Order Number 8801751

The effects of race on face recognition

Regan, Susan, Ph.D.
City University of New York, 1987

U·M·I
300 N. Zeeb Rd.
Ann Arbor, MI 48106

THE EFFECTS OF RACE ON FACE RECOGNITION

by

SUSAN REGAN

A dissertation submitted to the Graduate Faculty in Psychology in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York.

1987

This manuscript has been read and accepted for the Graduate Faculty in Psychology in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

9/22/87 — Robert Buckhout
Date — Chair of Examining Committee

9/24/87 — Herbert D. Saltzstein
Date — Executive Officer

Neil Macmillan

Don Scarborough

Arthur Reber

Michael Siegel

Supervisory Committee

The City University of New York

Abstract

THE EFFECTS OF RACE ON FACE RECOGNTION

by

Susan Regan

Adviser: Professor Robert Buckhout

Three experiments explored the effects of race on face recognition. In Experiment 1, white and black subjects saw slides of own- and other-race faces in upright and inverted orientations. All subjects showed an "own-race effect:" superior performance on own-race faces, for upright faces. White, but not black, subjects showed the same effect for inverted faces. All subjects performed poorly on inverted black faces. False alarm rates were highest where sensitivity was lowest: for other-race and inverted faces. The results fail to support the notion that lack of experience with other-race faces is responsible for the own-race effect. In Experiment 2, white and black subjects judged whether two simultaneously-presented slides of the same face were in the same left-right orientation. Performance was higher on white faces, but there was no own-race effect. Subjects made more <u>Same</u> than <u>Different</u> judgments, especially for black faces. The results failed to support the notion

that the own-race effect is due to impoverished perception of other-race faces and suggested that the poor performance on inverted black faces might be due to their greater symmetry. Experiment 3 differed from Experiment 2 only in that the judgment was made from memory of a standard. Performance was poor, but among those scoring above chance, all of the findings of Experiment 2 were replicated, showing that the difference between black and white faces persists in a memory task. However, there was no correlation between percent correct (for a given face) in the left-right orientation judgment and inverted recognition, and therefore no evidence of a connection between the poor performance on black faces in Experiments 1 and 2. Race emerged as a variable that effects recognition performance reliably but somewhat weakly. It was concluded that motivation is probably involved in face recognition, and that the own-race effect reflects memory rather than perception of faces.

Table of Contents

List of Tables

Suppose you are sitting in a subway car. You haven't brought any reading material with you, and so you look idly at the advertisements and at the other people in the car. Your gaze settles briefly on a young woman as the train pulls into the station. The doors open and she jumps up and quickly leaves the car. Would you recognize her if you saw her again?

Contrary to the expectations of many lawyers and jurors, research on face recognition will never be able to provide a definite answer to such a question. This work has, however, identified several factors that influence our ability to recognize faces. Some of them are obvious and affect visual memory for any kind of stimulus -- the better the lighting and the more time you spent looking at the young woman, the more likely you are to remember her (Laughery, Alexander, & Lane, 1971) -- but others are not. Context plays a role: you are more likely to recognize her if you see her in another subway car than if she walks into your office (Watkins, Ho, & Tulving, 1976; Winograd & Rivers-Bulkeley, 1977). We know from the work of Yin (1969) on inversion of visual stimuli that if she were upside down during the encounter you would be less likely to recognize her face later on (although she might make more of an impression on you). One factor that appears to be irrelevant is whether or not you try to remember her. Deliberate attempts to remember a face are no more successful than casual observation (Brigham & Barkowitz, 1978; Devine & Malpass, 1985). No conclusions can be drawn about the likelihood of a face being recognized without a consideration of who is to recognize it. If you were under the age of 17, you would be less likely to remember the woman than adult passengers in that subway car (Carey, Diamond, & Woods, 1980). The bulk of the evidence suggests that you are in a better position to recognize her if you are a woman (Shepherd, 1981). Perhaps the most crucial factor is race, but not simply your race or hers. It is the combination that counts. You are more likely to recognize her if she is of your own race.

The fact that we are better at recognizing the faces of people of our own race than those of other races, the "own-race effect", is reflected in such remarks as "They all look alike to me." This longstanding bit of folk wisdom has lately been corroborated by a series of laboratory findings. In most of this research the same simple procedure is used. The races in question are almost always black and white. The stimulus faces are usually obtained from yearbook pictures. Each picture is rephotographed and mounted as a black and white slide. The subject is shown a subset of these faces, designated the study set, which contains an equal number of faces from each race. Each face in the study set is presented individually for the same fixed period of time. After a retention interval (which ranges in length from minutes to days), a recognition test is administered. The subject is presented with a set of faces, the test set, that includes the study set plus one or two times as many novel faces. The subject's task is to report which of the faces in the test set were part of the study set. Typically, a signal detection analysis is performed on the data.

All of the published attempts to obtain the own-race effect have been successful, but the effect does not always take the same form. Devine and Malpass (1985), Brigham and Barkowitz (1978), and Galper (1973) all found a "symmetrical" own-race effect, i.e. the other-race deficit was of equal size for both their white and black subjects. Asymmetrical own-race effects have also been demonstrated. Barkowitz and Brigham (1982), Cross, Cross, and Daly (1971), and Malpass and Kravitz (1969) all found evidence of an own-race effect for whites but not for blacks. Brigham and Williamson (1979) found the opposite; their white subjects had no other-race deficit but their black subjects had a large one. In the experiments of Ellis and Deregowski (1981) and Malpass, Lavigueur, and Weldon (1973), all subjects showed a sizable other-race deficit but that shown by the black subjects was significantly larger. Chance, Turner, and Goldstein (1982) and Elliott, Wills and Goldstein (1973) have found an other-race deficit for Asian faces in their white subjects. Because they did not employ Asian subjects we do not know whether the effect was symmetrical.

The lack of uniformity in the data is puzzling. It is tempting to attribute the variability in the symmetry of the effect to the stimuli -- because no two faces are exactly alike, no two groups of faces can be equivalent -- but there is evidence that the stimuli are not at fault. Malpass and Kravitz (1969) and Malpass et al. (1973) used the same stimuli and produced the opposite results: in the earlier study there was no other-race deficit in blacks; in the later one the own-race effect was stronger in blacks than in whites. The differences in the results are probably due to the subjects rather than to the stimuli. Whatever factors are responsible for the effect are not uniformly distributed among subjects, black or white.

Apparently the own-race effect is fairly easy to replicate, but it has been difficult to explain. Five hypotheses have been put forward to account for the poor other-race recognition usually found:

1. The "they really <u>do</u> all look alike" hypothesis: the difficulty in discriminating between other-race faces is caused by a lack of variability in the (non-white) faces themselves.

2. The "depth-of-processing" hypothesis: own-race faces are remembered better because they are processed more deeply (see Craik & Lockhart, 1972) than other-race faces. Because we have fewer associations with other-race faces, we encode them in terms of "shallow" characteristics -- facial features and broad categories like race and sex.

3. The prejudice hypothesis: we have difficulty remembering other-race faces because racial prejudice

prevents us from giving them the same degree of attention that we give own-race faces. Two versions of this hypothesis are possible. One suggests that if we were forced to pay attention to other-race faces, we would remember them as well as we remember own-race faces. The other claims that our history of inattention has prevented us from learning how to perceive and remember other-race faces. Increasing attention to a particular face cannot overcome this handicap.

4. The motivation hypothesis: because our social interaction is primarily with members of our own race, we are more interested in remembering own-race faces. It is possible to distinguish two versions on the same basis as with the prejudice hypothesis: one claims that motivation has its impact while we are viewing a particular face; the other claims that, without motivation, we will never learn to remember other-race faces. The second predicts that temporarily increasing motivation will not affect memory for other-race faces.

5. The experience hypothesis: our poor other-race recognition is the result of the lack of experience we have in perceiving and discriminating between them. This hypothesis implies that we gain knowledge about faces that is race-specific, that we learn to attend to aspects of faces that are helpful in discriminating between members of our own race but that are generally (but not necessarily) less useful with other-race faces. At present, there is

no convincing evidence in support of any of these hypotheses.

The "They Really Do All Look Alike" Hypothesis

This suggestion was put forward by Malpass and Kravitz (1969) when they found the own-race effect only among their white subjects, and, indeed, evidence that their black subjects were somewhat better at recognizing white than black faces. Later research (Brigham & Williamson, 1979; Malpass et al., 1973) demonstrated that this is not usually the case. Goldstein (1979a, b) has shown, with anthropometric data, that among whites, blacks and Japanese, the highest degree of variation of feature shape is found in the faces of Japanese women. There were no other significant differences. As classes of physical objects, the faces of blacks, whites, and Japanese are equally varied.

However, all features are not equally salient (Shepherd, Davies & Ellis, 1981). Some, like the hair and eyes, are usually found to be more psychologically prominent than others, such as the lips and chin. (These data must be taken with a grain of salt, however, because they are based, almost entirely, on the judgments of white subjects.) Perhaps, because we do not distribute our attention evenly across the face, we fail to take advantage of the fact that other-race faces are as varied as own-race faces, but in different ways. The possibility exists that we come to rely on the features that vary most among faces of our own-race and do not modify the perceptual weights applied to each feature when we look at other-race faces. As a result, other-race faces seem to be less objectively distinguishable when in fact they are not. The source of this difference in discriminability rests in the perceiver.

It is not clear from Goldstein's data whether it is all features taken individually, or faces as a whole, that are equally varied across races. In the latter case, it would be important to know which race is more diverse in a given feature. Suppose, for example, it turned out that eyes and hair were most variable among white faces, and nose and chin were most variable among black faces. Then, to someone who attends to the eyes and hair, black faces look more alike than white faces, as Malpass and Kravitz suggest, although, in fact, they are not. While this state of affairs is conceivable, it is inconsistent with the fact that both blacks and whites are better at recognizing members of their own race.

The Depth-of-Processing Hypothesis

The issue of depth-of-processing is generally explored by manipulating the orienting task the subject is engaged in while viewing the study set. In a typical experiment, subjects are told to judge the sex or personality of the faces in the study set and are unaware that their memory will be tested. Several investigators (e.g. Bower & Karlin, 1974; Strnad & Mueller, 1977) have found that subjects instructed to judge faces on physical characteristics such as gender have poorer memory for those faces than subjects who made personality judgments about them. The interpretation given to such results is that personality judgments result in deeper, more elaborate memory traces that are easier to retrieve.

It has been suggested that poor other-race face memory reflects shallow processing strategies. Chance and Goldstein (1979) found that, when asked to give their "first reaction" to a face, subjects make shallower responses to other-race faces. Devine and Malpass (1985) manipulated instructions within a cross-racial recognition study. They argue that if the other-race deficit is the result of a shallow processing strategy, then providing the subject with a better strategy should reduce the deficit. They replicated the usual finding in that performance was better when subjects made personality judgments than when they evaluated physical features. Even deep processing, however, was no more effective than strategies elicited by simply telling the subject to remember the faces. Further, the other-race deficit was of the same magnitude in all conditions. Other-race recognition does benefit from a deep processing strategy but only to the extent that own-race recognition does. Shallow processing alone cannot explain the own-race effect. Winograd (1981) has offered another interpretation of the depth-of-processing effect on memory for faces. He has demonstrated that the superiority of deep processing in face recognition is not the result a more elaborate memory trace per se. The amount of information the subject encodes is less important than the nature of that information. Subjects who are asked to scan faces for distinctive features remember them as well as subjects who make personality judgments about them, and better than those who are asked to make judgments about specific facial features. Making judgments about features pre-selected for high distinctiveness and scanning faces for distinctive features produce equally good recognition. Winograd concludes that deep processing is advantageous not simply because it encourages broader sampling of facial features but because it increases the likelihood that a distinguishing feature will be sampled.

Given this interpretation of deep processing, we can restate the depth-of-processing hypothesis: the own-race effect is caused by a relative failure to encode distinctive features less often in other- race faces. If this is so, then why doesn't a processing strategy that encourages broader sampling (such as the personality judgments used by Devine and Malpass) close the gap? To answer this we must assume that, in addition to racial differences in scanning faces, we are less able to encode the features of an other-race face. It seems reasonable to suppose that if such differences exist, they are the result of our greater experience in perceiving and remembering members of our own race. In short, the depth-of-processing hypothesis cannot stand without the assumptions of the experience hypothesis.

The Prejudice Hypothesis

The issue of racial prejudice is a difficult one to study because it depends on the subject's willingness to make a frank self-report of his racial attitudes. Brigham and Barkowitz (1978) approached it by having their subjects fill out a series of questionnaires, including one that surveyed racial attitudes, after participating in a face recognition experiment. The correlation between racial attitudes and recognition of other-race races was near zero for both black and white subjects. There are two reasons why it would be hasty to reject prejudice as a factor simply on this basis. First, in the academic settings where most psychological research takes place, racial prejudice is not socially acceptable. A prejudiced subject might therefore be more apt to lie about his attitudes than to admit them. Second, Brigham and Barkowitz are only testing one possible version of the prejudice hypothesis. They assume that the more prejudiced an individual is, the more likely he is to forget an other-race face. An equally plausible interpretation is that highly prejudiced people are hyper-sensitive to other-race faces and might actually be better at recognizing them than those with more neutral attitudes would be. This version of the hypothesis predicts a curvilinear relationship between attitude and recognition performance. Unfortunately, Brigham and Barkowitz's presentation of their data does not permit its evaluation.

Galper (1973) took a different, and more questionable, tack in approaching the effects of racial attitude on recognition. Instead of seeking to demonstrate that prejudiced attitudes are associated with poor other-race memory, she looked for superior memory in subjects who might be expected to have positive attitudes towards another race: white students in a black studies class. The prejudice hypothesis predicts that such subjects would not show an own-race effect, and in fact, they did not. Galper points out that these white students were slightly more likely ($.10 < p < .20$) to remember black faces than white ones. She feels this "affirmative action" reverse own-race effect, although nonsignificant, is additional evidence in favor of an attitudinal explanation.

