Recognizing familiar and unfamiliar faces

ROBERTA L. KLATZKY and FIONA H. FORREST
University of California, Santa Barbara, California

These experiments addressed why, in episodic-memory tests, familiar faces are recognized better than unfamiliar faces. Memory for faces of well-known public figures and unfamiliar persons was tested, not only with old/new recognition tests, in which initially viewed faces were discriminated from distractors, but also with tests of memory for specific information. These included: detail recall, in which a masked feature had to be described; orientation recognition, in which discrimination between originally seen faces and mirror-image reversals was required; and recognition and recall of labels for the public figures. Experiments 1 and 2 showed that memory for orientation and featural details was not robustly related either to facial familiarity or to old/new recognition rates. Experiment 3 showed that memory for labels was not the exclusive determinant of the famous-face advantage in recognition, since famous faces were highly recognizable even they were not labelable or when labels were forgotten. These results suggest that the familiarity effect, and face recognition in general, may reflect a nonverbal memory representation that is relatively abstract.

This paper concerns the nature of the memorial representation that is encoded when a face is viewed on a particular occasion—that representation to be called a “facial episode”—and how that representation differs according to the a priori familiarity of the face. To start with, we assume that a facial episode is a composite of several types of information, or components, an assumption consistent with models of episodic memory in general (e.g., Kintsch, 1974). It may include verbally coded information, visual details, more abstract visual properties, or semantic and affective attributes. Whatever its components, this representation is capable of supporting face recognition at high levels of performance—typically 70%-85% in psychological experiments (Goldstein, 1977).

Familiar faces are recognized even better than unfamiliar ones (e.g., Ellis, Shepherd, & Davies, 1979; Yarmey, 1971; the present studies). Assuming that the recognition rate reflects the memory “strength” of facial episodes, as determined by the contributions of their various components, this finding indicates that the facial episodes of familiar faces are stronger. That is, some of the discriminative information in the episode for a familiar face is qualitatively or quantitatively superior. The question addressed here concerns which components of the facial episode vary with familiarity and thus underlie the recognition advantage for familiar faces. [We do not make particular process assumptions about recognition. For example, the components in question might contrib-

ute to perceptual fluency judgments (Jacoby & Dallas, 1981) or to contextual associations.]

We focus here on two types of potential information in facial representations: highly concrete and specific to the particular picture that was viewed, and highly abstract and verbalizable. (By the concreteness of a memory representation, we mean how narrowly it constrains the set of external objects that it represents. The more concrete the representation, the fewer objects can be mapped into or matched with it.) In regard to the first of these, the literature on face recognition leaves open the question of how concretely faces are represented in memory, and whether concrete components of facial episodes might be affected by a priori familiarity with a viewed face. Theories of face memory often assume, at the least, that an individual's idiosyncratic physical features are encoded. This assumption underlies such questions as: whether features are perceived analytically or holistically (e.g., Smith & Nielsen, 1970); whether a face's recognizability depends on the encoding of unusual or distinctive features (Winograd, 1981); and which features are most important in face perception and memory (Davies, Ellis, & Shepherd, 1977; Klatzky & Thompson, 1975; Smith & Nielsen, 1970). There is even support for the idea that face representations might be specific to a particular viewing, in the finding that small changes in a face between its initial viewing and a subsequent recognition reduce the level of performance (Galper & Hochberg, 1971; Patterson & Baddeley, 1977).

On the other hand, there are also indications that representations that commonly mediate face processing are relatively abstract, in the sense that they are not one-to-one mappings from visual displays. Harmon (1973) pointed out that pictures of faces can be recognized despite transformations that essentially eliminate specific featural information. Multidimensional scaling analyses

This research was supported by NIMH Grant 25090 to the first author. We thank Gale Martin and Maureen McGrath-McGuire for their help with the research. Requests for reprints should be addressed to the first author at: Department of Psychology, University of California, Santa Barbara, California 93106.

have indicated that most of the dimensions underlying similarity-of-face judgments are not physical features, like size of nose, but more global and inferential, like attractiveness (Hirshberg, Jones, & Haggerty, 1978). Observation of everyday life also provides anecdotal evidence for the abstractness of facial representations. We may not be able to recall even enduring attributes of people we know well, such as whether they have facial hair. We often know that a friend's face has changed since a previous viewing, but not how: Was the hair cut? Were glasses added or removed?

At an extreme of abstraction, a face might be represented with a verbalizable category, into which can be mapped a wide variety of particular viewings. Familiarity with a person provides a highly available category label, in the form of a name or personal description, but even unfamiliar faces might be represented categorically ("Burt Reynolds clone," "hairstylist type"—see Klatzky, Martin, & Kane, 1982a, 1982b).

As was stated above, we assume that the superior recognition of familiar faces derives from their memorial representations' being stronger in some component(s). But which? An obvious candidate is the categorical component, which could provide a verbal label for familiar faces. But this need not be the only basis for the familiar-face advantage in recognition. The memory episodes representing familiar faces might also be superior in their more concrete, view-specific components. In particular, this could occur if knowledge about a familiar face, gained from many different viewings, facilitated the encoding of visual details from any one instance and thus enhanced its subsequent recognizability (see Klatzky et al., 1982b). Such an effect would resemble the finding that background knowledge about a text can enhance memory for its wording (e.g., Dooling & Lachman, 1971). On the other hand, categorical knowledge about a face might lead to the assimilation of any one instance, resulting in impaired memory for view-specific attributes. Again, there are parallels in research on memory for text (e.g., Graesser, Gordon, & Sawyer, 1979). This would mean that the advantage in recognizing familiar faces must work against an inferior representation of their visual details.

Some evidence relevant to these possibilities comes from studies of meaningful-interpretation effects on memory for visual stimuli. To the extent that these studies show that applying a priori knowledge (i.e., interpretation) enhances memory for picture-like details, it would suggest that familiarity with a person's face might increase memory for a specific viewing. However, the evidence is mixed. Some studies using faces as stimuli have indeed found that interpretation facilitates subsequent recognition; however, the effect may occur not because interpretation directly enhances detail encoding but, rather, because it induces subjects to view the face as a whole (McKelvie, 1976; Winograd, 1981). Other studies have indicated that simply relating a face to a se-

mantic category does not improve memory for its particular features (Klatzky et al., 1982b). Similarly, manipulations designed to increase interpretation have been found not to affect memory for idiosyncratic details of objects in scenes (Mandler & Ritchey, 1977) and cartoon-like drawings (Rafel & Klatzky, 1978). There is also evidence that categorizing a visual stimulus can impair memory for its appearance, by producing assimilation into a more abstract categorical representation (Daniel, 1972). In short, it remains unclear whether having categorical knowledge relevant to a face, as occurs when the person is familiar, should enhance relatively concrete components of its memory episode.

