

Recognition of faces and verbal labels

JUNE CHANCE and ALVIN G. GOLDSTEIN
University of Missouri-Columbia, Columbia, Missouri 65201

Three groups of subjects who viewed briefly exposed target portraits were instructed to expect a recognition test in 1 week; two of these groups were told to produce written verbal labels whenever possible in the study trial. A fourth group viewed portraits and made age judgments but were unaware of the impending recognition test. Both accuracy of recognition and recall of verbal labels were measured after a 1-week delay. Instructions affecting the way a portrait is labeled had a mild effect on recognition memory; recall data offers weak support for a verbal encoding hypothesis of face recognition, but one kind of instruction did elicit better recall than another.

Whether use of verbal labels or descriptors facilitates recognition of once-seen faces when they reappear is an experimental question of both practical and theoretical significance. The role of verbal labeling in discrimination learning and perceptual memory has been examined in many studies (Gibson, 1969), but their outcomes still leave many unresolved issues pertaining to uses of verbal labeling with natural, pictorial stimuli. A potential finding that use of verbal descriptors could increase face recognition accuracy would also have important practical implications (Goldstein, 1976).

Studies of pictorial recognition memory employing heterogeneous arrays of pictorial stimuli often demonstrate truly astounding recognition performances of subjects (Haber, 1970; Nickerson, 1965; Shepard, 1967). When more homogeneous stimuli—all pictured objects of the same class—are used, recognition accuracy is much reduced. Goldstein and Chance (1974) have noted that the impressive memory performances obtained when arrays of picture stimuli are heterogeneous probably reflect the subjects' greater use of partial learning of stimuli, and less extraexperimental interference among pictures than among words. Heterogeneous arrays also lend themselves more readily to verbal coding than do more homogeneous arrays.

Faces, presented as a class, are more homogeneous than pictures of scenes or objects but still capable of eliciting some verbal labeling. Nevertheless, Goldstein and Chance (1971) found that while faces elicit more verbal labels than either inkblots or snowflakes, ease of labeling a particular stimulus is not related to ease of recognizing it. The present study tests whether use of two kinds of self-generated verbal labels for faces is related to subjects' accuracy in recognizing these faces 1 week later. It also tests how well the labels given are

recalled by subjects when the face is presented again.

METHOD

A standard recognition memory paradigm was used. During the study session, 20 full-face portraits (black-white slides, 10 men, 10 women, culled from yearbook pictures) were shown to 80 undergraduate women randomly assigned to four groups. After a delay of 1 week, all subjects were shown the 20 target faces randomly interspersed among 80 distractor (40 men) faces. Two sets of target faces were used to minimize the effects of fortuitously high or low target recognizability.

The four groups of subjects were treated alike except for instructions given to them at the beginning of the study session. Groups A, B, and C were told they would see 20 faces in the first session and would be asked to recognize these 20 faces from among a larger number of faces in the second session 1 week later. Groups A and B were asked to write a response (on an appropriately numbered response sheet) to each of the 20 target faces. Subjects in Group A were told to look at each briefly presented face and, immediately after the picture disappeared, write "some one thing about the face which you believe will help you recognize it when you see it again." Subjects in Group B were told to write "some one thing which the face reminds you of or looks like to you." Responses to each face were not forced, but subjects were requested to insert the word "No" in response to those faces which elicited no verbal response. Subjects in Group C made no written response during the study trial, but they were informed of the nature of the task and were prepared to recognize the target faces.

Group D subjects were deceived, at the beginning of the study session, with the information that they were to be subjects in a small-group decision-making experiment, but the principal part of the experiment would take place 1 week later. This deception was enhanced by requiring them to complete a questionnaire (modified Coopersmith Self-Esteem; it was administered also to all subjects in the other three groups at the beginning of their study sessions, but without explanation as to its purpose) the data of which ostensibly would be used to determine the composition of the decision-making groups. When all subjects in Group D completed the questionnaire, one of the investigators (A.G.G.) entered the testing room, carrying a slide tray, and conferred with the experimenter. The essence of this manipulation was to make it appear as though Professor X, who was in no way connected to this study, was in need of several people to make age judgments of a few faces. Following appropriate "instructions" to the group, the 20 target faces were projected, and each subject judged (in writing) the age of the individuals portrayed in the pictures.

