tasks with theoretical levels of processing have been discussed by Postman (1975), and similar problems may apply to the concept of "relevance." In any case, recog-

This research was supported by Grant GB-18703 to the author from the National Science Foundation. I am grateful to Claire Coles and Nancy Hepler for their assistance.

nition of faces was better following judgments about pleasantness than for the standard or height conditions, the last two conditions not differing from each other. As in the Bower and Karlin (1974) study, the kind of orienting task engaged in led to a difference in memory.

In the present experiment, nine ordering tasks were chosen. All were presented in the form of questions allowing for an answer of yes or no. Three questions concerned physical characteristics (straight hair, big nose, heaviness), three concerned traits or attributes which were chosen to represent psychologically continuous dimensions (intelligence, anxiety, friendliness), and three questions asked about occupational roles (does he look like an actor, a teacher, a businessman?) which were viewed a priori as being dichotomous rather than continuous. Craik and Tulving (1975) found better recognition memory for words classified positively than negatively in phonetic (does it rhyme with weight?) and semantic (is it a kind of fish?) classification tasks; however, when questions about physical dimensions were asked, e.g., size and weight, equivalent memory outcomes were obtained following positive and negative decisions. We were interested in whether an analogous outcome would result when faces were tested, relying on our intuitions that such traits as anxiety and intelligence are psychologically continuous, while roles are not. Superior recognition for positive over negative occupational judgments, together with a finding of equivalent recognition for faces categorized either way on traits, would support our intuitions and extend Craik and Tulving's (1975) finding to a different class of stimuli.

In terms of levels of processing, we anticipated that judgments about physical characteristics would lead to poorer recognition than those dealing with traits or roles, with no expectation of an overall memory difference between the trait and role judgments. Since relatively little is known about the recognition of faces following different orienting tasks, this research was entered into in an exploratory frame of mind and the findings are presented to aid others contemplating research on related problems.

Table 1
Recognition Performance and Judgments

	Physical Characteristics			Traits			Occupations		
	Big Nose	Straight Hair	Heavy	Intelligent	Anxious	Friendly	Actor	Businessman	Teacher
Overall proportion correct	.62	.62	.75	.77	.69	.78	.81	.76	.75
Total positive responses	282	294	122	336	247	279	263	235	245
Total negative responses	294	282	454	240	329	297	313	341	331
Proportion correct for positive responses	.66	.63	.82	.74	.70	.80	.84	.70	.75
Proportion correct for negative responses	.58	.61	.73	.76	.69	.77	.79	.80	.76

METHOD

Subjects

The subjects were 73 students of both sexes enrolled in the author's undergraduate course on human learning and memory. The experiment was conducted during the first week of the quarter as part of the normal course work. Data from one subject were discarded prior to scoring for purposes of counterbalancing.

Materials and Design

One hundred and forty-four 35-mm black-and-white slides were made from pictures in a British casting directory. All the pictures used were of white males of varied ages, chosen to be unknown to American viewers. Faces with facial hair or spectacles were not used. Most of the pictures showed both head and shoulders. The study set consisted of 72 faces; the test set contained duplicates of these plus an additional 72 faces presented in random order.

The design was completely within-subjects, with each subject rating eight faces on each of nine characteristics. Each subject received a booklet containing 72 questions, each of which was followed by a yes and no. The nine questions were as follows: Does he have straight hair? Does he have a big nose? Does he look heavy? (physical characteristics). Does he look intelligent? Does he look anxious? Does he look friendly? (traits). Does he look like a teacher? Does he look like a businessman? Does he look like an actor? (roles). The questions were ordered so that no question appeared twice before each of the other questions had appeared. A 9 by 9 Latin square was used to determine the ordering of the questions on the nine different booklets. Thus, there were nine groups of eight subjects each, each group receiving the questions in the same order. In other words, while each face was on the screen in front of the room, all nine questions were being asked at the same time, each question being asked of eight subjects.

Procedure

The subjects were told that we were primarily concerned with their perceptions of different characteristics of the faces they were to see. It was emphasized that we were interested in their personal impressions of the faces and that there was no correct answer. They were also told that their memory for the faces would be tested later but that the answers to the questions about the faces were the major concern of the study. The 72 old faces were then shown at an 8-sec rate to allow time for the subjects to read the question, study the face, and circle either "Yes" or "No" on the sheet. Immediately following the last slide, test sheets containing 144 numbered spaces were passed out. The subjects were told that half of the faces on the test series were old and half were new and asked to indicate their decision about each face by writing a "Y" for yes, if they thought it looked familiar, or an "N" for no, if it did not look familiar. The test proceeded at a 5-sec rate.