In order to investigate what types of information in facial representations contribute to the superior recognition of familiar persons' faces, these studies incorporated not only a standard old/new test, in which subjects discriminated between faces seen previously in the experiment (old items) and new faces, but also tests that probed memory for certain information. Experiments 1 and 2 tested memory for orientation-specific information and details of facial features, respectively. The assumption behind these studies was that if view-specific information is critical to the familiar-face advantage in old/new recognition memory, tests that probe directly for such information should demonstrate superior performance with familiar faces. (Such a result would constitute supporting evidence for the hypothesis, although it would remain possible that superior memory for pictorial details of familiar faces is not causally related to their superiority in old/new recognition tests.) Using a somewhat different approach, Experiment 3 asked to what extent the superior recognition of familiar faces lay in the contribution of verbalizable, personal-category information to their memory episodes. Recognition was assessed for faces that subjects knew were famous persons but could not label or whose labels they had forgotten, with the assumption that performance in these conditions should suffer to the extent that the verbal information was critical to the recognition advantage of familiar faces.

EXPERIMENT 1

Experiment 1 was very straightforward. The subjects were first shown famous and nonfamous faces and then tested not only on old/new recognition, but also on recognition of the right/left orientation in which each face had originally been presented. (No face was in full frontal position.) The principal interest was in comparing performance on the orientation test with familiar and unfamiliar persons' faces. The orientation task was intended to be a measure of concrete visual information in the memory episode of a face. If famous faces derive an advantage in old/new recognition from this sort of information, performance on the orientation task should be superior for these faces. One should also find that orien-

tation-recognition performance is higher for more famous faces, leading to a positive correlation with a measure of face "famousness."

It is also interesting to compare the magnitude of the familiarity effect in orientation recognition with its magnitude in the old/new recognition test. Although one cannot specify what amount of memory for orientation-specific information would produce what magnitude of old/new recognition, gross discrepancies between the two effects would suggest that the better recognition of familiar faces is not due entirely to greater storage of picture-specific details.

Method

Subjects. The subjects were 34 students at the University of California, Santa Barbara, who participated in partial fulfillment of a course requirement.

Stimuli. The stimuli were 80 slides of faces, with no clothing being visible about the neck. Forty slides composed a pool of "famous faces." These were persons from the fields of politics and entertainment, whose names and faces were frequently in the mass media. There were 20 males and 20 females, spanning an age range from adolescent to elderly. The remaining 40 faces composed a "nonfamous" pool. The poses in the slides ranged from profile to approximately 7/8 full; none of the faces was depicted in a fully frontal pose. Each famous face was matched with a nonfamous face for approximate age, sex, and pose. Two copies of each slide were made for projection in opposite orientations.

Procedure. The experiment was run on a PDP-11/34 computer, capable of monitoring three subjects simultaneously. All subjects saw a common slide display from three adjacent booths; the slides were viewed at a distance of approximately 80 in. The slides were presented from a rear-screen projector at eye level. In each subject's booth was a CRT and a typewriter keyboard.

The subjects took part in a single session of approximately 50 min. In Phase 1, 40 faces were presented for 2 sec each, with 3 sec between them. Half the faces were famous; the others were their nonfamous counterparts. The faces within each level of fame were equally sampled from each combination of sex and direction (facing primarily right vs. left), and the distribution of poses (profile vs. partly frontal) in the sample of presented faces approximated that of the overall pool. Phase 2 was a retention interval of approximately 10 min, in which subjects read a passage on statistics and were tested on its verbatim content.

In Phase 3, the subjects took part in a test on the originally seen faces. They were shown all 80 faces in the stimulus pool. Of these, half were new, and they sampled the experimental factors as had the old faces in Phase 1. The rest were the old faces. Of these, half were shown in a direction identical to that of the original; the others were mirror images of the originals. Famous and nonfamous faces of each sex were represented equally in the identical and mirror-image conditions. Left-to-right mirror reversals were of the same number as right-to-left, and the direction of reversal was sampled as equally as possible within each combination of fame and sex.

Each face appeared on the test until all subjects had responded. The subjects were instructed to make one of three responses, indicating whether the face was a new item, an old identical item, or an old item that had been mirror reversed. Across subjects, each face occurred as both an old and a new item, was presented in both directions, and was tested as an identical and a mirror-reversed item. Presentation and test orders were randomized.

Half the subjects were assigned, in approximately alternating sessions, to a recall phase between the retention interval and the recognition test. They were requested to recall names or identifying descriptions of as many as possible of the famous people that they had seen; this took about 5 min.

Results

The principal measures were the percentages of correct responses on the two components of the test phase—old/new recognition and, for items to which subjects responded "old," orientation recognition. The old/new data are shown in Table 1 by famousness and test-item status (new; old, identical; and old, reversed). Table 2 shows the orientation responses to old items that were called old, by famousness and orientation (identical, reversed). (Orientation responses to new items called old were of little interest, since they do not reflect memory for visual appearance. The percentages of new items called old, identical and old, reversed were 8.2% and 6.6%, respectively, for famous faces and 9.6% and 9.1%, respectively, for nonfamous faces.) Also shown are a d' and a β value from a signal-detection analysis, in which a hit was defined as a correct orientation-test response to an identical item, and a false alarm was defined as calling an item identical to the original when it was actually reversed. Data are averaged over recall and nonrecall subjects, since group showed no main effects or interactions in analyses reported below.

An analysis of variance with the factors group (recall, nonrecall), fame of face in slide, and test-item status (old, new) was conducted on the old/new data. The data for old items were pooled over orientation, since the hit rates of identical and mirror items did not differ significantly for either famous or nonfamous faces; indeed, they were virtually identical. Effects are reported first with subjects treated as a random factor (data pooled over slides for each subject), second with slides treated as random (data pooled over subjects for each slide), unless stated otherwise. There were main effects of fame [$F(1,32) = 110.7$; $F(1,78) = 37.2$; both $ps < .001$] and a fame $\times$ status interaction [$F(1,32) = 32.0$; $F(1,78) = 17.0$; both $ps < .001$]. These data show that famous faces were recognized better when old, but not when new.