For all groups, each of the targets was displayed for 3 sec; interstimulus intervals were 10 sec, during which time a blank, translucent slide was displayed.

This paper was presented at the meetings of the Psychonomic Society, Boston, Massachusetts, November, 1974.

Table 1
Percentage Correct, Means, and Standard Deviations of Correct Recognition of Faces for Four Task Instruction Groups

	Group Instructed to			
	A Describe	B Associate	C Look	D Judge Ages
Average percentage correct recognition	82	71	74	74
Mean number correct recognition (20 possible)	16.35	14.20	14.70	14.75
SD	2.23	3.00	2.13	3.34
N	20	20	20	20

RESULTS AND DISCUSSION

Before analysis of the data was initiated, protocols with less than 19, or more than 21, faces checked as "seen before" were eliminated (less than 1%).¹ Effects of the four instruction conditions on accuracy of recognition, summarized in Table 1, were examined in an analysis of variance ($F = 2.63; 3/72; p < .10 > .05$). The value of F obtained suggests only a slight trend toward superiority of performance for the groups which verbalized about the faces, and then only for Group A ("describe"). Indeed, Group B's ("associate") recognition performance was no different from the performance of subjects who only looked at the faces (Group C), or the incidental learning subjects in Group D (71%, 74%, and 74%, respectively).²

The overall outcome of this aspect of the study indicates that verbalization about pictures of faces per se does not necessarily contribute to accuracy of recognition performance and that even the slightly more helpful form of verbal coding found in the "describing" group may have little to offer the subject beyond any other task instruction which would be likely to insure that he paid close attention to the face stimuli when they were presented in the study trial.

In the second aspect of the study, we asked subjects in Groups A and B to recall the verbal labels they had given the "seen-before" pictures a week ago. Thirty additional subjects were also added to each of the two instructional conditions, and all subsequent analyses are based on the data of these 100 subjects. Note that the study session instructions emphasized recognition of faces, but omitted any reference to remembering the verbal responses later. Strictly speaking, "remember" is not the correct term here; during the recognition trials the subject could be considered to be in a cued recall situation since the face he thinks is familiar is on the screen at the time he is making his "remembered" verbal response. But, if the face in view is really the correct target stimulus, then the subject could be responding to the same facial features in the picture that set off his response in the

earlier study trial. If this were the case, "remember" would not be an appropriate term here. Obviously, if the latter did occur, one would expect fairly high rates of "remembering" of the verbal material.

In analyzing the verbal responses elicited during the recognition test, we did not require that the response be worded in the exact form that it was given in the study trial in order for it to be counted as accurately recalled. Instead we counted as "recalled" any response which had the same content, or expressed the same idea as the original. For instance, "wavy hair" and "kinky hair" were judged to express the same content; similarly "looks studious" and "looks like a student" were considered to be the same responses for scoring purposes. Two independent judges agreed 93% of the time about whether the response was accurately recalled or not.

Responses elicited by the "describe" instructions (A) were significantly better recalled than responses elicited by the "associate" instruction (B). When figured on the basis of 20 possible correct responses, 35% ($\bar{X} = 6.96$) were recalled by Group A and 23% ($\bar{X} = 4.50$) recalled by Group B ($F = 22.32; 1/96, p < .001$). Both groups were very much alike in the number of times they correctly recognized a particular face, but said that they could not recall what they said 1 week ago (i.e., they wrote the word "no" on the response sheet). The frequency of this response was 27% (5.38) in the describe group and 34% (6.76) in the associate group, a difference which was not reliable. Likewise, the groups were nearly identical in the number of instances where the face was falsely recognized, and a confabulated verbal response was produced. The "describe" group did this 15% of the time (3.08) and the "associate" group did it 16% (3.28).

