Retention interval and face recognition: Response latency measures

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

Effects of duration of retention interval on recognition memory for faces were tested immediately after study or after 2 or 7 days. In addition to counting correct and incorrect responses, we measured response latencies of subjects' decisions. Subjects were tested on photos of both Caucasian and Japanese faces. Length of retention interval was not associated with number of correct responses, although longer intervals were associated with more numerous false alarms. As retention intervals were greater, latencies of both hits and false alarms were longer, reflecting subjects' greater difficulty in retrieving their memories of faces. Length of retention interval did not differentially affect responses to Caucasian and the more difficult to recognize Japanese faces.

Although the research literature would lead us to expect that memory for faces, as for other stimuli, should become increasingly vague as length of retention interval increases, many studies of effects of retention intervals on memory for faces seen briefly and only once could be interpreted to suggest that faces are somehow exempt from the laws of memory that apply to other experiences (Chance, Goldstein, & McBride, 1975; Deffenbacher, Carr, & Leu, 1981; Egan, Pittner, & Goldstein, 1977; Goldstein & Chance, 1970; Laughery, Fessler, Lenorovitz, & Yoblick, 1974; Shepherd, 1983). In contrast, other studies do show that rates of correct recognition of such faces are lower as the length of time between study and test is longer (Barkowitz & Brigham, 1982; Brigham, Maass, Snyder, & Spaulding, 1982; Courtois & Mueller, 1981; Krouse, 1981; Malpass & Devine, 1981; Shepherd & Ellis, 1973; Yarmey, 1979).

The question that needs to be asked is "Why does delay sometimes seem to affect face recognition and sometimes not?" Several possible reasons can be examined. First, studies failing to demonstrate memory loss actually often do show some loss, although not a significant amount. Second, examining studies that address the effects of delay reveals a variety of differences among these studies. For example, they differ in whether and how false alarm measures were obtained, in whether study and testing were done with identically posed stimuli, in what medium—photos, videos, live targets—was used for presentation of study faces, and in whether testing was done within the same medium or in a different one. Differences also occur in the length of delay interval used, the number and distinctiveness of faces shown, and the degree to which the original presentation of targets was attention compelling. Finally, few previous investigations have employed

any measures of strength of face memory other than frequencies of correct recognition responses and errors.

In the present study, we investigated whether simple counts of responses may be insensitive to decrements in face memory with lengthened delay intervals. Comparing studies showing decreased memory with delayed recognition testing with studies that did not find differences suggested to us that some of the disparity in findings may be related to the relative difficulty of the face recognition task employed. More difficult tasks may be more sensitive to memory losses.

The question of the durability of face memory is an important one. Is there a separate kind of memory for faces—and perhaps for other sorts of visual stimuli—that is more durable than our memory for words? Is most of our accumulated knowledge of the processes of verbal learning irrelevant to face learning? Also on the practical level, can it be assumed, as it was until quite recently by the American justice system, that the person memory of eyewitnesses is everlasting? Only in relatively recent years has the Supreme Court of the United States indicated that the length of time between crime and identification should be considered in assessing the value of eyewitness testimony (e.g., *Manson v. Brathwaite*, 1977; *Neil v. Biggers*, 1972).

In the present study, the subjects were tested for face recognition memory at three different retention intervals: immediately after study trials, or 2 or 7 days later. We measured response latencies and counted correct and incorrect responses. We also attempted to test the impact of relative task difficulty on the effects of delay by testing all subjects for memory for Caucasian faces and for relatively more difficult Japanese faces (see Shepherd, 1981, for a summary of memory for other-group faces). All subjects were Caucasian.

METHOD

The subjects were 59 undergraduate men and women assigned in approximately equal numbers to three delay conditions. At the outset of

the study trials, each subject was told that we were interested in how accurately briefly seen faces of strangers could be recognized. The subject was told to pay attention to the faces in the study slides, because he/she later would be asked to recognize these faces among similar new faces. The subject was also told that he/she would see some words and be asked to recognize them later. Target faces and words were projected on a screen for 5 sec. One second elapsed between slides.

Each subject saw two sets of face study slides: 16 Caucasian photos and 16 Japanese photos. Sets were presented in counterbalanced order to approximately half the subjects of each sex within each delay condition. Words were projected last. Face slides did not include unusual faces or those showing nonfacial cues. Words were a list of 5 animal names. All sets of study slides were shown, with brief pauses between sets, before any tests of recognition were performed.