The second analysis of variance was performed on the

Table 1
Percent Correct Responses in Experiment 1, Old/New
Test, by Fame and Test-Item Status

	Old		New
	Identical	Reversed	
Famous Faces	89.6	91.2	85.2
Nonfamous Faces	69.5	65.9	81.2

Table 2
Performance on Orientation Test in Experiment 1

	Percent Correct Orientation Response			
	Identical Items	Reversed Items	Orienta- tion d'	Orienta- tion β
Famous Faces	85.3	48.5	1.30	.55
Nonfamous Faces	78.4	44.4	.84	1.97

percentage-correct data from the orientation test, shown in Table 2. Performance on this test was above chance (.5) for both famous and nonfamous slides [$t(33) = 7.08$ and 4.58 , respectively, $ps < .01$]. Fixed factors in the analysis were fame, group, and orientation (identical, mirror). There was a main effect of orientation, reflecting the fact that responses to identical items were more accurate [$F(1,32) = 71.4$; $F(1,78) = 99.0$; both $ps < .001$]. This undoubtedly reflects a bias to respond "identical." There was a main effect of fame when slides was the random factor [$F(1,78) = 6.44$, $p < .05$], but when subjects was random, the effect was only marginal [$F(1,32) = 2.38$, $p < .15$]. These effects suggest a slight advantage on the orientation test for famous slides. (It should be noted that this analysis selects for the more memorable items, since it includes only old items called old, which could promote an advantage for famous faces.)

To evaluate further the fame effect on the orientation test, analyses of the d' and β values, defined above, were performed. An analysis on d' with fixed factors group and fame, and with subjects treated as a random factor, revealed no significant effects. For fame, the analysis showed $F(1,32) = 2.8$, $p = .10$. However, an analysis of the same factors on β revealed that the criterion for calling an item identical in orientation was lower for the famous faces [$F(1,32) = 5.2$, $p < .05$].

Experiment 3, described in full below, provided for each face a measure of its "famousness," based on the number of subjects who called it famous and provided a name. For the 38 famous faces common to Experiments 1 and 3, this famousness score was correlated with percentage correct on each of the tests—old- and new-item recognition, orientation recognition (averaging identical and reversed items), and recall of labels. These correlations are shown in Table 3, along with relevant values from the other experiments. (Experiment 3 also provided a score for each nonfamous face, but due to renumbering of slides between experiments, this could not be applied to the present data.)

Further analyses were conducted on the recall data. The mean recall was 41.8%. Recall was divided broadly over items, with only two faces being recalled by no subjects. However, the more famous faces did tend to be recalled more (see Table 3), which means analyses con-

ditionalized on recall could reflect item-selection effects. The old/new recognition rate for recalled items was 98%. The percentage correct on the orientation test for recalled faces was 94.1% for identical items and 38.4% for mirror items, again exhibiting a response bias. The average of these two percentages was essentially identical to the 65.4% correct on orientation for the same subjects over all the famous faces, recalled and nonrecalled. Thus, there was no advantage on the orientation test for recalled items, despite their being the more famous (and better remembered).

Discussion

Experiment 1 showed three results of interest. First, the subjects were superior at recognizing famous faces. This cannot be attributed to response bias, because it is confined to the hit rate on old items; there is no advantage for rejecting new faces of familiar persons. Moreover, the correlations between face famousness and recognition are positive for old items but essentially zero for new faces, again indicating a recognition advantage for the famous faces, not merely a bias to call famous faces old. The overall pattern suggests that the advantage of famous faces derives from the episode that is encoded when a face is viewed, an advantage that can be obtained only when the face is an old item. (This is not to say a new-item advantage could not be induced by judicious selection of stimuli.)

The second result was that when subjects were able to recognize that faces had been presented, they also recognized their orientations at above-chance levels. This indicates that subjects retained some concrete information about the particular visual configuration of the faces that they had seen. However—and this is the third major finding—famous faces showed little advantage in the orientation test (despite an inherent bias toward including the more famous). The orientations of famous faces were recognized slightly better—about 5%—than those of nonfamous faces, but this difference failed to reach significance in a comparison of the d' values or percentages correct. The orientation recognition rate for faces whose labels had been recalled was no better than the overall rate, which confirms the lack of an orientation-memory advantage even for faces that subjects surely knew were

Table 3
Correlations Between Face Famousness and Test Performance

Measure of Performance	Experiment 1	Experiment 2		Experiment 3	
	Famous Faces	Famous Faces	Nonfamous Faces	Famous Faces	Nonfamous Faces
Correct Recognition, Old Faces	.45*	.61*	.40*	.34*	.56*
Correct Recognition, New Faces	.06	.15	-.05	-.02	-.16
Recall of Labels	.51*	.50*			
Recognition of Labels				.57*	
Detail Recall		.24	.09		
Orientation Recognition	.01				

* $p < .05$.

famous and remembered as having appeared during the presentation phase. Finally, the correlation between the famousness score and orientation recognition was essentially zero, again indicating that familiar faces do not provide better representations of visual detail of this sort.

EXPERIMENT 2

Experiment 2 was similar to Experiment 1 except for the nature of the concrete information that was tapped with a memory test. Here, we tested memory for two specific details in faces, regarding the configuration of the mouth and the hairline near the ears. Like the orientation information in Experiment 1, these details are highly concrete, idiosyncratic to the particular stimulus picture. The basis on which the current questions can be answered is locally specified, however, whereas the orientation question might be answered on the basis of a variety of specific details, the face's outline, or other global knowledge. The two details tested here also present an interesting contrast. The mouth configuration in a face (open/closed) is transient and in a highly informative, frequently fixated, area (Cook, 1978). It is also a strong affective signal. In contrast, the ear area is relatively stable, receives few fixations, and is not directly related to affect.

Again the test was straightforward. The subjects were first shown famous and nonfamous faces and then were tested not only on old/new recognition, but also on recall of facial details. During the recognition test, the faces were shown with either the mouth or the ear(s) covered. The subjects were asked, for faces they thought they had seen in the initial presentation, to indicate whether the mouth had been open or closed or, alternatively, whether or not the hair had obscured the ears. For faces they thought were new, the subjects were to guess answers to the same questions, which provided a baseline level of performance.

Method

Subjects. The subjects were 36 students at UCSB who participated in partial fulfillment of a course requirement.