There is a major problem with the way the experiment was conducted. Galper departed from the general paradigm outlined above by presenting the complete study set all at once. The subjects could thus allocate their viewing time as they chose. There is reason to believe that white students in a black studies class would feel a certain social pressure to remember the black faces and would therefore take the opportunity to concentrate on them. Perhaps a more fundamental problem with this experiment is that it is designed to test the null hypothesis. A lack of an other-race deficit is interpreted as evidence for the experimental hypothesis. This is unsound not only on general principles but also in light of the variability data cited above. Neither whites nor blacks always show an other-race deficit. This is underscored by another, unpredicted finding in this experiment: blacks in the class did not show an own-race effect, although both whites and blacks in a comparison group did.

The Motivation Hypothesis

A curious finding indicates that, while lack of other-race experience may contribute to the own-race effect, there are other factors involved. Several investigators have tried training subjects in face recognition. A variety of procedures have been employed, ranging from the rudimentary (having subjects learn face-number paired associates) to the elaborate (having subjects learn feature-cluster "concepts"). Of the studies in which subjects were trained in own-race recognition (Baddeley, 1979; Elliot et al., 1973; Malpass, 1981; Malpass et al., 1973) none has demonstrated any posttraining improvement. Of the studies that have attempted to increase other-race recognition (Elliot et al., 1973; Goldstein & Chance, 1985;

Lavrakus et al., 1976; Malpass et al., 1973), all produced some improvement. The training that Lavrakus et al.'s subjects received was sufficient to wipe out a significant pretraining own-race effect in an immediate posttraining test. However, one week later all benefit from the training was lost. While training does not improve own-race recognition, it can eliminate an other-race deficit, but only temporarily.

It is clear that the subjects are not "learning" anything from the training. If they were, why should the improvement be so transient? What could a paired associates task teach them that they had not already had the opportunity to learn? Training may simply serve to focus attention on a class of stimuli, and increase motivation to remember them. This is also evident in the data of Malpass et al., who found that feedback in the form of a shock increases recognition of blacks by white subjects. In a less motivating verbal feedback condition, no learning occurred. Motivation may be a major factor in the own-race effect.

Barkowitz and Brigham (1982) pursued this by offering a monetary reward to their subjects. Subjects in a high-incentive condition won a lottery ticket for every correct identification they made and lost half a ticket for each false alarm. The lottery prize, of at least five dollars, was awarded immediately after the recognition test. White subjects in both the high- and no-incentive conditions showed an own-race effect. No comparison between recognition of whites by blacks in the two conditions is reported. We can only assume that they did not differ. The authors conclude that motivation cannot overcome the other-race deficit. This is not warranted. It is debatable how much of an incentive the lottery tickets really represented. They were, after all, only tickets, not cash. The scheme for awarding them, while not exactly complicated, is not so straightforward as to preclude confusion. Further, three different retention intervals were used: five minutes, two days, and seven days. This means that some subjects had to wait a full week before receiving their reward. This clearly reduced the motivational value of the incentive to practically nothing: the drop-out rate in the high- and no-incentive groups was the same. The incentive did not even motivate the subjects to return for the recognition test.

The Experience Hypothesis

According to the experience hypothesis, our difficulties in recognizing other-race faces stem from a deficiency in perceiving and encoding them. At the heart of this argument are two assumptions: that we acquire knowledge about faces through exposure to them and that the knowledge we acquire about faces of one race is less helpful when we are confronted with a person of another race. Although this sounds like a reasonable proposition, it has not found much support. The recognition performance of Brigham and Barkowitz's (1978) subjects was uncorrelated with the amount of their cross-racial experience. Lavrakus et al. (1976) actually found a significant negative correlation between quantity of experience and recognition in their subjects.

Cross et al. (1971) used demographic data (degree of integration of schools and neighborhoods) to estimate the amount of cross-racial experience their subjects had. The segregated whites made more incorrect identifications (false alarms) of blacks than did the integrated whites. For this reason, these data are usually (see Shepherd, 1981, for example) interpreted as evidence in favor of the experience hypothesis. A closer look at the data indicates that this is not justified. Cross et al. followed the procedure for presenting their stimuli that Galper used: the entire study set was presented at once. The segregated whites were free to use their viewing time primarily on the white faces, and apparently this is what they did. Table 1

Table 1

Values of *d'* from the data of Cross, Cross, & Daly (1971).

	Integrated Whites	Segregated Whites
Black faces	.63	.41
White faces	.92	1.70

lists $\underline{d}'$ values calculated from Cross et al.'s data. The difference between the integrated and segregated whites is far greater for white faces than black ones. It seems that the poor performance on black faces is the cost of good performance on white faces, not the product of a lack of experience with blacks.

Carroo (1986) compared black American and black African subjects and found the Americans to be superior at recognizing white faces. Further, she found a significant correlation between reported inter-racial experience and recognition accuracy. While it apparently supports the experience hypothesis, this study can be criticized on several grounds. The sample size was small (ten subjects in each group), only ten target photographs were used, and these were yearbook photographs from an American college. Most importantly, only photographs of white faces were used. We have no idea how these subjects would have performed on black faces.

It can be argued that none of these experimenters was testing the experience hypothesis at all. The hypothesis concerns differences in recognition performance <u>across</u> races, e.g. the fact that blacks recognize blacks better than whites do. It accounts for this by the differences in experience <u>across</u> races, e.g. that blacks have more experience with blacks than whites do. What all of these experiments have done is to look at differences in experience within a race. They have compared whites who had a lot of experience with blacks with whites who had little cross-racial experience. The difference between these two groups of whites is probably trivial compared to the massive differences between (most) blacks and whites. There is no evidence that the large difference across races is not responsible for the own-race effect. The experience hypothesis remains a viable one.

Chance, Turner and Goldstein (1982) and Goldstein and Chance (1980) have made an argument for the experience hypothesis based on developmental trends. They refer to a "face schema" that contains a child's growing knowledge of faces and serves as an "information processing structure (program) which increases efficiency of attending selectively in order to make relevant discriminations among faces" (Chance et al., p. 36). They make no specific claims for the face schema, but imply that it is a single, internalized representation of faces previously seen. The developing schema is responsible for the child's improving memory for faces. If a child is only or primarily exposed to faces of one race, then the schema will come to reflect the characteristics of that race. Supposedly, such a schema will be of little value when the child is asked to remember faces of other races. This argument is bolstered by data obtained from white children (Chance et al., 1982) First- and second-graders have relatively poor memory for white faces, presumably owing to their immature face schemas, and show no other-race deficit for Japanese faces. Fifth- through seventh-graders and college students have much better memory for faces, but show a significant own-race effect. These mid-westerners are developing face schemas that improve performance for both white and Japanese faces, but at a faster rate for white faces.

The ability to recognize own-race faces generally does not improve after adolescence, by which time performance is at a relatively high level. If the face schema theory is the correct explanation of the own-race effect, then we would expect other-race recognition to continue to improve beyond adolescence. A developmental study of older adults should show a declining other-race deficit, provided the subjects continue to encounter members of other races. Only one study has looked at older adults at all. Brigham and Williamson's subjects were whites and blacks whose mean age was 72. The theory is only partially supported by these data: as noted above, the blacks, but not the whites, showed an other-race deficit. Certainly, a 72-year-old black American has had ample opportunity to develop a white face schema. Schema theory alone cannot explain this result.

It is plain from the literature surveyed that the issue of race in face memory is a thorny one. That the own-race effect assumes different forms from one experiment to the next is evidence of the complexity of the underlying processes. Obviously, race does not affect perception and memory of faces in the same way and to the same extent in all people, as a variable such as illumination might. The data indicate, as if it needed to be shown, that there is no "black" or "white" mode of responding to a face. Anyone conducting research on face recognition must accept variability in the data. Anyone seeking to explain the own-race effect must deal with the possibility that it has multiple causes.

Of the hypotheses that have been entertained, only one, the "they all really do look alike" hypothesis, can really be discounted. Racial attitudes may influence other-race memory, but this question will probably never be resolved due to the social unacceptability of prejudice. It is just as well that this line of inquiry be dropped given its potential for increasing mistrust and misunderstanding. Despite the efforts of Barkowitz and Brigham, the influence of motivation is still largely unexplored. The role of experience remains an open question, but it must be approached in a more fruitful way, one that addresses between-race differences.

The Inversion Effect and Familiarity

The "inversion effect" was first demonstrated by Yin (1969), who showed subjects pictures of faces, houses, airplanes, and drawings of "men in motion". These stimuli were selected because they are mono-oriented -- generally seen only in the upright orientation. When viewed in their usual orientation, recognition performance was best for faces, but when the stimuli were inverted, performance was worst for faces. All three study/test combinations, inverted/inverted, inverted/upright and upright/inverted, yielded this pattern. Yarmey (1971) and Scapinello and Yarmey (1970) replicated the inversion effect for both familiar and unfamiliar faces using buildings and faces of dogs as comparison stimuli.

Yin (1970) concluded that human faces are "special stimuli", and that there is some neurological structure or specialization that is responsible for face recognition. He and others who share this view (see Ellis, 1983 for a review) generally place this structure in the right hemisphere, where lesions have been known to cause a disorder known as prosopagnosia -- an inability to recognize or to form new memories of faces. Yin found that this impairment existed only for upright faces. Subjects with right hemisphere lesions performed no worse than normals on inverted faces.

Diamond and Carey (1986) question the notion that faces are processed differently than other visual stimuli. Indeed, it is debatable whether the deficits caused by right hemisphere lesions are restricted to memory for faces or are generalized problems in visual memory (Benton, 1980; Ellis, 1983; Ellis, 1975; Hecaen, 1981). Diamond and Carey argue that faces are special only by virtue of their physical characteristics (see below) and the vast number of them that we have committed to memory, and not because of any specialized face processing structure -- any stimulus that shares these properties will be perceived and remembered in the same way.

Diamond and Carey draw a distinction between isolated features that involve only one part of the face, and relational features that encompass several. For examples they offer eye color as an isolated feature and "wide-set eyes with a low forehead" as a relational feature (Diamond & Carey, 1986, p. 108). This is not a clear-cut distinction. To which category does one assign the shape of the eyes, for example? On the one hand, it is a local feature, but on the other hand, eye shape can affect the way the cheek bones are perceived. While there is no sharp dichotomy between isolated and relational features, it is still meaningful to characterize features in this fashion. Faces differ from other classes of complex stimuli in isolated features such as the number, location, and overall shape of their features and, barring disfigurement, faces are reasonably uniform. It would seem that isolated features are more useful in separating faces from non-faces than in distinguishing one face from another. Isolated features alone are not sufficient for the recognition of individual faces. This presumably makes relational features more salient.

There is reason to believe that relational features serve as the primary basis for distinguishing between individual faces. Harmon (1973) has studied face recognition using pictures in which the high spatial frequencies have been filtered out. Such pictures degrade local or isolated features while largely preserving global or relational information. Familiar faces are easily recognized from pictures treated in this way. Haig (1984) has explored the role of individual features and their location within the face. He used image-processing equipment that can display and alter black and white photographs. This allowed him to change relational features by moving features within the face. Haig's subjects were sensitive to minute vertical displacement of the mouth, eyes, and nose, and to inward displacement of the eyes. Czigler (1985) has shown that relational information can interfere with judgments about isolated features. He showed subjects pairs of composite photographs that differed in one to four features or were identical. Subjects were asked to determine whether the faces shared a target feature (eyes or hair). When the target feature was in fact the same in the two faces, reaction time increased with the number of non-target features that differed, indicating that the subjects had difficulty ignoring the influence of extraneous relational information.

Diamond and Carey (1986) suggest that many of the comparison stimuli used in earlier investigations of the inversion effect (houses, headless figures wearing costumes, landscapes) are unlike faces in that they are easily recognizable from isolated features alone. An unusual window, a distinctive posture, or a large rock can trigger a correct "old" response despite the fact that the rest of the picture does not look particularly familiar. They have also presented evidence (Diamond & Carey, 1986) that isolated features are not responsible for the inversion effect. Isolated features are apparently no harder to perceive in inverted faces than in other inverted stimuli. They asked subjects to determine whether relatively isolated features were present in photographs of faces, landscapes, and houses. For each class of stimuli, reaction time was longer if the photograph was inverted, but there was no interaction between reaction time and type of stimulus. Turning a picture upside down has no greater effect on decisions about isolated features of faces than it does on other kinds of stimuli. The inversion effect, therefore, is largely due to the effects of inversion on relational features. Comparison stimuli that obviate attention to relational features do not provide a fair test of the uniqueness of the inversion effect to faces.

Not all of the data suffer from this flaw. Yarmey (1971) and Scapinello and Yarmey (1970) found an inversion effect with comparison stimuli abundant in relevant relational features: faces of dogs. Unfortunately, we cannot draw any conclusions from these data because faces of dogs and faces of people are not psychologically equivalent. We are much more experienced in discriminating between faces of people than dogs. Most people cannot recognize very many dogs by their faces, but dog show judges can, and Diamond and Carey enlisted some judges as subjects. By using dogs as comparison stimuli and show dog judges as subjects, they were able to test the inversion effect with stimuli that were both high in relational features and very familiar to their subjects.

Diamond and Carey (1986) showed "experts" (dog show judges) and "novices" (undergraduates) upright and inverted pictures of show dogs and human faces. Full-body profiles, rather than face shots, were used for the dogs because that is how they are judged in dog shows. In a forced-choice recognition test, both novices and experts showed an "inversion deficit" for faces, performing better on upright than on inverted faces. The experts showed an equally large inversion deficit for the show dogs, but the novices showed none. Subjects inexperienced with dogs performed as well when the dogs were inverted as they did when the dogs were upright. Further, the novices recognized more of the inverted dogs than did the experts. These data demonstrate that the inversion effect can be found with stimuli, other than human faces, that are mono-oriented and rich in relational features if the viewer is familiar with them. There is no need to hypothesize the existence of a special "face processor". Faces are special only to the extent that our expertise is limited to them.