Before the memory test, the subjects were reminded that the purpose of the experiment was to find out how accurately they recognized briefly seen faces. Also, they were told that we were interested in how quickly they could tell us whether a projected face or word was old or new. The subject rested his/her right index finger lightly between two microswitches mounted on a small panel firmly attached to the desk in front of the subject. The right and left switches were labeled "OLD" and "NEW," respectively. Two large signs on the projector screen also identified left as "NEW" and right as "OLD."

The subjects were told to decide as each stimulus was presented whether it was old or new, and then to push the appropriate switch quickly. After making a response, the subject returned his/her finger to the point between the switches. The subjects then practiced signaling responses with the switches by taking a recognition test for the 5 animal names seen earlier, embedded in a series with 12 other nouns. Instructions to subjects emphasized striking a balance between speed and accuracy of response.

Each time a slide was presented, a light-sensitive switch mounted on the front of the projector lens activated a standard timer, which ran until the subject pressed one of the microswitches. Pressing either switch closed the circuit, stopped the clock, and activated a light signaling the experimenter about which switch had been pressed. The experimenter recorded the subject's choice and response time, reset the clock, and then gave a verbal ready signal before projecting a new slide. Reliability of the clock was approximately $\pm .03$ sec when tested repeatedly with a signal of precise, known duration.

After the practice trials, the first test set of faces was shown. Each subject was shown a set of test faces in random order that included 8 of the original 16 target slides from that set plus 24 new faces of the same sort. Approximately half the subjects within each delay condition saw one set of 8 targets and the other half saw a test set including the remaining 8 target slides. Subjects were tested on the two sets of face slides (Caucasian or Japanese) in the same order in which they had been seen during the study trials.

RESULTS

Mean frequencies of hits and false alarms made by the subjects on the recognition tests for both Caucasian and Japanese face portraits after the three different retention intervals are summarized in Table 1. Two two-way analyses of variance (3 delay intervals × 2 picture types, with picture type as a repeated measure) were performed on these data.

An analysis using number of correct identifications of target pictures (hits) as the dependent measure shows that the subjects recognized Caucasian portraits more often than they recognized Japanese portraits at all three delay intervals [$F(1,56) = 44.94, p < .001$]. However, number of hits is not related to length of delay interval [$F(2,56) < 1$], and there is no significant interaction between delay interval and picture type [$F(2,56) < 1$].

An analysis of false alarms indicates that the subjects committed more false alarms after a delay than they did on immediate recognition testing [$F(2,56) = 8.84, p < .001$]. Examined by means of Scheffé tests, the number of false alarms made by the subjects tested immediately is less than the numbers made by the subjects tested after 2 or 7 days, but no difference was observed between the latter two groups. The subjects also made more false alarms on Japanese than on Caucasian portraits [$F(1,56) = 73.61, p < .001$]. Contrary to our prediction, false alarm rates also showed no evidence of an interaction between length of delay and picture type [$F(2,56) = 1.60$, n.s.].

A summary of means of response latencies and their reciprocals for hits and false alarms is presented in Table 2. Mean time reciprocals for correct rejections (correctly identifying a distractor picture as new) and for misses (failing to identify a target picture as old) are also included in this table. Two-way analyses of variance (3 delay intervals × 2 picture types, with picture type as a repeated measure) were performed using the reciprocal time measures for each of these four measures of recognition accuracy.

Both the time taken to make correct identification of targets and the time taken to misidentify distractor pictures as targets (false alarm) were longer as retention interval was longer [$F(2,56) = 3.95, p < .025$, and $F(2,56) = 15.39, p < .001$, respectively]. Comparison of these means by Scheffé tests shows that the response times for both hits and false alarms were shorter on immediate testing than on later testing, but that response times did not differ significantly between 2 and 7 days. However, neither the time taken to make correct rejections of nontargets (another kind of accuracy) nor the time taken to incorrectly reject targets as not seen before are significantly affected by length of retention interval.