Stimuli. The stimuli were 80 slides of faces, 76 of which had been used in Experiment 1, again 50% of each sex, and with famous and nonfamous faces matched for age. Further, these were chosen so that within each level of famousness and each sex, there were five faces in each of four classes: mouth open, mouth closed, ear exposed, and ear unexposed. The meaning of ear exposure was slightly different for the two sexes: For males, "exposed" meant that no part of the ear in the photograph was obscured by hair. For the females, it meant that the ear was not entirely covered by hair; any ear region in the photograph showed at least the earlobe or an earring. Two copies of each slide were made, one intact, and one with the relevant part of that particular face (ear or mouth region) covered by tape.

Procedure. Subjects took part in a single session of approximately 50 minutes, using the same apparatus as Experiment 1. In Phase 1, 40 slides were presented to the subject for 2 sec each, with a 4 sec interval between them. These sampled, as equally as possible, faces with each combination of fame, questioned feature (mouth or ear), answer to that question (open/exposed v.

closed/unexposed), and sex. The slides shown in this phase were always intact, not partially covered by tape. Phase 2 was again a retention interval of about 10 min, in which the subjects read a brief passage on statistics and then took a test on its verbatim content.

In the last phase, the subjects were given an old/new recognition test on all 80 faces. Half of the faces had been shown in Phase 1; the remainder were new faces that sampled the experimental factors in the same manner as the old faces did. Each face was shown with either the mouth or the ear region (both ear regions, if necessary) covered. It was accompanied by a test question on the CRT, which asked about the hidden feature. The CRT gave instructions that the subject should answer with one of four numbers, indicating if the face was old or new, and if the questioned feature was open or closed (mouth), or exposed or unexposed (ear). (The meaning of exposure was defined for the subjects.) The subjects were instructed to answer on the basis of what they remembered about old faces, and to guess the answer for new faces. Each face appeared until all subjects had responded.

Half the subjects were assigned to a recall group, as in Experiment 1. Across subjects, each face occurred as both an old and a new test item, so that the guessing rates for the subjects seeing faces as new items could be compared with the performance of other subjects who had previously seen the same faces and were responding on the basis of memory. Presentation and test orders were randomized.

Results

The principal measures were the percentage of correct responses on the two components of the test phase, old/new recognition, and detail recall (answering the question about the mouth or the ear). Percentage correct on details was computed only over items for which the subjects had responded correctly in the old/new test (old items that were hit and new items that were correctly rejected), since it is not possible to determine on what basis subjects answer a detail question about items they are misrecognizing. These data are shown in Tables 4 and 5. They are pooled over subjects who recalled and those who did not, because the two groups were very

Table 4
Percent Correct Responses in Experiment 2, Old/New Test,
by Fame, Questioned Detail, and Test-Item Status

	Mouth Question		Ear Question	
	Old Items	New Items	Old Items	New Items
Famous Faces	87.2	87.7	86.9	83.4
Nonfamous Faces	57.4	83.4	65.4	85.9

Table 5
Percent Correct Responses in Experiment 2, Detail Recall Test,
by Fame, Questioned Detail, and Test-Item Status

	Mouth Question		Ear Question	
	Hit Old Items	Rejected New Items	Hit Old Items	Rejected New Items
Famous Faces	65.2	60.2	60.4	64.7
Nonfamous Faces	55.4	51.8	75.9	73.4

similar and the group variable produced no main effects or interactions in the analyses of variance reported below.

An analysis of variance was conducted on the old/new recognition data with fixed factors group (recall, nonrecall), fame of face in slide, questioned feature (mouth, ear), and test-item status (old, new). The F values for the two random factors—subjects (averaging data over slides) and slides (averaging over subjects)—will be reported in succession. There was a main effect of fame [$F(1,34) = 57.5$; $F(1,76) = 38.3$; both $ps < .001$], with famous faces being remembered better. The fame $\times$ status interaction was significant [$F(1,34) = 35.9$; $F(1,76) = 25.1$; both $ps < .001$], indicating that the famous faces were recognized better when old, but that performance on new items was equivalent for famous and nonfamous faces. The main effect of item status was also significant [$F(1,34) = 11.5$; $F(1,76) = 16.8$; both $ps < .01$]. Finally, the fame $\times$ question interaction was significant when subjects was treated as a random factor [$F(1,34) = 5.8$, $p < .05$], but not when slides was [$F(1,76) = 1.6$]. The pattern of the interaction indicates that obscuring the mouth of nonfamous slides led to worse performance than obscuring the ear, but the two questions were equivalent for the famous slides. The failure to find this interaction with slides treated as a random factor indicates that it does not generalize to other samples of slides.

Although Experiments 1 and 2 were run at different times and were not intended to be directly compared, it is interesting to note that performance on the old/new recognition test was very similar across experiments, despite the obscuring of facial details in Experiment 2. For old and new items, and within each level of famousness, the mean percentage correct in Experiment 1 was compared with the means for mouth- and ear-obscured faces in Experiment 2. For new items, there were no significant differences. Comparisons of the hit rates revealed one significant difference: The performance in Experiment 2 with nonfamous, mouth-obscured faces was lower than the overall hit rate in Experiment 1 [$t(68) = 2.14$, $p < .05$].

A second analysis of variance was conducted on the data from the detail test. Fixed factors were again group, fame, questioned feature, and test-item status. There were significant effects of question and the fame $\times$ question interaction when subjects was treated as a random factor [$F(1,34) = 36.2$, $p < .001$, and $F(1,34) = 31.0$, $p < .001$, respectively], but not when slides was treated as random [$F(1,76) = 2.6$, $p < .15$, and $F(1,76) = 3.6$, $p < .10$]. Most importantly, the main effects of fame and test-item status were not significant (all $Fs < 1$). This indicates that the subjects were no better at recalling details of faces they had seen than at guessing details of faces they had not seen. This occurred despite the fact that only items that were correctly recognized as old/new were included, which would select for the more famous and better remembered old items. The signifi-

cant effects of this analysis suggest that guesses on the ear question were better than on the mouth for nonfamous faces, but not for famous faces, although this only marginally generalizes over slides.

Other data concern the performance on recalled faces, within the recall group. The subjects recalled labels for an average of 38% of old faces. Again, the more famous faces tended to be recalled more, but recall was broadly distributed over faces, with all but three faces being recalled at least once. The hit rate on the old/new test was 100% for recalled faces. The percentage correct on the detail question for recalled faces was 72% on the mouth question and 77% on the ear. The percentage correct recall of details (pooled over the two questions) for faces whose labels were recalled was above the guessing rate for famous faces [$t(17) = 1.79$] and was also above the percentage correct on the detail question for all old famous faces combined (recalled and nonrecalled) [$t(17) = 2.03$, $ps < .05$, one-tailed]. Thus, details of recalled famous faces were remembered at a better-than-chance rate and better than those of famous faces in general, although these effects were not strong.

