THE OTHER-RACE EFFECT IN EYEWITNESS IDENTIFICATION
What Do We Do About It?

Gary L. Wells and Elizabeth A. Olson
Iowa State University

The greater difficulty of other-race relative to own-race facial recognition is one of the most researched issues in eyewitness identification. The authors discuss articles in the March 2001 issue of Psychology, Public Policy, and Law as well as some methodological and interpretational problems regarding demonstrations of the other-race identification effect. In spite of some problems with the literature, they conclude that the other-race identification effect is real and that it is relevant to the criminal justice system. The authors propose a system-variable approach in which other-race identification cases receive special preventive treatment at the time of the identification test rather than after-the-fact attempts to assist triers of fact. These preventive procedures include using same-race lineup constructors, larger numbers of fillers in other-race cases, dual lineup techniques in other-race cases, or a combination of these procedures.

The other-race effect hypothesis is simple: Eyewitnesses are less likely to misidentify someone of their own race than they are to misidentify someone of another race.¹ This is no longer just a hypothesis. As the articles in the current issue of this journal make clear, there is a great deal of empirical support for the other-race effect. At this point, it is reasonable to conclude that there is internal validity to the studies showing the other-race effect. Although any particular study might leave some questions about whether the observed effect was due to some flaw in the design or was the result of stimulus sampling problems, the consistency of the findings across a variety of methodologies converges on the conclusion that the other-race effect is real.

We have no particular reason to think that the other-race effect is restricted to the lab and does not apply to eyewitnesses in actual criminal cases. In fact, one of the characteristics of lab studies is that they tend to force the participant-witness to pay equal amounts of attention to the faces, regardless of race. In real cases, eyewitnesses might attend less to faces of another race than to faces of one’s own race, thereby making the other-race effect even stronger in real cases than in the lab studies.

Although we do not know the extent of the other-race problem in actual criminal cases, DNA-exoneration cases suggest that there might be a pattern

¹We prefer to call this the other-race effect rather than other-race bias (or own-race bias) because of the connotation in the word bias that at least partly implies prejudice. The word effect strikes us as being more neutral regarding potential causes.
worthy of closer study. In their analysis of 77 mistaken eyewitness identifications, Scheck, Neufeld, and Dwyer (2000) reported that 35% were cases of White victim–witnesses misidentifying Black suspects, whereas only 28% were cases of White victims misidentifying White suspects. We think that there is some significance to this figure because most White victim–witnesses are victimized by other White people. If most victimization is within race categories rather than across race categories, then it seems to us that White victim–witnesses should misidentify White suspects more often than they misidentify Black suspects. The other-race identification effect is not the only possible explanation for this apparent overrepresentation of other-race misidentifications in the DNA-exoneration cases. Among other possible problems, the DNA-exoneration cases may not be representative of eyewitness misidentification cases in general, and the sample size of DNA-exoneration cases is still relatively modest. Nevertheless, the DNA-exoneration cases represent our best single source of actual misidentifications of criminal suspects to date, and it is noteworthy that other-race misidentifications constitute the largest single category.

The magnitude of the other-race effect obtained in research experiments can be estimated from a set of likelihood ratios that were calculated by Meissner and Brigham (2001). They report in their meta-analysis that the chances of a mistaken identification is 1.56 times greater in other-race than in same-race conditions. Hence, the other-race effect is not merely a statistically reliable effect but is actually a rather large effect in these studies. If we assume that these likelihood ratios are operating in real cases, for instance, then a Black innocent suspect has a 56% greater chance of being misidentified by a White eyewitness than by a Black eyewitness. The chance of an innocent suspect being misidentified can be rather high even when the witness and suspect are of the same race (see Levi, 1998). The fact that race can so dramatically increase the chance of misidentification is disturbing.

Naturally, the question arises as to what the justice system can do to address the other-race-effect problem. As we explain later in this article, we believe that expert testimony is a weak response to the problem. Although perhaps necessary under certain circumstances, expert testimony on the other-race effect is an attempt to postdict an event. We think that there may be ways in which the justice system can prevent the other-race identification error from occurring in the first place.