The remaining difference in the recall of verbal responses to the faces occurred in instances where the picture was correctly recognized as a target face, but a confabulated verbal response was produced. The "describe" group had significantly more of these responses than did the "associate group," 27% vs. 18% (5.40 vs. 3.50; $F = 8.12, 1/96; p < .01$).

Several inferences and speculations seem warranted when recognition memory and verbal recall data are considered together. Instructions which affect the way in which a subject labels the face the first time he sees it have a mild effect on recognition memory for that face following a 1-week study-test interval. This could be interpreted as evidence that verbal coding is involved in the registration of and recognition memory for faces. However, this involvement appears quite weak since only one of the two verbal instruction groups performed better than the nonverbal instruction groups, and, moreover, the improvement in their performance was not impressive. In the same vein, the label recall data also offered only weak evidence for a verbal encoding hypothesis. Only 35% of verbal responses were correctly recalled by subjects,

and these responses were elicited in the presence of the target face to which the verbal response had initially been made. Relatively few verbally coded responses to faces survive the study-test interval even under the best conditions for their elicitation. We are forced to conclude that verbal coding as an aid in recognition memory for faces is normally of relatively minor importance to the human subject.

Further support for this position may be derived from some simple analysis of the characteristics of verbal responses produced by subjects in this experiment. We are struck with their undifferentiated nature, an observation also noted earlier (Goldstein & Chance, 1971). Responses like "dark hair" or "looks mean" or "looks intelligent" seem unlikely to possess sufficient uniqueness to provide the subject with much advantage in later recognition of particular faces. Interestingly, we had found in some pilot research that subjects given the task to produce verbal labels for faces tended to produce compound labels. Roger Brown (1958) and others have suggested that the length of responses—assessed in several different ways—is an index of the codability of a stimulus; the shorter the response, the more codable the stimulus, and, presumably, the higher its availability to the subject. Our findings showed that subjects' verbalizations, when not restricted as to length, are compound about 65% of the time when they are asked to describe the face, and about 85% of the time when they are asked to say what the face reminds them of. These figures are in keeping with our conclusions and suggest that verbal codes for faces are not readily available to most subjects and may be less available for one instruction than the other.

REFERENCES

- BROWN, R. *Words and things*. Glencoe, Ill: Free Press, 1958.
- GIBSON, E. J. *Principles of perceptual learning and development*. New York: Appleton-Century-Crofts, 1969.
- GOLDSTEIN, A. G. The fallibility of the eyewitness: Psychological evidence. In B. D. Sales (Ed.), *Psychology in the legal process*. Englewood Cliffs, N. J: Prentice Hall, 1976.
- GOLDSTEIN, A. G., & CHANCE, J. Visual recognition memory for complex configurations. *Perception & Psychophysics*, 1971, 9, 237-241.
- GOLDSTEIN, A. G., & CHANCE, J. Some factors in picture recognition memory. *Journal of General Psychology*, 1974, 90, 69-85.
- HABER, R. N. How we remember what we see. *Scientific American*, 1970, 222, 104-112.
- NICKERSON, R. S. Short-term memory for complex meaningful visual configurations: A demonstration of capacity. *Canadian Journal of Psychology*, 1965, 19, 155-160.
- SHEPARD, R. N. Recognition memory for words, sentences, and pictures. *Journal of Verbal Learning and Verbal Behavior*, 1967, 6, 156-163.

NOTES

1. Subjects had been instructed at the time of the recognition test to try to limit their seen-before responses to 20.
2. Even though the F-value approached but did not reach the conventionally accepted level of significance, Duncan's multiple range test was applied to the data indicating that only the difference between Groups A and B is significant ($p < .05$). It seems reasonable to conclude that there is some evidence here of a mild effect of one kind of verbal coding on recognition memory for faces. Fortunately, this conclusion received additional support in another study where the instruction conditions of Groups A and B were essentially replicated with different subjects and when the data from both studies are combined, the difference between the means reaches a respectable level of significance; 85% correct for Group A vs. 74% for Group B ($p < .01$).

(Received for publication January 8, 1976.)