Table 3 again shows correlations of interest. Over slides, the famousness scores from Experiment 3 were correlated with percentage recall of faces by label, percentage correct recognition of old and new faces, and percentage correct on the detail-recall test (corrected for guessing by subtracting the rates for new items from old).

Discussion

Experiment 2 again showed a clear advantage for famous faces in episodic recognition, such that "old" famous faces were better recognized as having recently been seen, but famousness did not facilitate detecting that a face was a new item. The correlations between percentage correct and famousness also show famousness to be directly related to percentage correct for old items, but not for new items. (See Table 3.)

Next we turn to the principal question of this study, which was whether the better recognition of famous faces could be attributed to enhanced representation of details of the particular pictures that had been seen. The relevant data, which come from the detail test, fail to support this hypothesis. In general, memory for details was very poor, and was no better for famous faces than for nonfamous faces. Overall, the details of neither famous faces nor nonfamous faces were recalled at better-than-guessing levels. In addition to the lack of an overall effect of fame, the famousness of faces did not correlate with performance on recall of their details (Table 3).

Given the generally low levels of detail recall, it is of special interest that detail performance was above chance when measured for famous faces that had been verbally recalled. This might be explained by assuming that the verbally recalled faces are the more famous, and that there must be a stronger detail representation for the most famous faces. However, as described in the re-

sults, almost every famous face was recalled. A more plausible explanation for the finding of better detail performance on recalled faces would attribute it to a selection bias for the better remembered faces, rather than for the more famous faces. Specifically, if the recalled faces tend to be those with the stronger representations of view-specific information, one would see better detail memory for recalled faces. This explanation suggests that recall is not exclusively based on memory for verbal labels, independent of visual information. (Interestingly, this suggestion is supported by positive correlations over subjects between percentage recall (of famous faces) and old/new recognition accuracy for nonfamous faces, which are generally unlabelable—.67 in this experiment, .51 in Experiment 1.)

A final point that merits discussion is the effect of which question, regarding the mouth or the ear, was asked. For the nonfamous faces, the answer to the ear question appears to have been easier to guess, and old/new recognition of faces with the ear covered (or nothing obscured, as in Experiment 1) was higher than when the mouth was covered. For the famous faces, the questions were equivalent. (These trends generalized over subjects but not over slides.) Taken together, these findings suggest a somewhat speculative argument: The ability to recognize a nonfamous face may depend on filling in, by guessing and imaging, missing details. This is easier to do for the ear than for the mouth, and recognition is therefore higher when the ear is covered. Famous faces may be recognized without such a fill-in process, perhaps because greater categorical contributions to the facial episode compensate for the lack of detail.

EXPERIMENT 3

Experiments 1 and 2 probed facial episodes at a concrete level of representation and found little evidence that this level is the basis for the advantage in recognizing familiar faces. Experiment 3 therefore probed at a more abstract level, asking whether familiar faces are more recognizable solely because subjects generate category labels at the time of viewing and store the categorical information with the corresponding facial episodes.

The subjects initially viewed faces, indicated whether the faces were famous, and generated labels for as many of the famous ones as they could. They were then given old/new recognition tests on the faces and on the labels of the famous persons. If the advantage of recognizing famous faces lies solely in memory for their labels, the data should show that performance on faces whose labels were not remembered—or for which labels were never generated—was worse than on those whose labels were remembered. A stronger claim is that the data should show that performance on the famous faces whose labels were not remembered or generated descended to the level of performance on nonfamous faces. However, in the present task, there were substantial en-

coding differences between famous and nonfamous faces (e.g., attempts to generate labels for the former); therefore, comparisons of this sort must be made somewhat tentatively.

Method

Subjects. The subjects were 27 students at the University of California at Santa Barbara who were participating in partial fulfillment of a course requirement.

Stimuli. These were the same slides as in Experiment 2.

Procedure. Each subject took part in two sessions, 2 weeks apart. In the first (Phase 1), they saw 40 slides of faces, half of which were famous and half, nonfamous, in random order, for 5 sec each. The subjects were instructed to look at the slide throughout the interval. When the slide terminated, the subjects were to indicate, by writing F or NF on a response sheet, whether they thought it was a famous or a nonfamous person. In addition, the subjects supplied a name or, if not a name, a category label to each F face whenever possible. They were given examples like "first president of U.S." and "played Moses in a movie" for the labels. The subjects pressed a key on their keyboard when they had responded, and the next slide appeared when all had responded.

In the second session, the subjects took part in two tests. The first (Phase 2) was an old/new recognition test on verbal labels. On the CRT screen were displayed verbal labels for each of the 40 famous persons in the stimulus pool, in random order. Each consisted of the person's name and a brief category description. These descriptions had been derived from pilot data on label generation and corresponded to descriptions provided by the present subjects as well. (An example is "Rosalynn Carter—wife of Jimmy, former first lady.") As each label appeared, the subjects were to decide if they had seen the labeled person's face in Session 1. They responded "old" or "new," accordingly, and also indicated one of three degrees of confidence in their response, by pressing marked keys on the keyboard. They were also given the option of pressing another key to indicate that they were unfamiliar with the labeled person. The next label appeared after all subjects had responded.

Immediately following the test on labels, the subjects took part in an old/new recognition test on all 80 faces in the stimulus pool, in random order (Phase 3). As each face was exposed, the subjects indicated an old or new response with one of three degrees of confidence, by pressing one of the marked keys, as in the previous test. (The additional key to indicate unfamiliarity with the person was not used.) The next face appeared when all subjects had responded. Across subjects, all faces were used as both old and new items.

Results

Phase 1 yielded data regarding the subjects' identification of faces as famous or nonfamous and their labeling of them. Of the famous faces, a mean over subjects of 69.5% were identified as famous and given a correct label. Another 17.9% were identified as famous but not correctly labeled; of these, 2.8% were actually mislabeled (more often than not, with uncertainty indicated by the subject), and the rest were not labeled at all. An average of 12.6% were called nonfamous. (In comparison, the subjects failed to recognize only 3% of the famous persons from the names and labels in Phase 2. The difficulty of recognizing famous faces from single exposures was noted in Yarmey, 1971.) Of the nonfamous faces, 72.3% were identified as nonfamous. Of the remainder, 3.3% were called famous and given an

incorrect label (usually with uncertainty indicated), and the rest were simply called famous. (The labeling of nonfamous faces as famous was widely distributed over slides and will be discussed further below.)