Although the primary purpose of this article is to address possible justice system responses to the problem of the other-race effect, we first make some observations regarding some of the key articles in this special theme issue of Psychology, Public Policy, and Law. We are particularly interested in why the other-race effect occurs because this might suggest some possible solutions. Indeed, a core reason for trying to establish cause–effect relations in psychology is to gain possible control over the effect. To the extent that the causes of the other-race effect are controllable by the justice system, proposed solutions can follow logically from our understanding of the cause–effect relations. There are successful examples of this approach in the eyewitness identification literature. For example, our understanding that mistaken identifications can be caused by eyewitnesses using relative judgments (Wells, 1984a) led to the development of the sequential lineup procedure (Lindsay & Wells, 1985).
Some General Concerns and Caveats About the Other-Race Effect

Having acknowledged that there is an other-race effect does not mean that we accept it as an omnipresent, invariant fact across all circumstances. The size of this effect is likely to change as a function of a number of variables, such as the extent of the eyewitness's meaningful experience with faces of another race (see Meissner & Brigham, 2001; Sporer, 2001b). Even being a serious National Basketball Association fan can mitigate the other-race identification effect (Li, Dunning, & Malpass, 1998). Furthermore, faces within a given race vary considerably in their recognizability, such that distinctive faces are more recognizable than typical faces (Brigham, 1990; Chiroro & Valentine, 1995; Hosie & Milne, 1995; see meta-analysis by Shapiro & Penrod, 1986). Accordingly, one could show that there are many Black people, for example, who have certain distinctive characteristics that make them more recognizable to White eyewitnesses than is the "typical" White person. There are other moderators of the other-race effect as well. MacLin, MacLin, and Malpass (2001), for instance, showed that the other-race effect is moderated by the extent of delay or time lapse between initial encoding of the face and the time of the recognition test. Fallshore and Schooler (1995) showed that the other-race effect is moderated by having people provide verbal descriptions of the face after viewing the face. There are likely to be other moderators of the other-race effect that we do not yet know about.

It is useful to recognize that the magnitude of the other-race effect is not a single value; there are many possible effect sizes depending on the presence and absence of currently known (and some as-of-yet unknown) moderators. In addition, we have some concerns about many of the experiments that have been used to estimate the size of the effect. Our concerns fall into three major categories. First, we are concerned about problems of stimulus sampling (see Wells & Windschitl, 1999). Many studies of the other-race effect use only a few faces of each race as stimuli. The assumption that a few Black stimulus faces are necessarily representative of the means and variances of the population of Black people in general is not a defensible way to conduct research of this type. Physiognomic variation of faces within race categories raises the problem of how to generate a representative sample of faces to represent an entire racial category. The issue is not simply a question of the number of stimulus faces that are used to represent the race category but also a question of whether these faces are representative of the race they are supposed to represent. Some type of stimulus sampling strategy should be used and described in studies of the other-race effect, because these experiments manipulate the race of the face by sampling instances from a high-variance population of possible faces. This issue of stimulus sampling is particularly problematic in "half-design" other-race identification studies. In a half-design other-race study, only participant–witnesses of one race are used (e.g., all White participant–witnesses) and the race of the stimulus faces is varied (e.g., Black vs. White stimulus faces). Showing that White participant–witnesses perform more poorly in identifying Black faces than they do in identifying White faces is not, in and of itself, a meaningful test of the other-race effect. The particular Black stimulus faces used in the study might be less recognizable than the average Black person's face, or the particular White stimulus faces might be
more recognizable than the average White face.\(^2\) The beauty of the full design (i.e., designs in which participant–witnesses are of both races and the stimulus faces are of both races) is that main effect differences due to stimulus sampling can be ruled out by the presence of a statistical interaction. (For the nonpsychologist reader, a statistical interaction is present when the effect of one variable [in this case the race of the suspect] changes as a function of another variable [in this case the race of the eyewitness].)

A second concern that we have about a large portion of all other-race experiments is that they use one or another variant of the "old–new" or "yes–no" paradigm rather than the lineup paradigm. In the old–new paradigm, participant–witnesses are shown a large number of face photographs at encoding and then these faces (or a subset) are mixed with new faces and presented one at a time in a recognition test. The participant's task is to decide whether the face was presented in the original set (old) or not (new). The lineup paradigm, in contrast, presents test arrays (e.g., six photos) in which one of the original faces is either present (target present) or not (target absent). Many, if not most, researchers would argue that these two tests are fundamentally equivalent for purposes of testing the other-race effect. We accept that possibility but are not convinced that it is so. In particular, we think it is important to note that all identifications of new faces as old faces in the old–new paradigm are "false alarms" (i.e., mistaken identifications). In the lineup paradigm, on the other hand, there is only one suspect and \(N-1\) of the lineup members are "known innocent" fillers (where \(N\) is the number of lineup members). Some published studies of the other-race effect have used a lineup paradigm (Brigham, Maas, Snyder, & Spaulding, 1982; Doty, 1998; Fallshore & Schooler, 1995; Platz & Hosch, 1988). However, these studies have not included a target-present versus target-absent design in which filler identifications are not collapsed into the mistaken identification category. The identification of a filler is a "known error" in real cases and does not count as a mistaken identification of a suspect, because no charges would be filed against a filler. We believe that this difference is more significant than simply a difference in the "counting rule" that applies to the old–new paradigm versus the lineup paradigm. Consider, for instance, a provocative conundrum first noted by Lindsay and Wells (1983): If it is true that innocent suspects are better protected from mistaken identification when the fillers are more similar to them than when the fillers are not similar to them, then the innocent suspect ought to be better protected in other-race cases than in same-race cases because of the greater perceived similarity among lineup members in other-race cases. This conundrum does not apply to the old–new paradigm, but it appears to apply to the lineup paradigm. Sporer (2001b) discussed the problem with using the old–new para-