A Prediction

Experts show an inversion deficit, novices do not. This distinction can be utilized in a test of the effects of experience on the own-race effect, because the presence of an inversion deficit can tell us something about the familiarity of a stimulus class to a given subject. If familiarity is responsible for superior own-race recognition then, in a sense, we are experts on own-race faces and relative novices on other-race faces. Extrapolating from Diamond and Carey's data, we would predict no inversion deficit for other-race faces, or at least one that is significantly smaller than that for own-race faces. The same prediction can be made from the perspective of Goldstein and Chance's (1980) notion of face schemas, according to which familiarity leads to own-race face schemas that are well-developed and mono-oriented. Other-race face schemas are less so. Children below the age of 10, who presumably possess immature face schemas, do not show an inversion deficit for faces (Carey, Diamond, & Woods, 1980). The ability to recognize inverted faces of familiar people has actually been shown to decline steadily from age 10 to 20 (Goldstein, 1975), during which period recognition of upright faces is improving (Carey et. al., 1980), and face schemas are becoming, if anything, more defined. Because this schema is of little or no use when we are confronted with a face in a novel orientation, we again predict little or no own-race effect for inverted faces.

Several investigators have concluded that the mnemonic strategies subjects employ differ for upright and inverted faces. Subjects usually report trying to form a general impression of an upright face, whereas an inverted face tends to elicit a search for a distinctive feature (Phillips & Rawles, 1979; Yin, 1969). Shifting strategies is advantageous: Phillips and Rawles (1979) found that applying a feature search strategy was significantly correlated with accuracy on inverted faces. One of two things is happening: either the subjects are not equally accomplished in the two methods, or many of them fail to switch. Yin (1969) found a negative correlation between subjects' abilities to recognize upright and inverted faces. Phillips and Rawles (1979) showed subjects upright and inverted woodcuts (taken from an edition of Aesop's fables) and faces. There was a significant positive correlation between the subjects' performance on the upright and inverted woodcuts but a near zero correlation for performance on the upright and inverted faces. These results are particularly striking in the light of Woodhead and Baddeley's (1981) finding that subjects are consistent in their upright face recognition accuracy from one experiment to the next. We can predict accuracy on upright faces from past performance, but cannot predict accuracy on inverted faces from performance on upright ones.

Optimal performance in an inversion study apparently demands that subjects switch strategies. When faces are upright, and we may assume subjects are using the general impression strategy, own-race faces are remembered better than other-race faces. Because own-race faces are recognized best when upright, we might expect these faces to suffer most from inversion, just as faces suffer more from inversion than airplanes. This should be true whether or not the subject changes tactics. If a subject abandons the general impression strategy and adopts feature search for the inverted faces, we have no reason to predict a continued advantage for own-race faces. Now the face is in a novel orientation and the subject is using a different mode of processing. On the other hand, if the subject persists in the general impression strategy, accuracy on inverted faces can be expected to be particularly bad overall, which might obscure any advantage for own-race faces. The poor performance on inverted faces will be highlighted by the good performance on upright own-race faces, resulting in a larger inversion deficit for own-race faces.

Three lines of reasoning converge on the same prediction: inversion should reduce or wipe out the own-race effect: (a) We expect more familiar stimuli to suffer more from inversion than less familiar stimuli; (b) More familiar classes of stimuli produce more entrenched mono-oriented schemas that enhance upright recognition but are defeated by inversion; and (c) Strategies for remembering faces in the familiar orientation, that favor own-race faces, are not fruitful for remembering inverted faces. Thus, if the own-race effect were found to persist despite inversion, we would have three reasons for rejecting the experience hypothesis.

Experiment 1

The primary purpose of this experiment was to test the experience hypothesis. White and black subjects were shown pictures of white and black faces in the upright and inverted orientations. Memory was assessed by a "yes-no" recognition test. An own-race effect for the upright faces is expected. Of interest is whether the effect extends to inverted faces. A finding of little or no own-race effect for inverted faces would favor the experience hypothesis. On the other hand, a significant own-race effect for inverted faces would support the notion that the difference in familiarity between own- and other-race faces is not sufficient to account for the own-race effect.

A further aim of the experiment was to replicate the own-race effect with upright faces with an improved stimulus set. Typically, research on the own-race effect has been carried out using rephotographed black and white pictures from high school yearbooks. Because the experimenter has only one picture of each stimulus person, subjects can succeed in the "face recognition" task by recognizing peripheral details such as clothing, pose, and expression, or even totally extraneous details such as shadows or specks of dust on the slide. The present research avoided such pitfalls by employing two photographs of each stimulus person that differ in clothing and expression. It was predicted that this stimulus set would prove more challenging because it tests face- and not picture-recognition, and therefore overall levels of performance would be somewhat lower than previously obtained.

In addition, the data were inspected for stimulus effects. It is conceivable that the own-race and inversion effects are created by a few isolated faces -- white faces that blacks never remember or one or two faces that are never remembered when inverted. No previous work on the own-race or inversion effects has included such a check and the possibility remained that both or either of the effects is merely an artifact of stimulus set.

Another question to be answered by the analysis concerned mnemonic strategies. The correlation between the probability that a face is recognized by whites subjects and black subjects is an indication of the degree of overlap in viewing strategies. Likewise, the probability that a face is recognized in the upright and inverted orientations will also be compared. Based on the findings of Phillips and Rawles (1979) and Yin (1969), it was predicted that there would be no correlation between upright and inverted recognition.

Method

Subjects. The subjects were 134 college students. Nineteen Brooklyn College students participated in the experiment in partial fulfillment of a class requirement. The rest of the subjects were from Manhattan Community College and Baruch College and participated on a voluntary basis as part of an in-class activity. Data from 33 of the subjects were not used in the analysis because of failure to follow instructions or to make a response for every slide in the test set. Of the remaining subjects, 52 were black, 26 white, 14 Hispanic, and 9 Asian. Race was determined by self-report. Except where noted, the Hispanic and Asian subjects were excluded from the analysis.

Stimuli. The stimuli were color slides of 40 black and 40 white male college students. The students' race was determined by self-report. Several head-and-shoulder shots were taken of each student. In one, the student wore a blue jacket (which completely covered his clothing) and smiled (jacket/smile), and in another his own clothing was exposed and he did not smile (no-jacket/no- smile). All of the photographs were taken in the same room, under the same lighting conditions, against a gray background. Side ligh-ting was used to eliminate any shadows on the face. Cues to the students' height were minimized: the students were photographed seated on a backless stool and the vertical position of the camera was changed to keep the subject's face in the center of the frame.

Procedure. Two overlapping study sets were be used. Each set was created by randomly selecting 40 target faces from the pool of 80, with the constraint that 20 were black and 20 were white. The purpose of dividing the faces into targets and distractors in two different ways was to reduce the chance that the targets as a group, or a subset of them, are especially hard or easy to discriminate from the distractors. Each set was randomly divided into two blocks, one to be presented upright and the other inverted, again with the constraint that half of the faces in each block were black and half were white. The jacket/smile slides were used in the study set, while the 80 no-jacket/no-smile slides made up the test set. The test set was also divided into two blocks, one upright and one inverted. The 20 targets from the upright study set appeared in the upright test set with an equal number of distractors. The same was true of the inverted test set. Two different presentation orders were used for each study set, constrained so that no more than three faces of the same race or of the same type (target or distractor) appeared consecutively.

Subjects were run in groups of 5 to 30. At the outset, they were told that their memory for upright and inverted faces would be tested. They were informed that a recognition test would be administered in which half of the faces would be old and half new, that the faces would appear in the same orientation in the test set as they had in the study set, and that a different photograph would be used in the test set.

Each subject saw only one study set. The study set was presented at a rate of five seconds per slide, with a blank slide between blocks to signal the change of orientation. The test set was presented immediately following the study set. During the recognition test the subjects filled out a response sheet offering them a choice of four responses: Old -- Sure, Old -- Half Sure, New -- Half Sure, and New -- Sure. The subjects were given as much time as they needed to make a response. For approximately half of the subjects the inverted block was presented first during both the study phase and the recognition test. At the end of the recognition test, the subjects were asked to indicate their race on the back of the response sheet.

Results

Log-Linear models. The data were submitted to a log-linear analysis which provides independent measures of sensitivity and response bias similar to those of Choice Theory (Macmillan & Kaplan, 1986), and allows testing of specific models. The analysis assumes a linear model, as in analysis of variance, in which the frequency of each cell is the product of a grand mean, and, potentially, each main effect, and all possible interactions. Models need not contain all possible effects; one that does is said to be saturated, and is frequently not of interest. Models are tested by a chi-square goodness-of-fit test in which expected frequencies are generated based on the parameters included in the model. A nonsignificant chi-square value indicates that the model predicts the data adequately. A saturated model, by definition, fits the data perfectly, but leaves no degrees of freedom. Fewer parameters will leave more degrees of freedom, but generally result in poorer fit. The goal of log-linear analysis is to determine which effects can be deleted without harming the fit unduly.

The significance of an effect can be assessed in two ways. First, the size of the effect under the saturated model can be determined. This value, called lambda, can be divided by its standard error, to provide a standardized parameter. Second, the importance of an effect can be gauged by a log-likelihood chi-square goodness-of-fit test. BMDP, a statistical software package, contains a log-linear analysis program that generates expected frequencies for non-saturated models and performs goodness-of-fit tests on them. The significance of an effect can be tested by comparing the goodness-of-fit of a model that contains that effect with one that does not but is otherwise identical. The difference between the likelihood ratio chi-square values indicates the goodness-of-fit gained by including the effect (and losing the associated degrees of freedom). The nature of the relationship between these two measures is not clear (N. A. Macmillan, personal communication, May, 1987), but they are generally in agreement.

For simplicity, the variables will be referred to as indicated in Table 2. In general, interactions are denoted by the initials of each of the variables involved (e.g. Knoke & Burke, 1980) and that practice will be followed

Table 2

Abbreviations for Variables in Experiment 1.

Name	Variable	Levels
(R)esponse	Subject's Response	Old or New
(S)tatus	Status of Slide	Target or Distractor
(F)ace	Race of Slide	Black or White
(O)rientation	Orientation of Slide	Upright or Inverted
(P)erson	Race of Subject	Black or White

Table 3

Observed Frequencies of Old and New Responses in Experiment 1.

Person	Orientation	Face	Status	Response Old	New	Total
White N = 26	Upright	White	Old	178	82	260
			New	84	176	260
			Total	262	258	520
		Black	Old	161	99	260
			New	110	150	260
			Total	271	249	520
	Inverted	White	Old	159	101	260
			New	81	179	260
			Total	240	280	520
		Black	Old	151	109	260
			New	147	113	260
			Total	298	222	520
Black N = 52	Upright	White	Old	346	174	520
			New	187	333	520
			Total	533	507	1040
		Black	Old	342	178	520
			New	138	382	520
			Total	480	560	1040
	Inverted	White	Old	300	220	520
			New	256	264	520
			Total	556	484	1040
		Black	Old	310	210	520
			New	276	244	520
			Total	586	454	1040

Table 4

Likelihood Ratio Chi-Square Values and Standardized Lambdas in Experiment 1.

Effect	X^2	*p*	Lambda	*p*
RSFOP	0.00	n.s.	0.034	n.s.
RSOP	8.82	<.01	2.969	<.01
RSFP	25.12	<.001	4.999	<.001
RSFO	5.11	<.05	2.118	<.05
RFOP	0.03	n.s.	0.176	n.s.
RSF	4.44	<.05	3.514	<.01
RSO	80.92	<.001	7.451	<.001
RFO	11.41	<.001	3.348	<.001
RFP	8.92	<.01	2.934	<.01
ROP	5.36	<.05	2.257	<.05
RS	289.23	<.001	16.030	<.001
RO	12.08	<.001	2.428	<.05
RF	1.30	n.s.	1.850	n.s.
R	7.21	<.01	2.235	<.05

Note. There is one degree of freedom for each effect.

here. Sensitivity is measured as the interaction between Response and Status (RS). Response bias is reflected in the Response effect and its interactions with Face, Orientation, and Person, that do not also include Status.

Sensitivity. The obtained cell frequencies are displayed in Table 3. Likelihood ratio chi-square values and standardized lambdas for the effects of interest are shown in Table 4. As predicted the own-race effect (RSFP) and the effect of inversion (RSO) were highly significant. However, the interaction of primary interest, the interaction of the own-race effect with Orientation (RSFOP) was not.

Overall, sensitivity interacted with Face (RSF) in an unexpected manner: sensitivity was higher for white faces. The advantage for white faces arose primarily from performance on inverted faces. This state of affairs is reflected in the interaction between Face, Orientation, and sensitivity (RSFO). In the top half of Table 5, the data are pooled across Person, and listed separately for upright and inverted faces. The interaction between sensitivity and Face (RSF in Table 5) is found only for inverted faces. When black and white subjects are considered separately, as in Table 6, the advantage for white faces when inverted (RSFO), is no longer significant, although both groups of subjects show the same trend. Presumably, this discrepancy is due to the loss of power resulting from the division of subjects. That the poorest performance was obtained with inverted black faces is clear from Table 7, which lists the lambdas and their standard errors for each condition.

Response bias. Overall, there was a significant response bias; subjects were more likely to respond Old than New. The data were reanalyzed by adding the Old--Half Sure responses to the New category. This procedure simply changed the direction of the response bias without altering sensitivity (the standardized lambda changed from 16.030 to 16.613) and so was abandoned. Table 4 indicates that there was a Response by Orientation (RO) interaction, in which

Table 5

Likelihood Ratio Chi-Square Values and Standardized Lambdas in Experiment 1, by Orientation, Pooled Across Person, and by Person, Pooled Across Orientation.

Effect	X^2	p	Lambda	p
		Upright Faces -- All Subjects		
RSF	0.01	n.s	0.087	n.s.
RS	337.28	<.001	17.854	<.001
RF	2.77	n.s.	1.666	n.s.
R	0.26	n.s.	0.532	n.s.
		Inverted Faces -- All Subjects		
RSF	8.93	<.01	2.985	<.01
RS	33.18	<.001	5.727	<.001
RF	10.09	<.01	3.132	<.01
R	18.48	<.001	4.293	<.001
		Black Subjects -- Both Orientations		
RSF	1.21	n.s.	1.099	n.s.
RS	188.70	<.001	13.580	<.001
RF	0.53	n.s.	0.713	n.s.
R	5.41	<.05	2.374	<.05
		White Subjects -- Both Orientations		
RSF	27.45	<.001	5.219	<.001
RS	100.42	<.001	9.326	<.001
RF	9.08	<.01	2.961	<.001
R	1.84	n.s.	1.248	n.s.