For each slide, a "famousness score" was calculated. Each nonfamous slide was given 4 points if it was called famous in Phase 1. A famous slide was given 4 points if it was both identified as famous and labeled correctly and 2 points if it was merely called famous. The scores were averaged over subjects for each slide. The mean score over slides was 3.1 for the famous slides and 1.1 for the nonfamous, and the standard deviations were .84 and .51, respectively.

The principal data for Phases 2 and 3 (recognition of labels and faces) were percentage correct responses and confidence ratings. The two measures gave very similar results, so analyses of percentage correct will be reported, and discrepancies with the confidence analyses will be noted when relevant. (Due to small error rates, especially for famous faces, confidence ratings were analyzed only for correct responses.) The mean percentage correct and the mean confidence in correct responses are given in Table 6, broken down by old versus new items, Phase 1 outcome (for old items), Phase 2 outcome (for famous items), stimulus type (name vs. face) and famousness.

The design of this experiment did not lend itself to any single analysis of variance. Rather, analyses were conducted on subdesigns to address particular questions. F tests will be reported first with subjects treated as a random factor and then with slides treated as random. The number of observations entered into these analyses varied, since only subjects (or, for analyses with slides treated as a random factor, only slides) who contributed data to all levels of repeated-measures factors were included. It should also be noted that some comparisons use pools of stimuli biased toward including the more famous and/or more memorable items. When relevant, such biases will be discussed.

The first analysis tested whether famous faces were recognized better than nonfamous faces. The fixed factors were fame and oldness (old items versus new items on the recognition test). The effect of fame was non-significant ($F_s < 1$), reflecting the less than 1% difference on the average percentage correct for famous and nonfamous faces. The effect of oldness was significant [$F(1,26) = 9.5$; $F(1,78) = 9.9$; both $ps < .01$], and more importantly, there was a significant fame $\times$ oldness interaction [$F(1,26) = 46.9$; $F(1,78) = 52.9$; both $ps < .001$]. Performance on old items was better with famous faces, but on new items, it was better with nonfamous faces. This may have been due to the test on names in Phase 2, which could well have increased the false-alarm rate to the faces in Phase 3. (This is suggested by the tendency to false alarm to a face after a false alarm to its name; see below.) Although the percentage correct data did not show famous faces to be remembered better, there was a significantly higher d' for the famous faces ($d' = 2.34$) than for the nonfamous ($d' = 1.87$) [$t(26) = 3.26$, $p < .01$], and the confidence ratings for correct responses were also higher for the famous faces [$F(1,26) = 23.9$; $F(1,78) = 24.9$; both $ps < .001$]. Thus, there is evidence for better recognition of famous faces despite the disadvantage induced by the test on names.

The second question concerned the effects of the accuracy of slide identification in Phase 1 on recognition of old faces in Phase 3. The fixed factors were famousness and accuracy of Phase 1 identification (identification was called accurate if famous faces were called famous, whether or not a label was given, and if nonfamous faces were called nonfamous). The effect of famousness was significant [$F(1,22) = 18.0$; $F(1,60) = 8.9$; both $ps < .01$]. The interaction was also significant [$F(1,22) = 25.2$; $F(1,60) = 8.2$; both $ps < .01$]. This indicates that correct identification of famousness was associated with higher hit rates for famous faces (as is consistent with a bias for correct identification of the

Table 6
Percent Correct (and Confidence in Correct Responses) on Phase 2 and Phase 3 Tests, by
Item Oldness, Phase 1 Outcome (Old Items Only), and Fame

	Old Items						New Items	
	Phase 1 Outcome							
	Face Called Famous, Labeled		Face Called Famous Only		Face Called Nonfamous			
	PC	CR	PC	CR	PC	CR	PC	CR
			Phase 2 Test of Famous-Label Recognition					
	91.7	2.76	49.7	2.11	41.3	2.22	64.7	2.01
			Phase 3 Test of Face Recognition					
Famous Faces								
Overall	96.6	2.86	97.5	2.54	88.5	2.46	65.8	2.00
Phase 2 Miss/False Alarm*	89.6	2.57	98.8	2.50	89.4	2.18	47.9	1.69
Nonfamous Faces			84.8	2.40	70.3	2.24	84.8	2.05

Note—PC = percent correct; CR = confidence (in correct responses) rating. *That is, given miss or false alarm to label in Phase 2.

more famous) but with lower hit rates for nonfamous faces. When famous faces were called nonfamous in Phase 1, their recognition was comparable to that of nonfamous faces called famous, but higher than the recognition rate of nonfamous faces called nonfamous [over subjects, $t(23) = 2.85, p < .05$].

A third analysis considered the famous category only. It compared the effects of the accuracy of Phase 1 slide identification on subsequent recognition of old labels (Phase 2) and faces (Phase 3). The fixed factors were Phase 1 outcome and stimulus type (label–Phase 2 test; face–Phase 3 test). The former included three levels, corresponding to the three headings at the top of Table 6. The effect of Phase 1 outcome was significant [$F(2,44) = 25.4$; $F(2,34) = 11.8$; both $ps < .001$], as was that of stimulus type [$F(1,22) = 53.4$; $F(1,17) = 39.3$; both $ps < .001$], with label recognition worse than face recognition. The interaction was also significant [$F(2,44) = 16.2$; $F(2,34) = 5.5$; both $ps < .01$], which reflects the fact that accurate label recognition depended on the subject's having initially given a label in Phase 1, whereas face recognition did not. (The confidence in correct responses did not show this interaction, however; $F_s \approx 1$.) These effects of Phase 1 outcome are generally consistent with the argument that the more famous faces are identified as famous and labeled, and are thus better recognized in both Phase 2 and Phase 3.

Tests were conducted to determine whether the probability of correctly recognizing a face as old depended on correctly recognizing its name. The first compared the overall hit rate for famous faces with the hit rate for faces whose names had been missed in Phase 2. When only faces that had been labeled in Phase 1 were considered, there were 12 subjects who supplied data to the comparison (the remainder failed to miss names in Phase 2 that they had supplied in Phase 1). Neither the t test on percentage correct [$t(11) = 1.02$] nor that on confidence in correct responses [$t(11) = 1.36$] was significant. When all famous faces were considered, ignoring Phase 1, the overall hit rate was again not significantly different from the hit rate for faces whose names had been missed [$t(23) = 1.19$], which was significantly higher than the hit rate for nonfamous faces [$t(23) = 4.59, p < .01$]. These effects run counter to what one would expect on the basis of item-selection bias, for which those faces whose names were missed should be the less memorable. In contrast to the independence of recognizing old names and faces, there were significantly fewer correct rejections of new faces after false alarms to names than there were correct rejections overall [$t(25) = 3.60, p < .01$].