\(^2\)Although we are unaware of any systematic research, some portrait photographers tell us that, to properly photograph persons of color, it takes different lighting/procedures than those used to photograph White people. They claim that the facial characteristics of some persons of color are very poorly represented on black-and-white film and rather poorly represented on color film. Hence, main-effect differences in the recognizability of Black versus White people in experiments could be attributable to this artifact. This is another reason why the focus must be on the interaction between the race of the participant witness and the race of the stimulus person, which half-design other-race studies cannot assess.
digm as a proxy for lineups and agrees with our view that there might be important differences. We do not yet have a good answer to the Lindsay and Wells conundrum, but we challenge researchers to address this matter directly so as to either dismiss it as a concern or explore its implications.

Third, there is considerable ambiguity about race as an experimental variable. Anthropologists, for example, do not consider Hispanics or Mexican Americans to be a different race from White Americans, yet there is evidence that White Americans have more difficulty identifying Mexican Americans than they do identifying other White Americans (Platz & Hosch, 1988). Is race really the operative variable? Sporer (2001a) clarified what he means by race, namely “differences in physiognomy” (p. 36). Although we refer to the phenomenon as the other-race effect in the current article, our underlying view is similar to Sporer’s in that the operative factor is perceived facial physiognomic characteristics, regardless of racial classification per se. Our view is that race is one of many possible dimensions along which there are patterns of physiognomic variation between groups. Race is, of course, a socially significant variable, and the concept of race itself is more significant than mere physiognomic appearances. Hence, there is ambiguity about the extent to which the other-race effect is a perceptual–memorial phenomenon or a social phenomenon.

With these caveats in mind, we discuss some of the interpretations that have been given to the other-race effect. Establishing that an effect occurs is relatively easy in contrast to establishing that a particular psychological process is operating to produce that effect.

Interpretations of the Other-Race Effect

For the most part, interpretations of the other-race effect have tended to involve cognitive, unintentional, and automatic processes rather than social, motivational, and intentional processes. The cognitive rather than social emphasis in the existing research literature is probably due to the relatively early dismissal of racial attitudes as a moderator of the other-race effect (e.g., Brigham & Barkowitz, 1978; Lavrakas, Buri, & Mayzner, 1976). Meissner and Brigham’s (2001) meta-analysis appears to confirm the idea that racial attitudes are not related to the other-race effect. Nevertheless, we think it premature to abandon completely the idea that racial attitudes are involved in the other-race effect. One of the problems with testing the hypothesis about racial attitudes is that people are generally unwilling to admit to racial prejudice. Self-reported negative racial attitudes have largely gone underground, especially from the period of 1940 to 1980 in the United States (Schuman, Steeh, & Bobo, 1985), whereas there is plenty of evidence that racial prejudice continues to exert influence on behaviors (Dovidio & Gaertner, 2000). The idea that people increasingly have been hiding their racial prejudices over the last several decades is consistent with the observation that the relation between the other-race effect and self-reported racial attitudes has decreased over the last three decades (Meissner & Brigham, 2001). Implicit measures of racial prejudice or more subtle measures of racial attitudes (e.g., Pettigrew & Meertens, 1995) might show associations between racial attitudes and the other-race identification effect that are obscured by social desirability concerns.
Among social interpretations, Doyle (2001) raised one of the most disturbing. Doyle suggested that White American eyewitnesses might be more willing to guess (hold a more lax or liberal criterion) when attempting to identify Black people from lineups than when attempting to identify White people from lineups. He raised the possibility that White eyewitnesses might regard the cost of an error as lower when the suspect is Black than when the suspect is White. Meissner and Brigham (2001) reported results that could be considered consistent with this idea: that participant–witnesses appear to use a more liberal criterion with other-race faces than with own-race faces. There are two potential problems with Doyle’s interpretation. First, an apparently more liberal response criterion can occur for reasons other than a shift in response criterion per se. For instance, the perceiver might actually “see” greater resemblance between the previously viewed face and an innocent member of the lineup in other-race cases because of poor encoding of the original face. Having a superficial encoding of the original face, the witness might not be able to see the differences between the person being identified and the original face. Another problem is that there appears to be symmetry to the response-bias in the sense that Black participant–witnesses also appear to use a more liberal criterion when identifying White faces than when identifying Black faces. Are we to also assume that Black participant–witnesses believe the cost of error to be lower when the suspect is White than when the suspect is Black?