Note. There is one degree of freedom for each effect.

Table 6

Likelihood Ratio Chi-Square Values and Standardized Lambdas in Experiment 1, for Black and White Subjects.

Effect	X^2	p	Lambda	p
		Black Subjects		
RSFO	3.48	n.s.	1.864	n.s.
RSO	81.81	<.001	9.018	<.001
RSF	1.38	n.s.	1.286	n.s.
RS	189.47	<.001	13.720	<.001
RF	0.57	n.s.	0.938	n.s.
RO	16.79	<.001	4.054	<.001
R	5.41	<.05	2.081	<.01
		White Subjects		
RSFO	1.63	n.s.	1.276	n.s.
RSO	7.91	<.01	2.746	<.01
RSF	27.68	<.001	5.215	<.001
RS	100.43	<.001	9.727	<.001
RF	9.08	<.01	2.930	<.01
RO	0.07	n.s.	0.105	n.s.
R	1.85	n.s.	0.029	n.s.

Note. There is one degree of freedom for each effect.

upright faces were more likely to be called New, and a Response by Face by Orientation (RFO) interaction, in whichwhite faces were more likely to be called Old when they were upright than when inverted, and black faces were more likely

to be called Old when they were inverted than when upright.

Black and white subjects showed different patterns of bias. While overall there was a tendency to call own-race faces New (RFP in Table 4), Table 6 reveals that the white subjects were responsible for this interaction (RF is

Table 7

Lambdas for Sensitivity in Experiment 1.

Black Subjects			White Subjects		
Race of Face	Orientation		Race of Face	Orientation	
	Upright	Inverted		Upright	Inverted
White	.316**	.085**	White	.378**	.312**
	(.033)	(.021)		(.047)	(.046)
Black	.418**	.067*	Black	.199**	.053
	(.034)	(.031)		(.045)	(.044)

Note. There is one degree of freedom for each effect. Standard errors are in parentheses.
* p < .0 ** p < .001

Table 8

Ratios of Response Bias to Sensitivity Parameters in Experiment 1.

Black Subjects			White Subjects		
Race of Face	Orientation		Race of Face	Orientation	
	Upright	Inverted		Upright	Inverted
White	0.089	0.824	White	0.024	-0.147
Black	-0.218	1.910	Black	0.090	2.773

Note. There is one degree of freedom for each effect. A positive number represents a bias toward responding Old.

significant only for the white subjects). The response bias of the black subjects was affected by orientation (RO in Table 6). They tended to call inverted faces New, despite an overall bias (R) toward responding Old.

The value of the Response parameter was calculated for each condition. The ratio of the Response to the sensitivity parameters are given in Table 8. The subjects appear to be more biased, and biased toward responding Old, when sensitivity is low, that is with inverted faces, and particularly inverted black faces.

Analysis for stimulus effects. The data were inspected for the possibility that the own-race effect and inversion deficit were caused by a few isolated faces -- faces of one race that are only remembered by members of that race, or faces that are never remembered when inverted -- rather than a general difficulty in recognizing other-race and inverted faces. This danger is especially acute when proportion correct is relatively low, as it is here (.61 overall). Only the 18 black faces and 21 white faces that appeared in both orientations were included in this analysis. For each face, proportion correct was calculated by race of subject and orientation. These were submitted to an arc-sin transformation, and the group means are presented in Table 9. An analysis of variance was conducted in which the between-group factor was race of face, and race of subject, and orientation were within-group factors. The error variance reflects differences in the probability that individual faces are recognized. There was no effect of subject race, $F(1,37) = 3.15$, $p > .05$, or interaction between orientation and race of face, $F(1,37) = 2.47$, $p > .10$. The critical effects were present, however. Both the main effect of orientation, $F(1,37) = 15.77$, $p < .001$, and the interaction between race of subject and race of face, $F(1,37) = 13.76$, $p < .001$, were significant. Because the error variance was not sufficient to wipe out the group differences across orientation, or the interaction with race of subject, we can conclude that the effects of race and orientation cause genuine deficits in recognition, independent of the individual face.

Processing strategies. For all 80 faces, proportion correct was calculated for each condition in which the face appeared: in each orientation, and presented to each race of subjects. These were submitted to an arc-sine transformation and correlations were calculated between transformed proportion correct across race and across orientation. These are presented in Table 10. The results are quite clear cut. The performance of white and black subjects on a given face is similar when it is upright and when it is inverted. On the other hand, there is no relationship between recognition of a face in the two orientations.

A related question is to what extent face recognition skills generalize to different types of faces and viewing

Table 9

Arc-Sine-Transformed Proportion Correct by Face in Experiment 1.

White Faces			Black Faces		
Race of Subject	Orientation		Race of Subject	Orientation	
	Upright	Inverted		Upright	Inverted
White	.868	.783	White	.750	.531
Black	.740	.595	Black	.852	.540

Table 10

Correlations between Arc-Sine-Transformed Proportion Correct for Each Slide for Black and White Subjects, and for Upright and Inverted Faces.

Race of Face	N	r	p	r	p
	Black vs. White Subjects				
		Upright		Inverted	
Black	29	.507	<.01	.596	<.001
White	30	.698	<.001	.570	<.001
Combined	59	.549	<.001	.562	<.001
	Upright vs. Inverted Orientation				
		Black Subjects		White Subjects	
Black	18	.251	n.s.	-.043	n.s.
White	21	.290	n.s.	.309	n.s.
Combined	39	.235	n.s.	.233	n.s.

Table 11

Correlations between Each Subject's d' for Black and White Faces, and for Upright and Inverted Faces.

Race of Subject	N	r	p	r	p
		Black vs. White Faces			
		Upright		Inverted	
Black	52	.032	n.s.	.203	n.s.
White	26	.525	<.01	-.007	n.s.
Combined[a]	101	.156	n.s.	.117	n.s.
		Upright vs. Inverted Orientation			
		Black Faces		White Faces	
Black	52	.233	n.s.	-.063	n.s.
White	26	.482	<.05	.181	n.s.
Combined[a]	101	.205	<.05	.115	n.s.

[a]includes 14 Hispanic and 9 Asian subjects.

conditions. Four d' values were calculated for each subject, for black and white upright and inverted faces. Table 11 displays the correlations between sensitivity in the various conditions. In general, sensitivity to white and to black faces was uncorrelated, as was sensitivity to upright and inverted faces. However, three correlations were significant: first, between recognition of black and white upright faces for white subjects, second, between recognition of upright and inverted black faces for white subjects and third, between recognition of upright and inverted black faces for all subjects combined.

It appeared from the scatterplots that one outlying data point might be responsible for the significance of these correlations. In each case the same subject was involved, a white subject with a negative d' for white upright faces and large negative d's for black faces. When this subject was removed from the analysis, only the correlation between white subjects' performance on white and black faces was still significant, r (N = 26) = .414, $p < .05$.

<u>Discussion</u>

These data provide the most powerful evidence to date of the reality of the own-race effect. Not only was the own-race effect obtained with a large and improved stimulus set, but it has now been demonstrated that the effect does not hinge on a few unusual stimuli.

In general, it seems that the level of performance was lower than has been previously reported, although it is difficult to make direct comparisons with other data owing to differences in the number of targets, duration of presentation, the other researchers' choice of statistic, and the fact that Experiment 1 had several different types of targets. Diamond and Carey (1986), using 5 second exposures of 18 targets in a two-alternative forced-choice procedure, observed 89% correct for upright faces and 70% correct for inverted faces in one experiment. A second experiment yielded similar levels of performance: 88% and 65% correct for upright and inverted faces respectively. With 3 second exposures to 20 targets and a 12 trial two-alternative forced-choice test, Yin obtained roughly equivalent performance, 92% correct for upright faces and 63% correct for inverted faces. These two studies found higher levels of performance than in the present data (69% for upright, own-race faces, 57% for inverted, own-race faces), but it is not clear how to interpret this because they used a procedure, forced-choice, that is known to produce higher levels of performance than the yes-no task employed in Experiment 1. Phillips and Rawles (1979) used a "yes-no" task, with 20 targets observed for 1.5 seconds each, and obtained d's for upright and inverted faces of 2.16 and 0.91 respectively. Mean *d'* for the 10 upright, own-race targets was 0.554, and for the 10 inverted, own-race targets it was 0.215, both substantially lower than those obtained by Phillips and Rawles.

Several of the studies of the own-race effect in which exposure to the targets was controlled have used *d'* as a measure of sensitivity. Of these, three have reported group means, and are listed in Table 12. Brigham and Barkowitz's (1978) results show the lowest levels of performance, and are most similar to the present data.

Table 12

Values of *d'* for Own- and Other-Race Recognition in Experiment 1 and Three Previous Studies.

Study	No. of Targets	Exposure	Own-Race *d'*	Other-Race *d'*
Barkowitz & Brigham (1979)	24	1.5 sec.	1.14	.65
Brigham & Barkowitz (1978)	24	1.8 sec.	1.60	.87
Malpass & Kravitz (1969)	20	2.0 sec.	1.89	1.49
Experiment 1	20	5.0 sec.	0.57	0.41

They also used the shortest exposure duration. Laughery etal. (1971) compared two different exposure times (10 and 32 seconds) and found that longer exposure led to a greater probability of recognition. Although they compared much longer exposures than those currently under discussion, it seems likely that, other things being equal, a 5 second exposure should lead to better recognition than a 1.5 second exposure. Therefore, one would expect higher levels of performance in Experiment 1 than in the other three studies listed in Table 12. This is not what was found, however. The most likely explanation of this discrepancy seems to be that research conducted with only one picture of each target over-estimates performance.

The major finding of this experiment was the failure of the own-race effect to interact with orientation. This result was quite clear-cut among the white subjects, who performed better on inverted white faces than on upright black faces. Although the black subjects performed about as well on inverted black faces as on inverted white faces, this appeared to be due to the particularly poor recognition of inverted black faces shown by all subjects, rather than particularly good recognition of inverted white faces. The white subjects, after all, were more successful at recognizing inverted white faces than were the black subjects. Thus, the familiarity hypothesis gains no support from these data. This experiment should have afforded a very sensitive test of the hypothesis because, unlike previous research, it compared the difference in experience across races. The failure to find support in these data, or in the previous attempts, indicates that the hypothesis is inadequate.

Indeed, why should familiarity bring greater sensitivity? There is no reason to assume that simply being exposed to faces should result in learning. We know that our representations of faces are insufficient only when we make a mistake, for example, when we greet someone we have never met. The feedback from such incidents supplies both the motivation and the only opportunity to learn to differentiate between faces. Until we have meaningful social contact with members of other races, we may never learn to focus on the most informative dimensions of other-race faces and so they may "all look alike" to us. Most people's social contacts, whether their neighborhood is integrated or not, provide more impetus to learn own-race faces. A teacher in an integrated school, on the other hand, would probably have good recognition of other-race faces because he must in order to carry out his job. This idea fits neatly with the work of Lavrakus et al. (1976) who found that other-race recognition was negatively correlated with interracial exposure, but positively correlated with reported number of interracial friendships. After all, one needs to be able to distinguish friends from passersby.

Logically, it is possible to draw a distinction between learning to remember faces and the act of remembering a particular face. Behind the results of Lavrakus et al. lurks the notion that the former process is an active rather than automatic process, one that requires attention and a desire to learn. This stands in contrast to the apparent automaticity of memory for individual faces: recall that intention to remember a particular face has little effect on subsequent recognition (Bower & Karlin, 1974; Devine & Malpass, 1981). In the distinction between the two processes we can see a parallel with the acquisition of cognitive skills as described by Anderson (1980). An unpracticed task demands the full attention of the actor, but, with repetition, it is eventually accomplished automatically, and we may even loose the ability to describe the procedures involved. While it is likely we are never able to explain how we recognize faces, it stands to reason that, with experience, the process becomes both more sophisticated and less amenable to verbalization. Klatzky (1984), for example, has remarked on our inability to articulate anything about how we recognize faces. All of this suggests that our ability to encode faces automatically is the fruit of over-learning a formerly effortful task.

Further insight into the processes underlying face recognition can be gained from the correlations between performance across race and across orientation. It appears that faces have properties that determine the likelihood that they will be recognized that are independent of the viewer's race. This was seen in the fact that the probabilities that a face would be recognized by white and by black subjects were correlated. Viewers, on the other hand, do not possess abilities to recognize faces that are independent of the race of the face. This was indicated by the absence of correlation between a subject's $\underline{d}'$ for white and black faces, which is striking when one considers that Woodhead and Baddeley (1981) found recognition of own-race faces to be consistent with past performance. What makes a face easy to recognize is independent of the viewer's race, but what makes a subject good at recognizing faces is not independent of the race of the face. The skills that the subject brings to the task are fairly stable for own-race faces but not predictable for other-race faces.

Recognition of upright and of inverted faces were unrelated, whether examined by subject or by face, confirming earlier reports (Phillips & Rawles, 1979; Yin, 1969) that the two tasks involved different abilities. The absence of correlation between a subject's performance on upright and inverted faces has two possible interpretations. First, the subject may fail to switch to an appropriate strategy for the inverted faces. His skill at applying the strategy on upright faces will be useless on inverted ones. Second, it could be that the two tasks (upright and inverted recognition) are fundamentally different. A subject who is good at one will not necessarily be good at the other. The latter possibility gains indirect support from the fact that the probabilities that a face is recognized when upright and when inverted are also unrelated. If the subjects persisted in applying a strategy suitable for upright faces, then presumably the same faces would be recognized, regardless of their orientation.

The most curious result was the poor performance on inverted black faces demonstrated by all subjects. Valentine and Bruce (1986a) report the same pattern of results. It cannot be attributed to greater homogeneity of the black faces because the best performance of all was by black subjects on upright black faces. Apparently, the process invoked by inversion, or the facial properties that it makes salient, favors white faces. If we can assume that subjects confronted with inverted faces scan them for distinctive features, then inverted black faces must have fewer such features. One obvious difference between black and white faces is color variance: white faces have a wider range of hair and eye color. Another involves the outline and texture of the hair which tends to be more varied in whites, who have (at least in this stimulus set) a greater variety of lengths, clear differences in texture, and are more likely to wear their hair parted.