Finally, the famousness score for the slides was correlated with the percentage correct recognition of old and new faces of famous and nonfamous persons, and with label recognition (averaging over old and new items) for famous persons. These correlations appear in Table 3.

Discussion

Experiment 3 again indicates that familiar faces have a recognition advantage. The advantage was not apparent in a comparison of percentage correct for famous and nonfamous faces, probably because the test on labels increased the false-alarm rate to famous faces in the subsequent test; however, it did emerge in comparisons of d' values and confidence in correct responses. In addition, there was a positive correlation between the famousness measure and recognition, even for the nonfamous faces.

To the extent that the famousness measure reflects label availability for the famous faces, the positive correlation between that measure and recognition of famous faces indicates that verbalizable information aids recognition. However, the principal question addressed in this experiment was not whether memory for categorical labels produces an advantage in recognition of famous faces, but, rather, whether it is the sole basis for the advantage. The answer appears to be negative. There are three indications in these data that famous faces have an advantage in recognizability that is not based on memory for labels.

First, face recognition is high even for famous faces for which label recognition is missing. This holds even when subjects were initially able to supply the label, as well as when they knew the face was famous but could not supply a label. Moreover, it runs counter to what would be predicted by item-selection biases. It might be argued that recognition of all famous faces is at ceiling level, so that a decrement in face recognition when label recognition fails simply cannot be observed. There is obviously some merit in this argument, but it should be noted in rebuttal that there is also little indication that confidence in face recognition suffers when a name is first generated and then forgotten.¹

It should also be noted that the strong recognition of famous faces whose labels were missed goes against a general trend to repeat the response made in the test on labels, when the test on faces is given. This trend can be observed in the higher false-alarm rate to new faces whose labels were erroneously called old than to those whose labels were correctly called new.

As a final point, the high proportion of hits on faces, given misses on labels, cannot be easily dismissed as random variability over consecutive tests. For comparison, consider the percentage of famous persons whose labels were hit that were missed in the subsequent face-recognition test. This was only 2.8% for faces that had been labeled in Phase 1, and 2.5% overall.

The second indication that recognizing famous faces does not depend on their labels comes from the differences in the effects of the outcome of Phase 1 on the subsequent tests of labels and faces. Whereas initial labeling appears to have been critical to recognition of the labels in Phase 2, it was not critical to recognition of faces in Phase 3. Faces that had been identified as famous and labeled were no more accurately recognized than were faces that had been identified as famous but

not labeled. Again, this runs against an item-selection bias, which should lead to the more famous—and memorable—faces being labeled; however, an effect here could be masked by ceiling effects.

The third indication that the recognition advantage for faces of familiar persons does not rely on labeling comes from examination of the data on nonfamous faces. When these faces were called famous, they were recognized better, even though labels were not supplied. Moreover, there was a strong positive correlation between the famousness score of nonfamous faces and their recognizability. Just why nonfamous faces were called famous is unclear. Examination of those faces that give the highest famousness score was not illuminating in this regard, since it reveals variations in age, sex, and source of picture. In some cases, the face may have been familiar from advertisements or news articles, or the photograph may have been mistaken for a familiar person. Another, somewhat speculative, idea is that these faces fit a variety of *a priori* categories for famous persons. For example, the pool of nonfamous faces included models, which fit the criterion that famous people are often attractive; it included successful business persons, which fit other aspects of appearance associated with fame, such as a well-groomed look. Whatever its cause, the effect on recognition of “looking famous” again indicates a basis other than verbalizability for the famous face advantage.

GENERAL DISCUSSION

This series of experiments began with the assumption that famous faces derive an advantage in old/new recognition because some component(s) of their memorial representation is (are) superior. The studies focused on components of two types: concrete, stimulus-specific details and more abstract information that can be captured in a categorical label.

In general, the studies confirmed that famous faces have an old/new recognition advantage (although this was obscured to a degree in Experiment 3 by an intervening test on names). However, it does not appear to derive from a more robust representation of details idiosyncratic to the particular picture viewed. Experiments 1 and 2 tested this possibility by probing explicitly for details of various sorts. With two probes—of memory for the configuration of the mouth and ears—not only was there no advantage for faces of familiar persons, but there was also little evidence that information of this sort is accessible in memory, at least with a recall test. With another measure—of memory for orientation-specific information—there was above-chance performance, but virtually no advantage for the familiar faces, certainly not enough to make orientation-specific information a likely candidate for the old/new recognition advantage those faces enjoy.

It is possible that other types of tests might reveal stronger evidence for the retention of idiosyncratic

pictorial details, and possibly even of an advantage for familiar faces in such retention. In particular, Bartlett, Till, and Levy (1980) distinguished between voluntary and involuntary forms of recognition, the former measured by overt reports and the latter, by comparing recognition rates between two conditions that differ in the extent to which they reinstate the original stimulus. Although the present focus was on voluntary reports of the content of memory, there are some data here that provide measures of involuntary recognition. One can examine the results of Experiment 1 for the effects of reversing a test picture on recognition, with the assumption that performance should suffer from reversal if the memory representation is highly orientation specific. There is no indication whatever of the retention of orientation information from this comparison: Hit rates on reversed and nonreversed items are virtually identical, for famous and nonfamous faces.

Similarly, one can investigate the importance of featural information such as that probed in Experiment 2, by comparing the recognition of old items when features are obscured (as in Experiment 2) with performance on the same items without covering the features (Experiment 1). This comparison (Tables 1 and 4) reveals very similar hit rates on famous faces, whether a feature was covered or not, but a decrement when the mouth is covered on a nonfamous face. The latter finding suggests that the mouth may critically contribute to recognizing an unfamiliar face. However, it still does not directly address whether the representation of the mouth is superior in a nonfamous (or famous) face. The present, more direct, investigations suggest that it is not.

Can verbalizable information entirely account for the famousness effect in old/new recognition? The importance of verbal information is suggested by the virtually perfect recognition rate of faces whose labels were recalled. However, it does not appear that the entire advantage for familiar faces can be attributed to the retention of category labels. Experiment 3 showed, somewhat surprisingly, that forgetting labels had little effect on the recognition rate for familiar faces, and that “looking famous” affected recognition even of faces for which labels were unavailable. Moreover, face labelability profoundly affected label recognition, but not face recognition, suggesting a disassociation between the episodic representations of category labels and faces.