In spite of these concerns, it is important to not dismiss Doyle’s (2001) hypothesis too readily. Among the shortcomings of the other-race research literature is that the experimental, laboratory nature of the participant–witnesses’s experiences do not simulate the consequence-based nature of real eyewitness identifications. If Doyle is correct that White eyewitnesses are more likely to guess in other-race cases because they regard the cost of error to be lower in those cases, we would not necessarily expect this dynamic process to surface in laboratory experiments. To test Doyle’s hypothesis, the participant-witnesses would have to believe that negative consequences would be attached to the person identified. Only a few studies have used simulations in which the eyewitnesses believed that charges would be brought against the person identified (e.g., Luus & Wells, 1994; Malpass & Devine, 1980; Murray & Wells, 1982; Wells, Rydell, & Seealu, 1993), and these were not other-race identification studies. Hence, we believe that Doyle’s hypothesis has not actually been tested.

Also, Doyle’s (2001) analysis leads us to consider additional levels at which there can exist bias against a suspect based on race. Rather than focusing narrowly on the eyewitness him- or herself, researchers should pull back and examine the processes that are involved in criminal investigations. Any time an innocent person is placed in a lineup, there is some probabilistic chance that the person will be misidentified. Often, the only way that an innocent suspect can avoid being placed in a lineup is to provide a convincing alibi. Where is the research on alibi believability? (The authors of the current article are launching a research program on the psychology of alibis, which is still in its early stages. We have found no empirical literature on alibis.) Are alibis from Black suspects as believable as alibis from White suspects even when they are otherwise identical alibis? Also, what happens when the eyewitness says “I do not recognize any of these people as the culprit?” Our recent work (Wells & Olson, 2000) indicates that such statements by eyewitnesses should serve to reduce the observers’ (e.g., detec-
tives’) beliefs that the suspect is the culprit. Is there a race effect here too? Are nonidentifications by eyewitnesses as likely to remove the suspect from suspicion if the suspect is Black as they are if the suspect is White? Our point is that there are many levels at which racial bias might enter the eyewitness problem, and not all of these are a function of the race of the eyewitness per se. We call for future research to expand the agenda on the problem of race and eyewitness identification to examine additional factors (such as how people end up in lineups in the first place and whether nonidentifications are treated as equally exonerating for one race versus another race).

Most cognitive interpretations of the other-race identification effect hold certain assumptions in common. The primary assumption running through each of these interpretations is that there are inherent differences between faces of one race and faces of another race in terms of frequencies or patterns of variability in the features of the faces. Physiognomic variability is critical for successful face recognition memory because recognition memory tasks require the perceiver to distinguish between previously seen faces and new ones. For the most part, we can reject the hypothesis that faces of one race show more or less actual variability overall than those of another race. For instance, the greater difficulty that Whites have in identifying Black faces cannot be attributed to less actual physiognomic variability in Black faces, because Black perceivers show the reverse pattern (more difficulty with White faces than with Black faces). Furthermore, work by Goldstein and Chance (1979) suggests that there are no differences in actual physiognomic variation among Black, White, and Japanese faces.

Cognitive interpretations of the other-race effect tend to focus on the idea that the type of physiognomic variability in the faces of one race is not the same as the type of variability in another race (even though the total amount of variability is probably the same). In general, effective face recognition requires that a person notice and encode characteristics that will distinguish this person from other people who are only superficially similar. Attention to one set of features (e.g., skin tone, hair color) rather than another set of features (e.g., nose, chin) will be effective for some faces and ineffective for other faces in terms of capitalizing on characteristics that permit reliable recognition judgments (Levin, 1996). If White faces show more variability in hair color and Black faces show more variability in skin tone, for instance, then using skin tone for White faces or hair color for Black faces might be ineffective for purposes of later recognition. Because people generally have more experience with members of their own race than with members of another race, the habitual or automatic processing learned from a person’s own race will be more efficient for own-race than for other-race identification purposes (Ellis, Deregowski, & Shepherd, 1975). Although the cognitive models advocated by various researchers to account for the other-race effect have important differences, most share implicitly or explicitly the assumptions of (a) differences in the pattern or type of variability in faces of one race versus another and (b) habitual processing of faces based on the perceiver’s experience with own-race faces.