Hits, false alarms, and correct rejections are made more slowly to Japanese than to Caucasian portraits. Values of F equaled 34.64 ($p < .001$), 5.088 ($p < .05$), and 9.025 ($p < .001$) for hits, false alarms, and correct rejections, respectively; all tests were with 1 and 56 degrees

Table 1
Mean Frequencies of Hits and False Alarms Made by Subjects When Recognizing Caucasian and Japanese Photos after Three Retention Intervals

	Caucasian Photos		Japanese Photos		Total Each Interval	
	Hits	False Alarms	Hits	False Alarms	Hits	False Alarms
Immediate Test (N=19)	6.4	2.3	5.5	4.6	6.0	3.5
After 2 Days (N=20)	6.6	3.5	5.5	7.3	6.1	5.4
After 7 Days (N=20)	6.4	5.4	5.4	8.2	5.9	6.8
Total All Intervals	6.5	3.7	5.5	6.7		

Table 2
Means of Response Latency Measures and Their Reciprocals for Hits, Correct Rejections,
False Alarms, and Misses After Three Retention Intervals

	Hits	Correct Rejections	False Alarms	Misses
		Caucasian Photos		
Immediate Test	.86 (1.16)	.86 (1.16)	.91 (1.10)	.80 (1.25)
After 2 Days	.73 (1.37)	.77 (1.30)	.64 (1.56)	.69 (1.45)
After 7 Days	.76 (1.32)	.73 (1.37)	.55 (1.82)	.72 (1.39)
		Japanese Photos		
Immediate Test	.78 (1.28)	.81 (1.23)	.77 (1.30)	.88 (1.14)
After 2 Days	.61 (1.64)	.70 (1.43)	.64 (1.56)	.71 (1.41)
After 7 Days	.62 (1.61)	.68 (1.47)	.49 (2.04)	.67 (1.49)
		All Photos		
Immediate Test	.82 (1.22)	.81 (1.23)	.84 (1.19)	.84 (1.19)
After 2 Days	.67 (1.19)	.73 (1.37)	.64 (1.56)	.68 (1.44)
After 7 Days	.69 (1.45)	.71 (1.41)	.52 (1.92)	.70 (1.43)

Note—The first number given is response time expressed as its reciprocal; lower numbers signify longer times. The number in parentheses is average time expressed in seconds.

of freedom. Nonetheless, consistent with the earlier finding concerning accuracy of recognition, no interaction effects on response time measures were observed between length of delay interval and the type of picture shown.

DISCUSSION

In this study we examined whether response times might provide a more sensitive measure of retention/forgetting of faces over time than do frequency counts of correct identifications or of false alarms. Previous studies in which correctness of response was the only measure taken of effects of delay on memory for faces had conflicting results. Many studies can be interpreted to suggest that face memory is impervious to delay, at least up to the limit of the delay intervals used in those studies. In contrast, other studies show memory losses. The question whether face memory is lost, as are other memories for briefly seen unrehearsed material, is an important one because it bears on the larger issue of whether memory for pictorially presented stimuli works differently than memory for stimuli presented in other media (Goldstein & Chance, 1974).

Although we found that the number of correct identifications did not differ when testing took place fairly soon after study trials from the number when testing took place 2 days or 1 week later, response times to make correct identifications were longer when testing took place either 2 or 7 days later than when it occurred immediately. Mean response times for hits on all faces were: immediately after study, 1.23 sec; after 2 days, 1.48 sec; and after 7 days, 1.45 sec. Slower response times after delay suggest that subjects had some difficulty retrieving the target faces, even though the faces were ultimately retrievable.

Although the response time differences observed as a function of delay are very small, they do reflect a small degree of memory failure that takes place fairly quickly after viewing faces and that is not detectable when only hits are counted. On the average, it took subjects an extra .23 sec to identify target faces after a 7-day delay compared with their response times when tested immediately. Nonetheless, small reaction time differences can indicate large processing increments. For example, average simple reaction time to a single stimulus is approximately .180 sec, whereas reaction time is .316 to a two-alternative choice and .364 sec to a three-alternative choice (Woodworth & Schlosberg, 1954).

False alarms, unlike hits, both increased in frequency and showed longer response times with longer delay between study and testing. Both these observations attest that memory for briefly seen faces becomes more difficult as recognition is delayed. Nontarget faces are more easily confused with targets, and recognition accuracy is impaired in the

sense that although actual target faces may be identified other nontarget faces are falsely identified. This latter sort of error is particularly serious in the real world of law courts and eyewitnesses. These findings support earlier ones by Egan et al. (1977), whose subjects saw two "culprits" but were asked later to make identifications from a photo array that contained only one "culprit." Hit rates remained undiminished up to 8 weeks later, but false alarm rates increased dramatically. Note that in both the earlier study and the present one, more targets were shown at study than at test, thus setting up a possible false alarm "trap" and enhancing false alarm rates (Ellis, 1984). Alternately, this procedure can be regarded as controlling a situational cue—because subjects know the number of targets originally seen—that the subject may use to support an uncertain memory on testing. Reducing the extraneous cue may make false alarms a more valid indicator of memory strength.