In short, the familiarity effect on the recognition of faces in an episodic memory task does not appear to derive from remembering idiosyncratic visual details of the particular configurations that were seen, and it does not derive entirely from memory for a verbal label for the familiar, but not the unfamiliar, faces. This leaves, by default, components that are both less concrete than the visual details tested here and less abstract than a category label. Further research should be directed to describing the nature of these components. Along that line, Klatzky et al. (1982a) investigated an advantage in

recognizing faces that were matched with a rather different sort of familiar category—occupational stereotypes. Underlying this familiarity effect appeared to be a relatively abstract representation, incorporating physical and affective attributes common to members of a stereotyped class, but not incorporating visual configurations particular to any one face.

The potentially abstract basis for recognition suggests one reason why memory for faces is fallible, making the present research relevant to eyewitness identification. By definition, abstract representations allow matches with many objects from the external world. The use of such representations will limit how finely faces can be discriminated in memory. If an abstract facial representation conveys only stereotyped attributes such as race, age, or even occupation, discrimination errors are especially likely. The lower recognition rates for members of races other than one's own may reflect the use of such representations (e.g., Cross, Cross, & Daly, 1971). On the other hand, to speak of abstract components of facial episodes does not mean that they must be very abstract. A very concrete representation was tested here, and the finding that faces are remembered at something less than this level of concreteness does not mean that facial episodes cannot allow fine discriminations.

REFERENCES

- BARTLETT, J. C., TILL, R. E., & LEVY, J. C. Retrieval characteristics of complex pictures: Effects of verbal encoding. *Journal of Verbal Learning and Verbal Behavior*, 1980, 19, 430-449.
- COOK, M. Eye movements during recognition of faces. In M. M. Greenberg, P. E. Morris, & R. N. Sykes (Eds.), *Practical aspects of memory*. London: Academic Press, 1978.
- CROSS, J. F., CROSS, J., & DALY, J. Sex, race, age and beauty as factors in recognition of faces. *Perception & Psychophysics*, 1971, 10, 393-396.
- DANIEL, T. C. Nature of the effect of verbal labels on recognition memory for form. *Journal of Experimental psychology*, 1972, 96, 152-157.
- DAVIES, G. M., ELLIS, H. D., & SHEPHERD, J. W. Cue saliency in faces as assessed by the Photofit technique. *Perception*, 1977, 6, 263-269.
- DOOLING, D. J., & LACHMAN, R. Effects of comprehension on retention of prose. *Journal of Experimental Psychology*, 1971, 88, 216-222.
- ELLIS, H. D., SHEPHERD, J. W., & DAVIES, G. M. Identification of familiar and unfamiliar faces from internal and external features: Some implications for theories of face recognition. *Perception*, 1979, 8, 431-439.
- GALPER, R. E., & HOCHBERG, J. Recognition memory for photographs of faces. *American Journal of Psychology*, 1971, 84, 351-354.
- GIBSON, J. J., & GIBSON, E. J. Perceptual learning: Differentiation or enrichment? *Psychological Review*, 1955, 62, 32-41.
- GOLDSTEIN, A. G. The fallibility of the eyewitness: Psychological evidence. In B. D. Sales (Ed.), *Psychology in the legal process*. New York: Spectrum, 1977.
- GRAESSER, A. C., GORDON, S. E., & SAWYER, J. D. Recognition memory for typical and atypical actions in scripted activities: Tests of a script pointer plus tag hypothesis. *Journal of Verbal Learning and Verbal Behavior*, 1979, 18, 319-332.
- HARMON, L. D. The recognition of faces. *Scientific American*, 1973, 229(5), 71-82.
- HIRSCHBERG, N., JONES, L. E., & HAGGERTY, M. What's in a face? Individual differences in face perception. *Journal of Research in Personality*, 1978, 12, 488-499.
- JACOBY, L. L., & DALLAS, M. On the relationship between autobiographical memory and perceptual learning. *Journal of Experimental Psychology: General*, 1981, 110, 306-340.
- KINTSCH, W. *The representation of meaning in memory*. Hillsdale, N.J.: Erlbaum, 1974.
- KLATZKY, R. L., MARTIN, G. L., & KANE, R. A. Influence of social-category activation on processing of visual information. *Social Cognition*, 1982, 1, 95-109. (a)
- KLATZKY, R. L., MARTIN, G. L., & KANE, R. A. Semantic interpretation effects on memory for faces. *Memory & Cognition*, 1982, 10, 195-206. (b)
- KLATZKY, R. L., & THOMPSON, A. Integration of features in comparing multifeature stimuli. *Perception & Psychophysics*, 1975, 18, 428-432.
- MANDLER, J. M., & RITCHEY, G. H. Long-term memory for pictures. *Journal of Experimental Psychology: Human Learning and Memory*, 1977, 3, 386-396.
- McKELVIE, S. J. The effects of verbal labeling on recognition memory for schematic faces. *Quarterly Journal of Experimental Psychology*, 1976, 28, 459-474.
- PATTERSON, K. E., & BADDELEY, A. D. When face recognition fails. *Journal of Experimental Psychology: Human Learning and Memory*, 1977, 3, 406-417.
- RAFANEL, K. J., & KLATZKY, R. L. Meaningful-interpretation effects on codes of nonsense pictures. *Journal of Experimental Psychology: Human Learning and Memory*, 1978, 4, 631-646.
- SMITH, E. E., & NIELSEN, G. D. Representations and retrieval processes in short-term memory: Recognition and recall of faces. *Journal of Experimental Psychology*, 1970, 85, 397-405.
- WINOGRAD, E. Elaboration and distinctiveness in memory for faces. *Journal of Experimental Psychology: Human Learning and Memory*, 1981, 7, 181-190.
- YARMEY, A. D. Recognition memory for familiar "public" faces: Effects of orientation and delay. *Psychonomic Science*, 1971, 24, 286-288.

NOTE

1. In a pilot version of this study, in which parts of faces were obscured during the test and recognition rates were lower, the results were similar. The average correct recognition of faces that were initially identified as famous, and whose labels were missed, was 81.8%, as compared with 88.4% for those whose labels were hit. The hit rate for nonfamous faces was only 40.5%. The percentages correct on new faces were not nearly as discrepant between famous and nonfamous—61% and 89%, respectively.

(Manuscript received April 13, 1983;
revision accepted for publication September 13, 1983.)