One variation on the general cognitive interpretation of the other-race effect is the configural-featured hypothesis (Rhodes, Brake, Taylor, & Tan, 1989). As people develop greater expertise with complex stimuli such as faces, they begin to process not only the features (e.g., nose, eyes) but also the relations (e.g.,
RACE AND EYEWITNESS IDENTIFICATION

...ordinal properties, distances) among the features (the *configuration*). Being less expert with faces of another race (presumably from less experience) than one’s own race, other-race faces tend to be processed featurally but not configurationally. A primary form of evidence for this interpretation is the work showing that inverting faces (to an upside-down position) disrupts own-race processing more than it disrupts other-race processing (Rhodes et al., 1989). An apparent assumption of the configural-featural hypothesis is that the other-race faces do not fit the configural processing strategy used for one’s own race, presumably because the physiognomic patterns are not the same across races. Another variation on this class of hypotheses is Valentine’s *multidimensional face space model* (Valentine, 1991). According to this model, faces are confused with each other to the extent that they tend to be close together in a multidimensional representation space in memory. According to this interpretation, own-race faces are spread rather widely in the mental representation, whereas other-race faces are more closely clustered. Chiroro and Valentine (1995) provided indirect evidence that other-race experience leads people to be more proficient in recognizing members of another race by encoding their distinctive dimensions and presumably spreading them out more in their multidimensional mental representation of the faces.

Sporer (2001a) offered yet another theoretical account of the other-race effect. His *in-group/out-group model* states that the first step in encountering a face involves social perception. When initially encountering an out-group (other race or other ethnic group) face, an out-group categorization cue is triggered, and the first step is that of categorizing the face (e.g., African American). This categorization in turn makes salient the dimensions along which in-group and out-group members differ. Attentional resources directed to the dimensions along which in-group and out-group members vary comes at the expense of attention to dimensions that would be efficient for distinguishing among individual members of the out-group. When encountering an in-group face, on the other hand, the group-categorization step is not involved at all. Instead, for in-group faces the perceiver automatically and immediately begins to process the face configurally in a manner that facilitates distinguishing among individual members of the in-group.

The work of MacLin and Malpass (2001) strikes us as perhaps providing support for Sporer’s in-group/out-group model (or models of that type). MacLin and Malpass created ambiguous-race faces that varied on only one feature (hair). Varying the single feature led participants to classify the face into either Hispanic or Black categories (Experiment 1). These faces were then used in a recognition paradigm, and the other-race effect was obtained (Experiment 2). It is somewhat surprising that changing only one feature had this effect unless we assume that the change creates some type of categorization effect that leads the perceiver to process the face differently when the face is thought to be a member of one’s own group versus some other group. The results of the MacLin and Malpass experiment are provocative. We wonder whether the effect might be found by merely changing the label of the ambiguous-race stimulus. Given an ambiguous-race face, for instance, would Hispanic and Black participants show different recognition performance if the person’s name were Juan versus Leroy?

It seems to us that most of these interpretations of the other-race effect are consistent with the findings of Smith, Lindsay, Pryke, and Dysart (2001), who showed
that the usual postdictors of eyewitness error do not seem to work with other-race identifications. In general, any factor that disrupts the normal way in which faces are recognized is likely to change the correlates that predict recognition memory. If certainty, decision time, and relative-versus-absolute judgments are not correlated with other-race identification errors, then what variables are correlated with other-race identification errors? The answer to that question might be very informative about the processes underlying the other-race effect.

We are impressed with the quality and sophistication of the theoretical treatments of the other-race effect that have emerged in recent years. The fact that no single explanation of the effect is agreed on by all researchers is hardly surprising, for two primary reasons. First, there is not yet a full understanding of how memory for faces works in general, regardless of the race of the face. Second, the other-race effect is probably the result of several causes rather than one alone. Race is not simply a matter of physiognomic differences among stimuli or familiarity with physiognomic patterns; race is also a social variable. As a social variable, race can trigger affective responses that might interfere with cognitive processes. In real cases involving crime and victim–witnesses, attitudes about race might have more of an effect than is found in controlled experiments in which the perceiver has no particular investment in the faces under consideration and there are no consequences to the persons identified. In the end, if there is such a thing, we believe that many factors will be needed to explain the other-race effect in actual cases of witnesses to crime, including both emotional and cognitive factors.

System Recommendations Regarding the Other-Race Effect

Our attempt to understand the causes of the other-race effect does not show much promise regarding how we might undo those causes. Suppose, for example, that the other-race effect is the result of having learned to use sets of facial features from one’s own race (because that is where we have the most experience) and applying them inappropriately to encoding faces of another race. Although we can imagine some type of massive other-race-recognition training program in schools or a more thorough societal integration of the races, neither of these ideas is a realistic response to the other-race identification effect. Furthermore, the data suggest that experience is playing a small role in the other-race effect. We believe that this is a situation in which realistic solutions cannot be derived directly from our understanding of the causes of the problem. Instead, we propose what might be called “effect remedies,” which attempt to minimize the effect of the phenomenon rather than correct the causes of the problem.