One reason why these features might be made more salient by inversion is that they are peripheral to the face itself. When a face is upright, we attend to the central features: the saliency of the eyes, nose, and mouth has been demonstrated through studies of eye movements (Luria & Strauss, 1978; Walker-Smith, Gale & Findlay, 1977). These aspects of a face are highly informative and are intrinsic parts of the face in a sense that the hair is not. Barring plastic surgery, they change only in a few, very predictable ways, as with age or expression, whereas hair can be radically transformed with a comb. The benefit that can be derived from our experience with these parts of the face may be subverted by inversion. This would explain why inverted faces are often described as looking "odd" or as "disturbing" instead of merely upside down. Consequently, we might rely on peripheral details, aspects of the face that may not be fully adequate for recognition, but do not force us to attempt to view this inverted "thing" as a face. Navon (1977) has demonstrated that, for non-face stimuli, global features can dominate local ones by inhibiting responses to local detail. If the tendency to respond to an inverted face as a face were suppressed at least partially, then the outline of the head might serve as a stronger cue than local features like the eyes. Interestingly, children, at an age when they show no inversion deficit, tend to rely on such features as hair, hats, and glasses in their identifications (Carey & Diamond, 1977). Whether or not adults use such strategies with inverted faces is an open question, but they offer an explanation for the poor recognition of inverted black faces.

Response bias was related to sensitivity in an interesting way: for other-race faces, where sensitivity was the lowest, bias toward responding "Old" was the greatest. Barkowitz and Brigham (1982), the only published study of the own-race effect to report a bias measure, also obtained this effect. This pattern suggests that sensitivity aids us more in discerning that a face has not been seen before, than in helping us to recognize the faces we have previously seen. If true, this notion has important implications for face perception and memory. It implies that unskilled perceptual processes tend to homogenize faces by omitting detail that might otherwise serve to distinguish between them, as might happen if we perceived new faces most readily in terms of their familiar attributes. Memory, therefore, captures primarily those aspects of faces that fit a prototypical or composite representation. The rest of the information available in a face is unabsorbed unless the face deviates substantially from the composite, that is, unless the face is "distinctive". When presented with an other-race face to identify as old or new, we compare our homogenized perception of it with our memory of earlier homogenized perceptions and, more often than not, find a match. Greater sensitivity would be associated with less perceptual distortion, and more diverse memory representations.

Alternatively, our perceptual processes might permit a veridical representation of an other-race face, but memory might fail to preserve the representation in as detailed a form as it would for an own-race face. Here, we must assume that we develop a kind of mental shorthand, possibly based upon knowledge of how faces vary, that allows us to store information about own-race faces more efficiently. When we are asked if we recognize an other-race face, we compare our accurate perception of it with our sketchy recollections of faces previously seen. Because our memory lacks sufficient detail to distinguish remembered faces from the current one, we say "yes". To increase sensitivity, we must expand the mental shorthand to accommodate other-race faces.

Face Prototypes

Whether memory, perception, or both are at fault, the result is the same, a high rate of false positives due to an over-emphasis on central tendencies, rather than the accurate representation of a face in memory. This type of phenomenon is not unknown: in studies of category learning, subjects are disposed toward extracting prototypes that embody the central tendencies of observed exemplars (Hayes-Roth & Hayes-Roth, 1977; Posner & Keele, 1970) and have trouble telling novel exemplars that fit the prototype from those they have previously seen. Hayes-Roth and Hayes-Roth had subjects study lists of descriptions of people that contained information about age, education, and marital status. When asked to identify the descriptions they had studied, the subjects tended to select a description that contained the modal values of each attribute, even though such a description was never actually studied. Subjects clearly retain information about specific descriptions, and specific faces -- they are more likely to say they recognize old ones than new ones -- but there is a tendency for this information to be blended in a way that leads them to "recognize" prototypical stimuli they have never seen before.

Valentine and Bruce (1986b) framed an argument for the existence of a face prototype based on reaction times to photographs of distinctive and typical faces. They asked subjects to determine if a photograph was that of a normal or scrambled face. Reaction time was much longer for normal faces that had been previously rated as distinctive. Valentine and Bruce argue that it takes longer to realize that these faces are actually faces because, being unusual, they are more distant from an internal prototype. Further, they suggest that faces are encoded with reference to this prototype. When they asked subjects if a face was familiar (i.e., the face of a celebrity), reaction time was shorter for distinctive faces (Valentine & Bruce, 1986c). The more unlike other faces a face is, the quicker we can decide that we have seen it before. This implies that we remember a face by its deviation from a prototype. As to how a face prototype is acquired, Valentine and Bruce make no specific claims except that each face updates the prototype. This implies a relatively passive process, the prototype is formed by mere exposure.

The problem with this conception of memory for faces is that, as it stands, it affords no explanation of the own-race effect, indeed, it predicts the opposite. If we suppose, as Valentine and Bruce claim, that we have a single face prototype, it will be biased toward own-race faces, which we see more often. Other-race faces should be easier to recognize in that case, because they differ more from the prototype. Should we assume that we develop race-specific prototypes, we would still have no basis for predicting superior own-race recognition. They will still differ from the prototype to the same extent that own-race faces do and should be as easy to recognize.

To rescue the notion of prototypes entails making an additional assumption: we must suppose that the other-race prototype is inadequate. As a result, using it as a reference by which to encode new faces and retrieve old ones leads more frequently to error. Why should the prototype for other-race faces be inferior to that for own-race faces? Simply having fewer exemplars contribute to an other-race prototype does not make other-race faces less distinctive. The fact that it has been extracted from fewer exemplars could make it relatively unstable. Each successive face should have a smaller influence on the prototype than the last one, but in its early stages of formation, each new face will alter the prototype substantially. Being more subject to modification, the other-race prototype is a poor standard for face memory. It is doubtful that this factor could be the sole source of the problem, however. The adults that participate in research on the own-race effect, particularly the New York City residents that served as subjects in Experiment 1, have been exposed to plenty of other-race faces, and the differences in the stabilty of their prototypes are not likely to be large enough to produce the differences in recognition that have been found. To fully account for the own-race effect we must return to the earlier suggestion that own-race faces benefit from superior perceptual and/or encoding processes. Because the prototype is a presumed product of these processes, it then stands to reason that the own-race prototype should also be superior. By now, however, we are far afield of Valentine and Bruce's conception of the face prototype.

Perception, Face Concepts, and Experience

Whether or not something like a prototype actually plays a role in face perception, it is true that, at least on a basic level, our knowledge of faces enhances perception. This is clear not only from the developmental research cited above, but from the "face superiority" effect explored by Homa, Haver and Schwartz (1976) and the related "face detection" effect noted by Purcell and Stewart (1986). In the face superiority effect, subjects are better able to identify a facial feature if it is part of a schematic face than if it is embedded in an arbitrarily arranged pattern of the same components. The face detection effect refers to a decrease in duration threshold for patterned stimuli when they are arranged to form a schematic face. When a stimulus is familiar, the subject's expectations and knowledge about its structure assists him in detecting it and recognizing its components.

Both of these experiments were concerned with a very crude level of "faceness", a configuration of eyes, nose, and a mouth in an oval, and yet they were able to demonstrate that prior experience with faces enhanced perception. It can be argued that the artificiality of the schematic faces makes the result all the more noteworthy -- it shows that the basic pattern of a face is so powerful a stimulus that even a degraded version will benefit from its familiarity.

What is important for the present discussion is whether more subtle attributes of faces can have a similar influence on perception. Suppose, for example, that black faces tend to be broad and have high cheekbones. Could those attributes be incorporated into a black person's concept of "faceness"? If so, perception of faces that conformed to that mold would be facilitated, while faces without those features would be made more difficult. There is no reason to limit the number of face concepts to one. If experience indicates that faces fall into different types, such as broad faces and narrow faces, wide-eyed faces and narrow-eyed faces, or white faces and black faces, then a different schema might develop for each. Face concepts would vary in their efficiency based on our familiarity with each type of face. Efficient concepts must involve attributes that divide faces into naturally-occuring categories, and they must be relatively specific. If a broad face is no more likely to have high cheekbones than low, then a concept involving the co-occurance of high cheekbones and broad faces is useless. Further, if a common configuration involved a weak chin, thin lips and an overbite, a schema that included only two of those features would not be optimal. One might argue that this complicates matters unreasonably, because multiple schemas entail a mechanism for schema selection, but we must not neglect the fact that some process is already responsible for activating a face concept instead of say, a chess board concept. The mere existence of the face superiority and detection effects shows that some selection is occurring.

Whatever the notion of face concepts might lack in simplicity, it makes up for in its ability to predict the own-race effect. Concepts are formed, based on faces previously seen, to facilitate recognition of those faces.

Therefore our concepts will concern the types of faces seen most often and that it is most important to remember, generally own-race faces. If perception of faces does proceed from the stimulation of a face concept, then we should have more difficulty in detecting detail in other-race faces than we do with own-race faces. Poorer perception can only lead to poorer recognition.

Impoverished perception of other-race faces is predicted by some of the explanations of the own-race effect. Prejudice and familiarity are supposed to operate by increasing attention or directing it to relevant aspects of a face, and not by producing a more durable memory trace. One of the versions of the motivation hypothesis also faults perception of other-race faces. If any of these hypotheses is true, then we should expect to find an own-race effect in a task that involves perceptual discrimination, without placing a demand on memory. Subjects in Experiment 2 were presented with such a task.

Lateral Reversal

Inversion is only one transformation that experimenters have applied to faces: lateral reversal has also received some attention. Although we ordinarily think of faces as symmetrical, they are not. The subtle differences between the two halves of a face can be made astonishingly clear through the use of chimeric photographs, artificial images created by removing one half of the face and doubling the other. The two chimeric photographs that can be derived from a face are quite different, to the extent that one of them, the right composite, looks more like the original face than the other (Gilbert & Bakan, 1973), and subjects ascribe different emotive qualities to the two images (Karch & Grant, 1978). When viewing normal images, however, the asymmetries are not salient, and it can be difficult to tell if a familiar face has been laterally reversed. Our capacity to do so hinges on perception and encoding of finely detailed information about a face, a level of detail beyond what is necessary for recognition.

Klatzky and Forrest (1984) showed subjects slides of target faces (celebrities and novel faces) and then tested their recognition memory with a set of slides that contained both reversed and non-reversed (normal) targets as well as distractors. The subjects were asked if the test slides were old or new, and, if old, whether or not they were reversed. Klatzky and Forrest found that normal targets were no more likely to be recognized than reversed ones, and that subjects were able to tell, at a level significantly above chance, which faces had been transformed. This suggests that we remember lateral organization but that it is independent of our judgment that the face has been seen before. On the other hand, McKelvie (1983; 1987), using essentially the same procedure, found that reversal did adversely affect recognition but claimed that his subjects were not able to identify which faces had been reversed. McKelvie's result indicates that lateral organization is not recalled but does play a role in perceived familiarity.

The stimulus sets used by McKelvie and Klatzky and Forrest differed in an important way that may explain this discrepancy. McKelvie used both full-face and three-quarter views, while Klatzky and Forrest used only three-quarter views. McKelvie (1983) turned up a surprising finding: reversal affected the recognition of full-face photographs more than that of three-quarter view photographs. Because a full-face photograph is more symmetrical, reversal results in a less obvious transformation, and one would expect that reversal of such photographs should have less, not more, of an impact. McKelvie's finding is counter-intuitive, but it reconciles his results with those of Klatzky and Forrest, who employed a stimulus set that was relatively immune to the negative effects of lateral reversal on recognition. The difference in stimulus sets also offers an explanation for the inconsistency in the accuracy of identification of orientation. McKelvie included faces that were less obviously asymmetrical and therefore his subjects had a harder time discerning that they were reversed. Perhaps this strange pair of findings can be reconciled in the following way. Klatzky and Forrest's subjects' had greater awareness of which faces were reversed. This helped them to correctly identify old faces that had been reversed because they could compensate for the reversal in their recognition judgment. To McKelvie's subjects the old but reversed faces simply looked unfamiliar because there were fewer cues to indicate that they should attempt to mentally reverse the faces.

Klatzky and Forrest, who found orientation identification performance on famous and novel faces to be identical, conclude that familiarity with a face does not increase the availability of "concrete visual information" (1984, p. 63). As we become more familiar with a particular face, we gain knowledge about it that facilitates recognition without providing additional information about asymmetry. McKelvie, on the other hand, concludes that "lateral orientation of a photographed face ... contributes to its memory representation" (1983, p. 405).

Neither opinion is warranted by the data upon which it is based. Klatzky and Forrest and McKelvie used the same photograph for the initial exposure and test phases of their experiments. Therefore, the subject's task was to recognize a photograph, not a face. This distinction is at least as crucial in orientation identification as it is in recognition. In both cases, memory for pose or clothing is sufficient for an accurate judgment. Information about pose may be part of the internal representation of a photograph but it is not an intrinsic property of a face or its representation.

That faces seem so symmetrical when, in fact, they are not, implies that our perceptual processes impose symmetricality on them by blurring assymetrical detail. This would occur if we employed a face concept that contained, in addition to information about the locations of major features, the expectation that faces be symmetrical. Because we are not ordinarily called upon to make judgments about the lateral organization of faces, and faces do not normally deviate too far from the symmetric, such a concept would be reasonably efficient. It would simplify our representations of faces without distorting them too much or reducing their utility. This notion predicts that subjects would be biased toward responding "Unreversed" in an orientation identification task, a prediction borne out by the data of McKelvie and Klatzky and Forrest. It must be recognized, though, that a more accurate, but less efficient, face concept would incorporate information about the ways in which faces are asymmetrical.

The question of whether lateral orientation is part of our representation of faces can be sensibly asked only with respect to the lateral organization of the face itself. Here the issue is whether we remember that a left eyebrow is higher or bushier than the right one, and not that a profile was facing left instead of right. The former is a question of memory for a property of a face, the latter is solely concerned with memory for pose. Therefore, only research involving full frontal views of faces is germane.