RESULTS

The overall proportion of old faces correctly recognized, or hits, is shown in the top row of Table 1 for each orienting task. Each proportion in the top row is based on 576 observations. The recognition scores seem to fall into two groups, one group containing questions dealing with size of nose and straightness of hair and another group containing all the other questions. A one-way analysis of variance on the mean hit scores for the nine tasks yielded a significant outcome [$F(8,568) = 10.85, p < .01, MSe = .028$]. The Newman-Keuls test showed that judgments about nose and hair led to significantly poorer recognition than did all other judgments ($p < .01$ in all cases). The only other significant difference was between the scores for "actor" and "anxious" ($p < .05$). The false positive rate for the experiment was .11.

The bottom four rows of Table 1 break down the recognition data according to whether positive or negative judgments were made at the time the face was studied. With these data, one can answer questions such as whether faces judged as appearing anxious are remembered better than those judged as not appearing anxious. It will be recalled that Craik and Tulving (1975) found better recognition of positively categorized words in all but one of their studies. One potentially important aspect of the present procedure is that, unlike Craik and Tulving's work with words, there is no correct answer to such questions as "Does he look intelligent?" Therefore, one cannot know in advance the distribution of positive and negative responses. The total number of positive and negative responses given to each question are given in rows 2 and 3 of Table 1, and the corresponding proportions of hits are given in rows 4 and 5. For all orienting tasks except heaviness, there is no greater split than 60-40 between positive and negative judgments of the faces; in the case of heaviness, 79% of the faces were deemed not heavy.

Do positive judgments about faces lead to better memory? In general, the answer is no. Across all nine tasks, the proportion of hits is identical at .73 following both positive and negative classifications. However, some potentially interesting findings concerning social percep-

tion may be gleaned from Table 1. Apparently, men who look like businessmen to college students are less likely to be remembered than those who do not conform to that stereotype, while the opposite seems to be the case for actors (the source of our faces was a casting directory) who look like actors. It may also be of interest that, while heaviness and a large nose are correlated with memorability, the semblance of intelligence is not.

DISCUSSION

When attention was called to a particular facial feature, recognition memory for that face was impaired relative to the other tasks investigated. Yet, classification of a face with respect to heaviness, a physical characteristic, produced recognition scores no poorer than those resulting from judgments of personality traits and occupational roles. It is possible that a more global or holistic assessment of a face is required in judging weight than in judging the size of a nose or straightness of hair. Heaviness is, after all, an attribute of the whole person. This interpretation is similar to Bower and Karlin's (1974) hypothesis that memory for faces improves as more features are examined. It seems a reasonable assumption that more features of a face are examined when a decision is demanded concerning the friendliness or intelligence of a person than whether he has a big nose. The more features encoded, the better for recognition. The results of Warrington and Ackroyd (1975) present a potential problem for this point of view in that poorer recognition of faces judged with respect to height than to pleasantness was found. Can one say with any confidence at present that more features

are sampled when faces are judged for heaviness than for height? An affirmative answer to this question would seem necessary to support the "number of features" hypothesis. Bower and Karlin's suggestion, that measuring the number of eye fixations during study may provide an index of feature sampling, deserves serious consideration in this regard.

REFERENCES

- BOWER, G. H., & KARLIN, M. B. Depth of processing pictures of faces and recognition memory. *Journal of Experimental Psychology*, 1974, **103**, 751-757.
- CRAIK, F. I. M., & LOCKHART, R. S. Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, 1972, **11**, 671-684.
- CRAIK, F. I. M., & TULVING, E. Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General*, 1975, **1**, 268-294.
- HYDE, T. S., & JENKINS, J. J. Recall for words as a function of semantic, graphic, and syntactic orienting tasks. *Journal of Verbal Learning and Verbal Behavior*, 1973, **12**, 471-480.
- POSTMAN, L. Verbal learning and memory. *Annual Review of Psychology*, 1975, **26**, 291-335.
- SCHULMAN, A. I. Recognition memory for targets from a scanned word list. *British Journal of Psychology*, 1971, **62**, 335-346.
- WARRINGTON, E. K., & ACKROYD, C. The effect of orienting tasks on recognition memory. *Memory & Cognition*, 1975, **3**, 140-142.

(Received for publication July 30, 1976.)