Before proposing our recommendations regarding the other-race effect, we think it is important to draw two distinctions that are useful for understanding our proposed recommendations. The first distinction is between system and estimator variables. Because this distinction is well known, we mention it only briefly. The second distinction is between general performance impairment variables and specific-suspect bias variables. This distinction is new to the literature, so we discuss its implications in greater detail.
RACE AND EYEWITNESS IDENTIFICATION

System Versus Estimator Variables

The question of how the justice system might deal with the other-race identification effect is a difficult one. The other-race effect is fundamentally different from certain other variables in eyewitness identification, such as structural lineup bias, because race is an estimator variable, not a system variable (Wells, 1978). The justice system can control how it selects fillers for a lineup, but it cannot control whether the witness and perpetrator are of the same race. The fact that the other-race variable is not a system variable is the reason why the national guidelines on eyewitness evidence developed by the U.S. Department of Justice (Technical Working Group for Eyewitness Evidence, 1999) did not deal with the other-race issue at all. These national guidelines were concerned with procedures for the collection of eyewitness evidence rather than the evaluation of eyewitness evidence.

Having classified the other-race effect as an estimator variable, however, does not automatically mean that the justice system can do nothing about the other-race identification problem. Theoretically, there might be moderator variables that the justice system controls (e.g., some type of pre-lineup instruction variable) that would reduce or eliminate the other-race effect. Frankly, we know of no such variables at this point, but we hold out the possibility that such variables might exist, and we encourage researchers to search for them. Meanwhile, as we state later, in the section Possible System Variable Responses to the Other-Race Effect, we believe that there are system variables that can reduce risk in other-race cases even though they do not moderate the other-race effect itself.

General Performance Impairment Versus Specific-Suspect Bias Variables

There is yet another qualitative difference between the other-race effect and certain other variables in eyewitness identification that, to our knowledge, has never been discussed in the literature. Specifically, we can make a distinction between eyewitness identification variables that impair performance generally versus those that produce a specific bias against the suspect. Consider, for instance, lighting conditions at the crime scene. Poor illumination should provide less opportunity for the eyewitness to acquire useful information in memory regarding the perpetrator such that, on viewing a lineup at a later date, the identification decision of the eyewitness should be less reliable than it would have been had the lighting conditions been good. However, the lighting conditions themselves do not create a bias against any particular person in the lineup. Poor lighting per se will fail to explain why a witness selected any particular lineup member. In this sense, poor lighting is a general impairment variable rather than a specific-suspect bias variable. The other-race effect is also a general impairment variable. Assuming a fair lineup procedure, the fact that a witness is White and the suspect is Black will fail to explain why a witness has identified the Black suspect rather than one of the Black fillers in the lineup.

Consider, in contrast, structural lineup bias. Structural lineup bias exists when
the suspect fits the description of the culprit that was provided by the eyewitness whereas the fillers do not fit that description (Malpass & Lindsay, 1999). Unlike general performance impairment variables (such as lighting conditions, race of witness and perpetrator, the duration between the witnessed event and the lineup, the amount of attention paid to the perpetrator), structural lineup bias is a specific-suspect bias variable. The presence of structural lineup bias leads the eyewitness toward a specific member of the lineup rather than distributing errors across lineup members. Unlike general performance impairment variables, specific-suspect bias variables explain why the witness selected a particular member of the lineup.

The distinction of general performance impairment versus specific-suspect bias is not redundant with the system variable/estimator variable distinction. A system variable can be either a general performance impairment variable (e.g., instructions prior to viewing) or a specific-suspect bias variable (e.g., structural-lineup bias). Similarly, an estimator variable can be either a general performance impairment variable (e.g., the other-race variable) or a specific-suspect bias variable (e.g., unconscious transference).

We think that the status of the other-race effect as a general performance impairment variable (rather than a specific-suspect bias variable) might be significant to our understanding of how jurors reason about the other-race variable. We suspect that there is a level at which jurors implicitly, if not explicitly, dismiss the other-race variable as something that is largely irrelevant to a specific case because the race variable begs the question as to why the witness selected that particular suspect. We imagine jurors deliberating and saying something like “If they ‘all looked alike’ to this eyewitness, then why did she identify the defendant rather than one of the other Black men in the lineup?” If we are correct in our speculation about juror reasoning, then jurors might not be persuaded by the other-race variable, because it fails to address the principal question of identity in the minds of the jurors.