Using such stimuli, McKelvie has shown that lateral organization affects our recognition of faces. He found that correct recognition of reversed and unreversed faces differed by nearly 10 percent. This indicates that some information about lateral orientation is encoded. It reduced the perceived familiarity of the reversed faces to the extent that 10 percent fewer of them were recognized.

McKelvie claims that the subjects were unable to identify the faces that had been reversed. He conducted separate t tests on proportion correct for reversed and unreversed faces, comparing them to chance (.5). He found a significant difference for unreversed faces, but not for reversed ones. This isn't suprising given that the subjects were biased toward responding "Unreversed", the correct response for unreversed faces. A closer look at his data shows that when proportion correct is combined across the two responses, it ranges from .58 to .61 in the four experiments in which this measure was taken. In other words, he not only found that his subjects performed about 10% better than chance, but he replicated the finding several times. From this we can conclude that, for approximately 10% of the faces we see, we can recall information about lateral organization. For another 10 percent, the amount of information we retain about lateral organization is enough to make a reversed face look unfamiliar, but not sufficient to trigger an awareness that the face is reversed. As far as the majority of faces goes, information about lateral organization is not preserved.

Experiment 2

Lateral reversal can be exploited to inform us about how much we perceive and remember of a face. If we can correctly judge the lateral orientation of a face, we have formed a representation of that face that is more detailed or more accurate than the representations of most faces. Given the assumption that perception of own-race faces is more veridical, we should expect greater accuracy in judging the lateral orientation of own-race faces. Experiment 2 presented subjects with the task of judging lateral orientation in own- and other-race faces. It was intended that this task be primarily a perceptual one. To this end, the subjects saw two photographs of the same person simultaneously, and made their judgment while the photographs were being inspected. Memory was involved only to the extent that they had to remember one photograph while they shifted their gaze to the other. If the predicted own-race advantage in orientation discrimination were found, it would only have relevance for the own-race effect in recognition if the advantage were found to persist when the judgment was made from memory. Experiment 3 addresses this issue.

A face concept facilitates detection of faces and perception of their features by providing an organizational framework for the interpretation of a visual stimulus. If that framework includes the expectation that faces are symmetrical then our perceptions should be biased toward symmetry. It is therefore predicted that orientation judgments will be subject to a bias toward responding Same.

The pool of faces from which the stimuli for Experiment 1 were drawn was used for Experiment 2 and 3. It meets the relevant requirements: several full-face photographs of each person wearing two different sets of clothing. It was assumed that an obvious asymmetry in the hair would provide a basis for orientation discrimination that obviated attention to the facial features themselves. Therefore, two sets of faces were used. Set 1 was selected on the basis of symmetry of the hair. None of the faces had an obvious part or irregularity in the hairline. It should be noted that only 34 such faces could be found in a pool of over 100, and that more of them were black (19) than white (15). Set 2 comprised 34 faces chosen randomly from among those that appeared in the inverted orientation in Experiment 1, with the constraint that half be black and half white. A comparison of inverted recognition performance with orientation discrimination performance was thought to be potentially interesting because, as McKelvie (1983) pointed out, inverting a slide by rotating it also involves a lateral reversal. It was expected that performance would be lower on Set 1 than Set 2 since one source of asymmetry was eliminated from the former.

Method

Subjects. The subjects were 43 Brooklyn College students who participated in the experiment in partial fulfillment of a course requirement and 4 volunteers who were paid $4. Sixteen were black, 27 were white, 1 was Asian, and 3 were Hispanic. Race was determined by self-report.

Stimuli. The stimuli were drawn from the same pool of slides from which the stimuli for Experiment 1 were taken. Two sets of 34 faces were chosen. Set 1 included only faces with symmetric hair-styles. This drastically reduced the pool of usable faces from over 100 to 34 (19 blacks and 15 whites). Set 2 consisted of 17 white and 17 black faces chosen randomly, and without regard to symmetry, from among the faces that appeared in the inverted orientation in Experiment 1. The two sets overlapped partially. For each face, the jacket/no-smile and no-jacket/no smile slides were used.

Procedure. The subject's task was to decide if two simultaneously presented slides of the same person were in the same left-right orientation or if one of the slides was reversed. For each set, the slides were arranged in two carousels, one containing the no-jacket/no-smile photographs and the other containing the jacket/no-smile photographs. In 17 of the pairs, the Different pairs, one of the slides had been turned around so that a mirror image of the actual photograph was projected. The rest were Same pairs in which both slides were presented in the normal orientation. Approximately half of the pairs in each race were Different. None of the pairs contained two reversed slides. In 8 of the Different pairs the jacket slide was reversed, and in the rest the no-jacket slide was reversed. In some of the no-jacket slides the man photographed is wearing a tee shirt with writing across it. These slides were always presented in the normal orientation, although they were not always part of Same pairs since the jacket slide may have been reversed. The slides were placed in random order with the restriction that no more than three consecutive pairs were of the same race or had the same orientation status. Each subject saw only one set. For approximately half of the subjects who viewed a given set, the slides were presented in reverse order.

The subjects were run individually or in groups of up to 10. All were seated within 10 feet of a large screen which insured that they had an adequate view of the stimuli. Great care was taken in explaining the task to the subjects. They were shown diagrams of schematic faces and letters to make it clear what left-right reversal meant and they were told that the mirror image was created by turning the slides around. The subjects filled out an answer sheet with columns marked Same and Different while the slides were on the screen. They were allowed as much viewing time as they needed to make their responses. After the last pair was presented the purpose of the experiment was explained to them and they were asked to indicate their race on the answer sheet.

At the end of the session, the subjects were asked what strategies they had used in making their decisions. This was handled very informally (no notes or records were made of the responses) and was intended primarily to insure that the subjects had understood the task and followed instructions. Every subject was asked if he had used specific features to make his judgments, and if so which ones, or if he had tried to make the comparison by forming a general impression of the face.

Results

Table 13 lists the experimental variables and their abbreviations. They will be referred to in the manner followed in Experiment 1. Observed frequencies for each Group in each condition are reported in Table 14. The data were first analyzed for any interactions between Group and Response or Group and sensitivity (RS). Chi-square values and standardized lambdas for these effects are shown in Table 15. Contrary to expectations, Set 1 was no more difficult than Set 2, as evidenced by the absence of an interaction between sensitivity and Group (RSG). Only one interaction was significant, that between sensitivity, Person, and Group (RSPG). This effect reflected the superior performance by the black subjects on Set 1 and by the white subjects on Set 2. It should be noted, however, that the interaction was significant by the goodness-of-fit test alone. Because of this, and because of the absence of

Table 13

Abbreviations for Variables in Experiments 2 and 3.

Name	Variable	Levels
(R)esponse	Subject's Response	Same or Different
(S)tatus	Orientation of Slide	Both Normal or One Reversed
(F)ace	Race of Slide	Black or White
(P)erson	Race of Subject	Black or White
(G)roup	Set of Slides	Set 1 or Set 2

Table 14

Observed Frequencies of Same and Different Responses in Experiment 2.

Group	Person	Face	Status	Response: Same	Diff	Total
One	White N = 10	White	Same	73	7	80
			Different	8	62	70
			Total	81	69	150
		Black	Same	75	15	90
			Different	22	78	100
			Total	97	93	190
	Black N = 7	White	Same	50	6	56
			Different	4	45	49
			Total	54	51	105
		Black	Same	51	12	63
			Different	15	55	70
			Total	66	67	133
Two	White N = 17	White	Same	141	12	153
			Different	10	126	136
			Total	151	138	289
		Black	Same	122	14	136
			Different	43	110	153
			Total	165	124	289
	Black N = 9	White	Same	70	11	81
			Different	10	62	72
			Total	80	73	153
		Black	Same	57	15	72
			Different	25	56	81
			Total	82	71	153

Table 15

Likelihood Ratio Chi-Square Values and Standardized Lambdas for Interactions with Group in Experiment 2.

Effect	X^2	p	Lambda	p
RSFPG	0.23	n.s.	0.481	n.s.
RSPG	2.84	n.s.	1.739	n.s.
RSFG	0.03	n.s.	0.087	n.s.
RSG	0.03	n.s.	0.480	n.s.
RPG	0.07	n.s.	0.145	n.s.
RFG	1.19	n.s.	0.983	n.s.
RG	2.31	n.s.	0.922	n.s.

Note. There is one degree of freedom for each effect.

Table 16

Likelihood Ratio Chi-Square Values and Standardized Lambdas Lambdas in Experiment 2, Collapsed across Group.

Effect	X^2	p	Lambda	p
RSFP	0.11	n.s.	0.327	n.s.
RSF	30.46	<.001	5.286	<.001
RSP	4.45	<.05.	2.118	<.05
RFP	0.28	n.s.	0.523	n.s.
RP	2.44	n.s.	1.009	n.s.
RS	777.37	<.001	22.174	<.001
R	5.54	<.05	1.906	n.s.

Note. There is one degree of freedom for each effect.

any other interaction, the data were collapsed across Group for the rest of the analysis.

Sensitivity. Chi-square values and standardized lambdas for the collapsed data are shown in Table 16. There were two unexpected effects: a tendency for white subjects to out-perform the black subjects (RSP), and a strong advantage for white faces (RSF). No evidence of an own-race effect (RSFP) was found.

Response bias. The predicted bias toward *Same* responses was significant only by the chi-square test, and must be regarded as marginal. Two bias effects that were obtained in Experiment 1, greater bias among the black subjects (RP) and greater bias with other-race faces (RFP) were not obtained here (see Table 16).

The interaction between bias and race of face (RF) was evaluated separately because the proportion of *Same* white pairs was greater than the proportion of *Same* black pairs, and therefore the correct expected frequencies could not be produced by the marginals. A Pearson chi-square test revealed a significant response bias for the black faces, X^2 (1 *df*, *N* = 765) = 12.34, $p < .01$, but not for the white faces, X^2 (1 *df*, *N* = 697) = 0.07, $p > .05$.

Discussion

Of the three predictions made, only one was even marginally borne out by the data: the bias toward responding Same. It is particularly puzzling that Set 1 was not, as expected, more difficult than Set 2. If the subjects were no less accurate with faces with symmetrical hair, then it stands to reason that they were ignoring hair as a cue. This is consistent with their reports: when asked about their strategies, the subjects rarely mentioned hair. It is not clear why such an informative cue should be ignored. The literature on cue saliency is of no help in understanding this result because how salient a feature is seems to be determined by how saliency is assessed. Consider, for example, these three findings: when subjects are asked to describe a face, hair is the most frequently mentioned feature (Shepherd, Davies, & Ellis, 1981); eye movement studies show that more fixations are made on the eyes, nose, and mouth (Walker-Smith, Gale, & Findlay, 1977); the nose is essentially irrelevant in identification tasks (Haig, 1986). Suggesting that the central features are fixated more often because they are more finely structured, and consequently hard to describe, might reconcile these findings, but it does not tell us why the hair should be overlooked in an orientation discrimination task.

An advantage for own-race faces, the third prediction, was not found. In and of itself, this result suggests that the own-race effect in recognition is not the product of an inability to "see" other-race faces. The detail is there, we simply fail to preserve it in memory. Had there not been an interaction between race of face and sensitivity, the next logical step would have been to look for an own-race effect in an orientation recognition task. One might object that orientation discrimination is an artificial task, that the processes it elicits are not like those in recognition, and therefore one cannot predict performance on one from the other. Both Klatzky and Forrest and McKelvie have shown, however, that we do remember orientation to some extent, even when we are not aware that we will be asked about it. Lateral organization is a natural part of some representations of faces.

Interpretation of the results is complicated by the unexpected finding that all subjects had greater difficulty with the black faces. Because black faces are, or seem to be, more symmetrical than white faces, it is not clear what form an own-race advantage would take. Should we still expect blacks to be better at black faces, or ignore performance on black faces and simply expect whites to be more accurate than blacks on white faces? To confuse matters further, the white subjects were more sensitive on both black and white faces. Salvaging the notion of an own-race effect in perception requires that one argue that the advantage for white faces is so great that it swamped the advantage for own-race faces. This is conceivable, but we have yet to find evidence that there is an own-race effect in perception.

It seems that it was the characteristics of the stimulus, rather than any race-specific aspect of the viewer's perceptual processes, that determined performance. In Experiment 1, we saw evidence that faces had qualities that determined their likelihood of being recognized that were independent of the race of the viewer. Taken together, these results suggest that whites and blacks do not perceive faces differently. Indirect evidence for this idea comes from a study of ratings of attractiveness. Using black, white, and Chinese faces and subjects, Bernstein, Lin, and McClellan (1982) found that cross- and within-race judgments of attractiveness were equally variable, which presumably reflects equality of perceived variation.

It is possible to construct versions of the prejudice and motivation hypotheses that suggest the relevant factor in the own-race effect is the amount of attention we pay when we are exposed to a face, and not our history of attention to other-race faces. Other things being equal, we should be as adept at perceiving other-race faces as we are at own-race faces. As we saw earlier, tests of this notion have produced mixed results: it is supported by the temporary "training" effects (e.g. Elliott et al., 1973), but not by Barkowitz and Brigham's (1982) direct attempt to increase motivation. This theory was not tested in Experiment 2, because the task demanded attention to the faces that the subjects might not ordinarily have paid. It is still reasonable to speculate that, although we are all capable of extracting the same kind and amount of information from a given face, we don't. We remember more own-race faces because we are able to create more detailed or accurate representations of them.

Experiment 2 does not differentiate between this notion, and a separate but similar one: it is memory, and not perception, that homogenizes other-race faces. The second hypothesis was one of two possibilities suggested by the high false alarm rate for other-race faces in Experiment 1. Both predict equal discriminability of own- and other-race faces, given equal attention, as in Experiment 2, but better memory for own-race faces when subjects allocate their attention as they choose, as in Experiment 1. The homogenization hypothesis is more attractive in that it alone explains why response bias should be greatest where sensitivity is lowest. This pattern, observed in Experiment 1, was continued in Experiment 2: bias was greater for black faces.