We were not surprised to find that subjects made both more hits and fewer false alarms on Caucasian than on Japanese portraits and that they took less time to do so. The fact that faces of other groups than one's own are harder to identify has been confirmed many times in various experimental contexts (Shepherd, 1981) and is quite robust.

The observations made earlier about reaction times and processing also have implications for the response time differences found in this study when subjects were responding to Caucasian and to Japanese faces. Average times to correctly identify faces were longer when subjects were identifying Japanese than Caucasian faces by .12, .27, and .29 sec, respectively, for our three retention intervals. These consistent differences in response times imply that there must be extra processing involved in the task of recognizing other-group as opposed to own-group faces.

Nonetheless, by using both Caucasian and Japanese faces in this study, we sought to test whether the effects of time delay on face memory would be more clearly discernible as the face task was more difficult, and it is disappointing to find that the effects of type of face shown does not interact with effects of delay intervals. Since using Japanese faces is only one of many parameters that might be varied to increase difficulty of the recognition task, further studies should address the role of difficulty using other manipulations.

REFERENCES

BARKOWITZ, P., & BRIGHAM, J. C. (1982). Recognition of faces: Own-race bias, incentive, and time delay. *Journal of Applied Social Psychology, 12,* 255-268.

BRIGHAM, J. C., MAASS, A., SNYDER, L. D., & SPAULDING, K. (1982). Accuracy of eyewitness identifications in a field setting. *Journal of Personality & Social Psychology, 42,* 673-681.

CHANCE, J., GOLDSTEIN, A. G., & McBRIDE, L. (1975). Differential experience and recognition memory for faces. *Journal of Social Psychology, 97*, 243-253.

COURTOIS, M. R., & MUELLER, J. H. (1981). Target and distractor typicality in facial recognition. *Journal of Applied Psychology, 66*, 639-645.

DEFFENBACHER, K., CARR, T. H., & LEU, J. R. (1981). Memory for words, pictures, and faces: Retroactive interference, forgetting, and reminiscence. *Journal of Experimental Psychology: Human Memory & Learning, 7*, 299-305.

EGAN, D., PITTNER, M., & GOLDSTEIN, A. G. (1977). Eyewitness identification: Photographs vs. live models. *Law & Human Behavior, 1*, 199-206.

ELLIS, H. D. (1984). Practical aspects of face memory. In G. L. Wells & E. F. Loftus (Eds.), *Eyewitness testimony: Psychological perspectives* (pp. 12-37). Cambridge, MA: Cambridge University Press.

GOLDSTEIN, A. G., & CHANCE, J. E. (1970). Visual recognition memory for complex configurations. *Perception & Psychophysics, 9*, 237-241.

GOLDSTEIN, A. G., & CHANCE, J. E. (1974). Some factors in picture recognition. *Journal of General Psychology, 90*, 69-85.

KROUSE, F. L. (1981). Effects of pose, pose change, and delay on face recognition performance. *Journal of Applied Psychology, 66*, 651-654.

LAUGHERY, K. R., FESSLER, P. K., LENOROVITZ, D. R., & YOBLICK, D. A. (1974). Time delay and similarity effects in facial recognition. *Journal of Applied Psychology, 59*, 490-496.

MALPASS, R. S., & DEVINE, P. G. (1981). Eyewitness identification: Lineup instructions and the absence of the offender. *Journal of Applied Psychology, 66*, 482-489.

MANSON V. BRATHWAITE, 432 U.S. 98 (1977).

NEIL V. BIGGERS, 409 U.S. 188 (1972).

SHEPHERD, J. (1981). Social factors in face recognition. In G. Davies, H. Ellis, & J. Shepherd (Eds.), *Perceiving and remembering faces* (pp. 55-79). London: Academic Press.

SHEPHERD, J. (1983). Identification after long delays. In S. M. A. Lloyd-Bostock & B. R. Clifford (Eds.), *Evaluating witness evidence* (pp. 173-187). New York: Wiley.

SHEPHERD, J. W., & ELLIS, H. D. (1973). The effect of attractiveness on recognition memory for faces. *American Journal of Psychology, 86*, 627-633.

WOODWORTH, R. S., & SCHLOSBERG, H. (1954). *Experimental psychology (rev. ed.)*. New York: Holt.

YARMEY, A. D. (1979). The effects of attractiveness, feature saliency, and liking on memory for faces. In M. Cook & G. Wilson (Eds.), *Love and attraction*. New York: Pergamon Press.