**Expert Testimony on the Other-Race Effect**

Given the status of the other-race effect as an estimator variable rather than a system variable, one of the more obvious ways to address the other-race problem is through expert testimony. Although we do not oppose expert testimony on the other-race effect, we believe that this is a weak solution that attempts to correct a problem at the wrong point in the process. Much has been written about the issue of expert testimony on eyewitness identification issues, and it is not our purpose to belabor the arguments for and against expert testimony here (see Buckhout, 1986; Cutler, Penrod, & Dexter, 1989; Egeth & McCloskey, 1984; Leippe, 1995; Lempert, 1986; Loftus, 1980; Pachella, 1986; Wells, 1984b). However, there remain many unresolved questions, such as the following: (a) To what extent do jurors already understand that other-race identifications are more difficult than within-race identifications? (b) Does expert testimony actually help jurors distinguish other-race identifications that are accurate from those that are inaccurate? and (c) If jurors need to be told by an expert to be skeptical in other-race cases, should they be told by an expert to be less skeptical when it is a within-race case?

Our primary concern about expert testimony on the other-race effect is that we
fear that it will be both costly and ineffective. The cost involves both court time (it lengthens trials) and money (e.g., fees, travel expenses). The ineffectiveness issue stems from the concern that jurors will dismiss the general argument as applicable to the case because it fails to address why the witness selected that particular person from the lineup. More important, expert testimony does nothing to prevent the mistaken identification in the first place, which ought to be a primary goal of our research programs. Furthermore, expert testimony on the other-race effect itself sends no prevention message to the justice system regarding how the justice system can improve the accuracy of eyewitness identification evidence.

Bartolomey (2001), a prosecutor, raised a number of concerns about the state of the literature on the other-race effect and is clearly opposed to both expert testimony on the other-race effect and any special instructions to the jury in other-race cases. She argued that the system already includes a large number of safeguards, such as pretrial suppression motions, cross-examination, closing arguments by defense stressing the race-factor argument, and instructions to the jury regarding the State's burden to prove guilt and identity beyond a reasonable doubt. In response to these points, we note work on the inadequacy of the traditional safeguards for all eyewitness identification evidence, regardless of the race factor (Cutler & Penrod, 1995; Stinson, Devenport, Cutler, & Kravitz, 1996; Stinson, Devenport, Cutler, & Kravitz, 1997). Indeed, all of these safeguards were in place when McKinley Cromedy, a Black male who was identified by a White eyewitness, was convicted in New Jersey (State v. Cromedy, 1999). After the New Jersey Supreme Court ordered a new trial for Cromedy (because the trial judge had refused to give cautionary instructions on other-race identification to the jury), forensic DNA tests proved Cromedy to be innocent. Whether cautionary instructions or expert testimony on the other-race effect would have prevented Cromedy's conviction in the original trial is something we can never know. However, the idea that existing safeguards are sufficient protection against the problem of mistaken identification is, in our opinion, not persuasive. Although we have concerns about the cost and effectiveness of expert testimony, preferring instead the types of system-variable improvements that we offer in the following section, we are not per se opposed to expert testimony on the other-race effect.

Possible System Variable Responses to the Other-Race Effect

Increasing the number of fillers. Suppose that the justice system were to consider the totality of the empirical data and conclude that other-race identifications run a higher risk of mistaken identification of innocent suspects than do own-race identifications. Are there methods to lower this risk in other-race cases? Clearly there are. Perhaps the most direct way to do this is to require that lineups include more fillers in other-race cases. Fillers are known-innocent members of lineups who serve to protect innocent suspects from mistaken identification by luring eyewitness choices away from the innocent suspect and toward the fillers. Although the identification of a filler is an error, it is a relatively harmless error in the sense that charges will not be brought against an innocent filler who is selected by an eyewitness.

Let's assume, for example, that 30% of the eyewitnesses in other-race cases
identify someone from a lineup in which the suspect is not the actual perpetrator and only 15% do so in an own-race case. Assume further that there are six members in the lineup, one of which is the suspect and the other five are fillers. Assuming that the lineup is fair, the innocent suspect has a 5.0% chance of being misidentified (1/6 of the 30%) in the other-race case and a 2.5% chance in the own-race case. If lineup size were doubled in the other-race case, the chances of misidentification in the other-race and own-race cases would be equal (each at 2.5%). We do not know the exact number of additional fillers required in other-race cases to erase the difference in risk to innocent suspects in actual cases. Nevertheless, the general principle of adding to the number of fillers in other-race cases strikes us as both simple and sound.

The selection of fillers. Our analysis assumes, of course, that the additional fillers possess the qualities for fulfilling their proper role as lures. The literature is fairly clear on what constitutes a proper filler, and the criteria are not particularly complex or difficult to achieve (see Wells, Seelau, Rydell, & Luus, 1994). However, one complication that can arise in other-race cases regarding the selection of fillers is evident in the work of Brigham and Ready (1985). They showed that lineup constructors of a different race than the suspect tended to select fillers for a lineup who did not resemble the suspect as closely as did lineup constructors of the same race as the suspect. This leads us to a second recommendation: Those who select fillers for use in a lineup should be of the same race as the suspect in the case, because it appears that other-race observers cannot readily detect potential biases that make the suspect stand out as distinctive. We might even go further to suggest that motions to suppress eyewitness identification evidence based on a biased lineup should be evaluated by a trial judge who is the same race as the defendant. The idea that a White trial judge can or should rule on the adequacy of a lineup of Black persons seems at least questionable given the findings of Brigham and Ready.