Another common thread in the results of Experiments 1 and 2 is the poor performance on black faces. In the first experiment this was restricted to inverted black faces, which hints that there may be a link between the two transformations, inversion and lateral reversal. McKelvie, as noted above, proposed such a connection, but with reference to recognition of a face that has been rotated 180 degrees between exposures. Such a face is not only upside down but has been laterally reversed. That was not an issue in Experiment 1 because the faces were presented in the same orientation in the study and test phases. Here, the potential link is in the strategies the subjects used. When we are trying to recognize inverted faces, or when we are trying to tell if a face is laterally reversed, there is a tendency to search for distinguishing features. Every subject in Experiment 2 named at least one feature that he had relied on in his judgments and none admitted to having used a "general impression" strategy. If subjects in an inversion experiment rely on distinguishing features for recognition, as Yin (1969) and Phillips and Rawles (1979) claim, then black faces might prove more difficult because they tend to lack one such feature: an obvious asymmetry.

In order to blame poor recognition of inverted black faces on symmetry, it must be shown that the advantage for white faces in a lateral reversal task persists when the judgment is made from memory. If it does not, we cannot conclude that remembering asymmetry gives inverted white faces their advantage in recognition. Experiment 3 addresses this issue.

Experiment 3

The procedures of Experiment 2 and 3 are identical with the exception that the subjects are shown one slide of each of the 34 pairs for five seconds each. Only after all the faces have been seen are they presented with the second slide of each pair and asked if that slide is reversed with respect to the other slide in that pair.

Bartlett, Gernsbacher and Till (1987) have proposed a model to explain the behavior of subjects in an experiment on memory for lateral orientation identical to those of McKelvie and Klatzky and Forrest with the exception that Bartlett et al. used photographs of landscapes. They claim that New, Old/Reversed, and Old/Unreversed pictures form three separate distributions along a dimension of familiarity. The reversed and unreversed distributions are very close together toward the high end of the scale, but relatively remote from the new distribution. Two criteria are established. Stimuli that fall below the lower criterion are judged New. If it falls above the first criterion the subject engages in image-sampling, a process that entails a "conscious recollection" of the original stimuli (p. 28), to determine whether or not the stimulus is reversed. Should image-sampling prove inconclusive, a second check on familiarity is made. If the stimulus falls above the second criterion it is judged Old/Unreversed, otherwise a guess is made as to its orientation.

Given response rules like these, removing the New distribution should have no effect on the probability of a correct response, except to eliminate the possibility of misses or false alarms for the old-new judgment. The Old/Reversed and Old/Unreversed distributions still remain the same distance apart on the familiarity axis. Therefore, it is not expected that differences in procedure between Experiment 3 and that of McKelvie will result in a higher level of performance here. In fact performance is likely to be somewhat poorer because the stimulus set consists exclusively of full-face photographs.

The predicted outcome is a repetition of the superior performance on white faces found in the discrimination task. If obtained, it would confirm the greater perceived symmetry of black faces, and demonstrate that a perceptual advantage is preserved in memory. Most importantly, it would suggest an explanation for the poor recognition of inverted black faces. Because orientation discrimination did not interact with race of subject, and no such interaction is expected in this experiment, it was not considered necessary to have comparable numbers of black and white subjects.

Method

Subjects. Thirty-six Brooklyn College students participated in the experiment in partial fulfillment of a course requirement. Twenty (17 whites and 3 blacks) of them viewed Set 1, and 16 saw Set 2 (13 whites, 2 blacks, and 1 Asian). Data from one of the white subjects who saw Set 2 were discarded because he failed to follow instructions. The Asian subject will be omitted from the analysis except where noted. An additional 17 Brooklyn College students viewed a subset of Set 2. They included 1 black, 1 Hispanic, and 15 whites. Race was determined by self-report.

Stimuli. The stimuli were the slides used in Experiment 2. Set 1 and Set 2 refer to the same grouping of slides that was used in that experiment.

Procedure. The subject's task and the procedure were the same as in Experiment 2, with the following exceptions. Instead of presenting the slides in the two carousels simultaneously, they were presented sequentially. First, one slide of each face was presented at a rate of 5 seconds each. Immediately after all 34 faces had been seen, the subjects were presented with the matching photographs. While these slides were projected they filled out an answer sheet. They subjects were again permitted as much time as they needed to respond. Approximately half the time the jacket slides were presented first. Because the photographs were arranged in the same order in the two carousels (the same random orders used in Experiment 2), approximately the same amount of time passed between the presentation of the two slides of each pair.

Again, great care was taken in explaining the task to the subjects. This was done in the same way described for Experiment 2 with the addition that the subjects were informed that they would see two different photographs of each face. It was pointed out to them that asymmetries in clothing or position would not help them make their decisions, and it was emphasized that they must search for asymmetries in the face itself. Each session lasted approximately 30 minutes.

<u>Results</u>

Observed frequencies are reported in Table 17 (see Table 10 for a list of the variables and their abbreviations). The data were first inspected for significant interactions of Group and of Person. As Table 18 indicates, there were none, and therefore the data were collapsed across these two variables. Chi-square values and standardized lambdas for the effects of interest are shown in Table 19. The only significant effects were sensitivity (RS) and Response. The subjects showed a marked bias toward <u>Same</u> responses. The expected superiority with white faces (RSF) was a non-significant trend.

Overall, proportion correct was quite low (.53) and only 19 of the 36 subjects scored above the level of random guessing. In an attempt to improve performance, 17 subjects were run using an abbreviated stimulus set, twenty faces (10 white and 10 black) randomly selected from Set 2. Seven of these subjects were tested using the procedure described above, and 10 were tested using a presentation rate of 10 seconds. The attempt to improve performance failed: only 9 of these subjects scored above chance, and the overall proportion correct (.53) was the same as found with the larger stimulus set.

It appeared that the low performance was largely a function of some of the subjects' unwillingness to

Table 17

Observed Frequencies of _Same_ and _Different_ Responses in Experiment 3.

Group	Person	Face	Status	Response		
				Same	Diff	Total
One	White N = 17	White	Same	96	40	136
			Different	66	53	119
			Total	162	93	255
		Black	Same	92	61	153
			Different	109	61	170
			Total	201	122	323
	Black N = 3	White	Same	16	8	24
			Different	12	9	21
			Total	28	17	45
		Black	Same	16	11	27
			Different	12	18	30
			Total	28	29	57
Two	White N = 12	White	Same	57	51	108
			Different	44	52	96
			Total	101	103	204
		Black	Same	60	36	96
			Different	60	48	108
			Total	120	84	204
	Black N = 2	White	Same	10	8	18
			Different	7	9	16
			Total	17	17	34
		Black	Same	7	9	16
			Different	9	9	18
			Total	16	18	34

Table 18

Likelihood Ratio Chi-Square Values and Standardized Lambdas for Interactions with Group and Person in Experiment 3.

Effect	X^2	p	Lambda	p
RSFPG	1.96	n.s.	1.399	n.s.
RFPG	0.01	n.s.	0.076	n.s.
RSPG	0.71	n.s.	0.753	n.s.
RSFP	0.33	n.s.	0.336	n.s.
RPG	0.03	n.s.	0.129	n.s.
RFG	3.15	n.s.	1.216	n.s.
RFP	1.97	n.s.	1.389	n.s.
RSG	0.01	n.s.	0.639	n.s.
RSP	0.30	n.s.	0.258	n.s.
RG	7.76	<.01	1.907	n.s.
RP	2.92	n.s.	1.514	n.s.

Note. There is one degree of freedom for each effect.

Table 19

Likelihood Ratio Chi-Square Values and Standardized Lambdas in Experiment 3, Pooled Across Group and Person.

Effect	X^2	p	Lambda	p
RSF	2.76	n.s.	1.660	n.s.
RS	4.53	<.05	2.228	<.05.
RF	0.56	n.s.	0.728	n.s.
R	31.37	<.001	5.409	<.001

Note. There is one degree of freedom for each effect.

persevere in the face of a relatively difficult task. This conclusion is based on discussions with the subjects. When asked what strategiesthey used, many subjects responded "Oh, I just guessed. It was too hard." or (on several occasions) "I paid attention to the way he tilted his head." (They were specifically warned that the latter strategy would do them no good.)

The data were reanalyzed, omitting the 17 subjects who had scored below chance on both black and white faces, and including the 9 subjects who had scored above chance with the smaller stimulus set. The justification for this was that subjects who showed no sensitivity at all could not possibly show differential sensitivity to different kinds of faces. A total of 36 subjects, including the Asian subject (17 had seen Set 1, 10 had seen Set 2, and 9 had seen the abbreviated Set 2).

Sensitivity. Overall, these subjects had much higher proportions correct (.53 for black faces, and .60 for white faces). This was evidenced in a highly significant sensitivity parameter (RS) and chi-square value, as Table 20 reveals. For this group of subjects, sensitivity was greater for white faces (RSF), replicating the finding of Experiment 2.

Table 20

Likelihood Ratio Chi-Square Values and Standardized Lambdas in Experiment 3, Pooled Across Group and Person, for Above-Chance Performers.

Effect	X^2	p	Lambda	p
RSF	4.43	<.05	2.104	<.05.
RS	20.11	<.001	4.562	<.001
R	0.00	n.s.	0.594	n.s.

Note. There is one degree of freedom for each effect.

Response bias. There was no main effect of response, as can be seen from Table 20. Again, as in Experiment 2 the response bias by race of face interaction had to be evaluated separately. This produced a highly significant Pearson chi-square value for the black faces, X^2 (1 df, N = 583) = 39.74, $p < .001$, but not for the white faces, X^2 (1 df, N = 515) = 3.60, $p > .05$.

Comparison with inverted recognition performance. Percent correct orientation discrimination (from Experiment 2) and identification (from Experiment 3) were calculated for each slide that appeared inverted in Experiment 1 and were compared with percent correct inverted recognition. The correlation between inverted recognition and orientation identification was not significant for the black faces, r (N = 23) = .052, $p > .05$, or for the white faces, r (N = 20) = .135, $p > .05$. The correlation between inverted recognition and orientation discrimination was not significant for black faces, r (N = 23) = .137, $p > .05$, or for white faces, r (N = 20) = .105, $p > .05$. There was also no correlation between orientation identification and discrimination, r (N = 48) = .206, $p > .05$, for black and white faces combined.

Discussion

As expected, performance was no better in Experiment 3 than in McKelvie's (1983) study, indeed, it was worse. This is probably due to the restriction of the stimuli in Experiment 3 to full-face photographs.

Among the subjects who were able to identify reversed faces, there was greater sensitivity for white faces, which demonstrates the perseverance of the advantage for white faces in the discrimination task. That performance declined so much, from approximately 85% correct for the discrimination task, to about 57% correct for identification, affirms that there is a strong tendency for information about lateral organization to be lost in memory.

Once again, there was greater response bias where sensitivity was lowest, for black faces. This is the fourth time this pattern has appeared, if one includes Barkowitz and Brigham (1982). The notion that low sensitivity is associated with a homogenized representation is thus strongly supported.

The absence of correlation between inverted recognition and orientation identification or discrimination is a disappointing aspect of the results. On its face, this finding suggests that symmetry is unrelated to the recognition of inverted faces. Despite this, it is hard to accept that, other things being equal, an inverted face with an obvious asymmetry would not be easier to recognize than a more symmetrical one. It is arguable that symmetry is only one of several factors that make inverted black faces hard to recognize and therefore the correlation between orientation identification and recognition should be weak. The advantage of this argument is that it saves us from having to accept that black faces have two peculiar, but unrelated, qualities: they are more symmetrical (or it is harder to perceive asymmetry in them) and they are hard to recognize when they are inverted. Unfortunately, we have no evidence for a connection between these two findings.

General Discussion

Three experiments have explored the effects of race on perceiving and remembering faces. Unfortunately, and perhaps not surprisingly, their findings resist being tied up neatly into an endorsement of a single theory.

The Own-Race Effect in Recognition

A fair test of the familiarity hypothesis was provided by Experiment 1, and the hypothesis was soundly defeated. Mere exposure to faces does not offer the opportunity to learn to remember them.

Two theories, the prejudice and motivation hypotheses, include versions that claim we are incapable of perceiving other-race faces in as much detail as we can own-race faces. Experiment 2 turned up no justification for this assumption. In fact, it demonstrated that, at least for one task, the viewer's race was irrelevant to the accuracy of his perception. If prejudice and motivation affect recognition of other-race faces, they do so by affecting memory and not perception. While we have no basis on which to reject a memory-based version of the prejudice hypothesis, neither do we any evidence in favor of it. On the other hand, the motivation hypothesis has received some support.

The most promising explanation is the notion that face recognition depends on the kinds of faces we have learned, through a conjunction of need and opportunity. This hypothesis can only be tested with special subject populations. A subject's history of remembering faces is something which, by definition, he walks into the experiment with, and cannot be manipulated. Interesting experiments might involve members of mixed-race families, or as mentioned above, teachers whose students are of other races. One would not expect such subjects to show an own-race effect.

The Organization of Memory for Faces

Several ideas about memory for faces have been discussed that revolve around the notion of internal representations of faces. Face prototypes, as outlined by Valentine and Bruce (1986b), can account for the effects of distinctiveness on certain kinds of judgments, they offer no explanation of the own-race effect in recognition. Face concepts, on the other hand, are attractive for two reasons. First, we have evidence, in the face superiority and face detection effects, of the existence of powerful, although crude, face concepts. It is reasonable to assume that the organization of memory for faces would not stop at this rudimentary level. Second, if we assume that the content of our face concepts reflects the kinds of faces we have seen and needed to recognize, then face concepts have the advantage of predicting the own-race effect. The problem with face concepts is that they also predict that we should perceive own-race faces more accurately given that our face concepts are biased to reflect own-race faces. This prediction has not received support.

The data of Experiments 1 and 2 show that, to a great extent, it is the physical qualities of a face, and not the race of the viewer, that determine whether it will be remembered and what will be perceived. Of course, we remember more own-race than other-race faces, but black and white subjects are likely to remember the same faces. This is not true of orientation: the likelihood that a face will be recognized when upright is unrelated to the likelihood of recognizing the same face when inverted. Faces do not have properties that make them memorable that are independent of orientation. Here we see that orientation is a more fundamental organizational property than race. This is echoed by the fact that orientation had a greater effect on sensitivity than the interaction between race of subject and race of face in Experiment 1.

One effect that was observed in all three experiments was the link between response bias and sensitivity. Where sensitivity was low, subjects seemed to feel that they had seen the face before. This is consistent with the idea that our memory holds a more generalized representation of faces for which sensitivity is low, a notion that has been referred to here as homogenization. Gibson and Gibson (1954) observed the same phenomenon in recognition of nonsense stimuli, and described increasing sensitivity as a process of "perceptual differentiation," of learning to respond differently to stimuli that initially seemed to be the same. They argue that their subjects do not learn to see more, but to distinguish between what they already see. The task Gibson and Gibson presented their subjects with was a memory task however, not one of simultaneous discrimination, which would be more appropriate for drawing conclusions about perception. We have seen no evidence in the present experiments for perceptual differentiation, but the response bias data do indicate that better memory is associated with differentiation of images in memory.

This differentiation could be a simple process of focusing attention on aspects of the stimulus that are most informative. If we assume that the most informative aspects of own- and other-race faces differ, then we can explain the own-race effect as misplaced attention: we center our attention on the parts of the face most relevant to own-race recognition, regardless of the face. This assumption has two implications: first, that by directing attention to the most informative aspects of other-race faces we should be able to eradicate the own-race effect, a testable hypothesis; Second, that the own-race effect is due to different viewing strategies on the part of black and white subjects, a notion that the present research has failed to support.

Alternatively, what Gibson and Gibson call perceptual differentiation might be the result of the development, with experience, of a more efficient means of encoding information about faces, an improvement of the mental shorthand, that results in a closer correspondence between perception and memory. This is essentially the homogenization hypothesis: faces are homogenized in memory if we are poorly equiped to remember them. To explain the own-race effect in this way presupposes that our experience in encoding own-race faces is less useful with other-race faces. Although this assumption not unreasonable, it is not clear how to test it.

Some Caveats

Strictly speaking, the conclusions offered here are intended to explain why subjects in face recognition experiments tend to remember more pictures of own-race faces than other-race faces. This is disappointing if what we would really like to know is why, or if, we are more likely to recognize people of our own race than those of other races. We can recognize people by voice, by body shape, by their dress, by the way they walk, even by the sound of their footsteps. When we fail to recognize someone, all of these factors are at fault. We do not know how these sources of information interact with face recognition.

Is recognizing faces in a psychology experiment different from recognizing people we have seen on the subway? Yes, in many ways, it is. In the laboratory, a subject sees a series of photographs of disembodied faces, stimuli that are far removed from their "real life" counterparts: faces attached to people going about their business. Normally, much contextual information is available when we look at a face, for example, what the person is doing, where he is doing it, how we feel at the moment, the nature of our interaction with that person. This context can form a framework of associations with the visual image of the person's face that can serve as cues for recognizing the face (Winograd & Rivers-Bulkeley, 1977). In the laboratory, the contextual information to which we are accustomed is considered by the experimenter to be a confounding variable, and it is removed in the interest of studying "pure" face recognition. All the faces the subject sees have essentially the same context, "seen in a psychology experiment," although serial position might offer some variations on this theme. Thus, the experimenter provides an artificial context that minimizes the differences between faces, and might make them harder to distinguish from one another. He also keeps the study and test contexts the same, which should make the faces easier to recognize. We do not know what effect these factors have on face recognition.

Baddeley and his colleagues have begun to tackle the issue of the relationship between face recognition in and out of the laboratory. They have shown that one factor that produces reliable effects on recognition in laboratory studies, pose of target photograph, is virtually irrelevant in natural settings (Logie, Baddeley, & Woodhead, 1987). They concluded that the effects of pose are negligible in comparison to factors like motivation. In addition, they showed that recognition of live targets in natural settings from photographs was far lower than the typical 90 percent correct obtained in laboratory experiments. In one study, for example, 33 subjects searched a public area for 6 targets. Out of a possible 198 correct recognitions, there were only 2.

In Experiment 1, race was shown to produce a small but significant effect on recognition. It is impossible to predict how large an effect race has on recognition in real life. We do not know the extent to which recognition is normally based on the face alone, the role that context plays, or the relationship between laboratory and natural face recognition. It may be that the laboratory is a neutralizing environment that mutes the social connotations of race.

These cautions are not meant to imply that laboratory work on the own-race effect does not contribute to our knowledge of behavior in natural settings. One must bear in mind that the own-race effect is not created by the experimenter. Its manifestation in the laboratory simply confirms a difference in sensitivity that was noted long ago by people who realized there were certain kinds of people they had trouble recognizing. Applied research is necessary, however, before we can say how great an effect race has on the recognition of people.

References

Baddeley, A.D. (1979). Applied cognitive and cognitive applied psychology: The case of face recognition. In L.G. Nilsson (Ed.), Perspectives on Memory Research. New Jersey: L. Erlbaum.

Bartlett, J.C., Gernsbacher, M.A., & Till, R.E. (1987). Remembering left-right orientation of pictures. Journal of Experimental Psychology: Learning, Memory, and, Cognition, 13, 27-35.

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

Bernstein, I.H., Lin, T.-D., & McClellan, P. (1982). Cross- vs. within racial judgments of attractiveness. Perception & Psychophysics, 32, 495-503.

Bower, G.H. & Karlin, M.B. (1974). Depth of processing pictures of faces and recognition memory. Journal of Experimental Psychology, 103, 751-757.

Brigham, J.C., & Barkowitz, P. (1978). Do "They all look alike?" The effect of race, sex, experience and attitudes on the ability to recognize faces. Journal of Applied Social Psychology, 8, 306-318.

Brigham, J.C., & Williamson, N.L. (1979). Cross-racial recognition and age: When you're over 60, do they still "All look alike?" Personality and Social Psychology Bulletin, 5, 218-222.

Carey, S., & Diamond, R. (1977). From piecemeal to configurational representation of faces. Science, 195, 312-314.

Carey, S., Diamond, R. & Woods, B. (1980). The development of face recognition -- A maturational component? Developmental Psychology, 16, 257-269.

Carroo, A.W. (1986). Other race recognition: A comparison of black American and African subjects. Perceptual and Motor Skills, 62, 135-138

Chance, J.E., Turner, A.L. & Goldstein, A.G. (1982). Development of differential recognition for own- and other-race faces. The Journal of Psychology, 112, 29-37.

Cross, J.F., Cross, J., & Daly, J. (1971). Sex, race, age and beauty as factors in recognition of faces. Perception & Psychophysics, 10, 393-396.

Craik, F.I.M., & Lockhart, R.S. (1972). Levels of processing: A framework for memory research. Journal of Verbal Learning and Verbal Behavior, 11, 671-684.

Czigler, I. (1985). Matching of facial features: Continuous processing, improper filtering, and holistic comparison. Perception & Psychophysics, 37, 257-265.

Devine, P.G., & Malpass, R.S. (1985). Orienting strategies in differential face recognition. Personality and Social Psychology Bulletin, 11, 33-40.

Diamond, R., & Carey, S. (1986). Why faces are and are not special: An effect of expertise. Journal of Experimental Psychology: General, 115, 107-117.

Elliott, E.S., Wills, E.J., & Goldstein, A.G. (1973). The effects of discrimination training on the recognition of white and oriental faces. Bulletin of the Psychonomic Society, 2, 71-73.

Ellis, H.D. (1975). Recognizing faces. British Journal of Psychology, 66, 409-426.

Ellis, H.D. (1983). The role of the right hemisphere in face perception. In A.W. Young (Ed.), Functions of the Right Cerebral Hemisphere. London: Academic Press.

Ellis, H.D., & Deregowski, J.B. (1981) Within-race and between-race recognition of transformed and untransformed faces. American Journal of Psychology, 94, 27-35.

Galper, R.E. (1973). "Functional race membership" and recognition of faces. Perceptual and Motor Skills, 37, 455-462.

Gibson, J.J., & Gibson, E.J. (1954). Perceptual learning: Differentiation or enrichment? Psychological Review, 62, 1954.

Goldstein, A.G. (1965). Learning of inverted and normally oriented faces in children and adults. Psychonomic Science, 3, 447-448.

Goldstein, A.G. (1975). Recognition of inverted photographs by children and adults. Journal of Genetic Psychology, 127, 109-123.

Goldstein, A.G. (1979a). Race-related variation of facial features: Anthropometric data I. Bulletin of the Psychonomic Society, 13, 187-190.

Goldstein, A.G. (1979b). Race-related variation of facial features: Anthropometric data II. Bulletin of the Psychonomic Society, 13, 191-193.

Goldstein, A.G., & Chance, J.E. (1985). Effects of training on Japanese recognition: Reduction of the other-race effect. Bulletin of the Psychonomic Society, 23, 211-214.

Goldstein, A.G., & Chance, J. (1980). Memory for faces and schema theory. Journal of Psychology, 105, 47-59.

Haig, N.D. (1984). The effect of feature displacement on face recognition. Perception, 13, 505-512.

Harmon, L.D. (1973). The recognition of faces. Scientific American, 229, 70-82.

Hays-Roth, B., & Hays-Roth, F. (1977). Concept learning and the recognition and classification of exemplars. Journal of Verbal Learning and Verbal Behavior, 16, 321-338.

Hecaen, H. (1981). The neuropsychology of face recognition. In G. Davies, H. Ellis & J. Shepherd (Eds.), Perceiving and Remembering Faces. London: Academic Press.

Homa, D., Haver, B., & Schwartz, T. (1976). Perceptibility of schematic face stimuli: Evidence for a perceptual gestalt. Memory & Cognition, 4, 176-185.

Karch, G.R., & Grant, C.W. (1978). Asymmetry in perception of the sides of the human face. Perceptual and Motor Skills, 47, 7627-734.

Klatzky, R.L. (1984). Memory and Awareness, New York: Freeman.

Klatzky, R.L., & Forrest, F.H. (1984). Recognizing familiar and unfamiliar faces. Memory & Cognition, 12, 60-70.

Knoke, D., & Burke, P.J. (1980). Log-Linear Models. Sage University Paper series on Quantitative Applications in the Social Sciences, 07-001. Beverly Hills and London: Sage Publications.

Laughery, K.R., Alexander, J.F., & Lane, A.B. (1971). Recognition of human faces: Effects of target exposure, time, target position, pose position, and type of photograph. Journal of Applied Psychology, 55, 477-483.

Lavrakas, P.J., Buri, J.R., & Mayzner, M.S. (1976). A perspective on the recognition of other-race faces. Perception & Psychophysics, 20, 475-481.

Logie, R.H., Baddeley, A.D., & Woodhead, M.M. (1987). Face recognition, pose, and ecological validity. Applied Cognitive Psychology, 1, 53-69.

Luria, S.M., & Strauss, M.S. (1978). Comparison of eye movements over faces in photographic positives and negatives. Perception, 7, 349-358.

Malpass, R.S. (1981). Training in face recognition. In G. Davies, H. Ellis, & J. Shepherd (Eds.), Perceiving and Remembering Faces. London: Academic Press.

Malpass, R.S., & Kravitz, J. (1969). Recognition for faces of own and other race. Journal Personality & Social Psychology, 13, 330-334.

Malpass, R.S., Lavigueur, H., & Weldon, D.E. (1973). Verbal and visual training in face recognition. Perception & Psychophysics, 14, 285-292.

Navon, D. (1977). Forest before trees: The precedence of global features in visual perception. Cognitive Psychology, 9, 353-383.

Phillips, R.J., & Rawles, R.E. (1979). Recognition of upright and inverted faces: A correlational study. Perception, 8, 577-583.

Posner, M.I., & Keele, S.W. (1970). Retention of abstract ideas. Journal of Experimental Psychology, 83, 304-308.

Purcell, D.G., & Stewart, A.L. (1986). The face-detection effect. Bulletin of the Psychonomic Society, 24, 118-120.

Scapinello, K.F., & Yarmey, A.D. (1970). The role of familiarity and orientation in immediate and delayed recognition of pictorial stimuli. Psychonomic Science, 21, 29-331.

Shepherd, J. (1981). Social factors in face recognition. In G. Davies, H. Ellis, & J. Shepherd (Eds.), Perceiving and Remembering Faces. London: Academic Press.

Strnad, B.N. & Mueller, J.H. (1977). Levels of processing facial recognition memory. Bulletin of the Psychonomic Society, 9, 17-18.

Valentine, T., & Bruce, V. (1986a). The effect of race, inversion, and encoding activity upon face recognition. Acta Psychologica, 61, 259-273.

Valentine, T., & Bruce, V. (1986b). The effects of distinctiveness in recognising and classifying faces. Perception, 15, 525-535.

Valentine, T., & Bruce, V. (1986c). Recognizing familiar faces: The role of distinctiveness and familiarity. Canadian Journal of Psychology, 40, 300-305.

Walker-Smith, G.J., Gale, A.G., & Findlay, J.M. (1977). Eye movement strategies involved in face perception. Perception, 6, 313-326.

Watkins, M.J., Ho, E., and Tulving, E. (1976). Context effects in recognition memory for faces. Journal of Verbal Learning and Verbal Behavior, 15, 505-517.

Winograd, E., (1981). Elaboration and distinctiveness in memory for faces. Journal of Experimental Psychology: Human Learning and Memory, 7, 181-190.

Winograd, E., & Rivers-Bulkeley, N.T. (1977). Effects of changing context on remembering faces. Journal of Experimental Psychology: Human Learning and Memory, 3, 397-405.

Woodhead, M.M., & Baddeley, A.D. (1981). Individual differences and memory for faces, pictures, and words. Memory & Cognition, 9, 368-370.

Yarmey, D.A. (1971). Recognition memory for familiar "public" faces: Effects of orientation and delay. Psychonomic Science, 24, 286-288.

Yin, R.K. (1969). Looking at upside-down faces. Journal of Experimental Psychology, 81, 141-145.

Yin, R.K. (1970). Face recognition by brain-injured patients: A dissociable ability. Neuropsychologia, 8, 395-402.