The blank lineup. Using additional fillers in lineups is one of the ways to use system variables to address the other-race identification problem. It is not, however, the only way to address the problem. The fact that eyewitnesses appear to use a more lenient criterion for making other-race than for making own-race identifications leads us to suggest that the blank lineup control procedure might be useful in other-race cases. The blank lineup control procedure involves the construction of two lineups: one containing the suspect (actual lineup) and the other containing only fillers (blank lineup) (Wells, 1984a). In the blank lineup procedure, the eyewitness is given the usual lineup instructions (“The perpetrator might or might not be in the lineup . . .”) but is shown a blank (fillers only) lineup. The eyewitness does not know that the lineup contains only fillers, of course, and is unaware that there is a second lineup. If the eyewitnesses are simply prone to pick someone, then they will likely pick a filler in the first lineup. It should be emphasized that there is no risk in the first lineup that they will pick an innocent suspect because there are no suspects in the first lineup, only known-innocent fillers. Any witness who does not make an identification from the first lineup, on the other hand, can then be shown the second (actual) lineup. If the eyewitness picks the suspect from the second lineup, then we can assume that the identification is relatively reliable. The blank lineup control procedure appears to be effective in weeding out eyewitnesses who are overly eager to make a selec-
tion, and the rate of misidentifications for those who survive the procedure is very low in comparison with those who were not subjected to the blank lineup prior to being shown the actual lineup (Wells, 1984a).

We realize that there is something a bit peculiar about our recommendations to use more fillers in other-race cases than in own-race cases, or to use the blank lineup control procedure in other-race cases but not in own-race cases. Clearly, both of these recommendations could be applied to own-race cases to further reduce jeopardy to innocent suspects, regardless of the race of the witness and suspect. So, what sense does it make to recommend these system improvements only in other-race cases? We do not have a perfect answer to that question. In general, we would rather see recommendations of this type applied to all lineups, not only in other-race cases. However, there are practical considerations associated with the addition of fillers or the running of two lineups rather than one, such as expense and time. Practical considerations of cost and time played a large role in the deliberations among police, prosecutors, and researchers in the development of the Department of Justice's national guidelines (Wells, Malpass et al., 2000). If these extra costs are borne in only the other-race cases, then it might be a more practical and acceptable recommendation to the justice system than that of bearing these extra costs in all eyewitness cases. Bearing the cost of extra fillers and or dual lineups in other-race cases might be especially attractive to legal policy makers if the alternative is to routinely permit expert testimony in other-race cases. The use of extra fillers or dual lineups would probably be less costly than expert testimony in other-race cases, and it has the additional advantage of helping to prevent mistaken other-race identifications from occurring in the first place.

Conclusions

Research supports the conclusion that people have more difficulty recognizing previously viewed faces of persons of another race than they do of their own race. A variety of procedures, materials, and experimental designs have been used in demonstrating the other-race effect. We have some concerns about the representativeness of the materials in some of these studies and the completeness of their designs. Nevertheless, the consistency of the findings suggests to us that the phenomenon is both real and relevant to the criminal justice system. The question of what should be done to remedy the problem is a difficult one. We are not opposed to expert testimony at trial. However, we prefer a system variable approach in which other-race identification cases receive special preventive treat-

3Regardless of the number of lineup fillers or whether a blank lineup procedure was used, a double-blind method should be used for administering the lineup (Wells et al., 1994). Because of the case with which lineup administrators can influence eyewitnesses, the lineup administrator should not be the detective in the case (or anyone else who knows which lineup member is the suspect). If Doyle (2001) is correct in his assertion that race permeates the behaviors of actors in the legal system, then we might expect the lineup administrator to be more "helpful" to some witnesses in picking out the culprit, depending on racial issues. By introducing double-blind testing in all cases (regardless of the race of the participants), any contribution of race to suggestive practices during the lineup procedure could be minimized, because in that case not even the lineup administrator him- or herself would know which lineup member is the suspect.
ment at the time of the identification test (e.g., use of more lineup fillers, use of the blank lineup control procedure) rather than an approach that attempts to undo an error at trial (e.g., introducing expert testimony or special jury instructions).

References


Smith, S. M., Lindsay, R. C. L., Pryke, S., & Dysart, J. E. (2001). Postdictors of


Received September 8, 2000
Revision received October 25, 2000
Accepted October 26, 2000