

Defendant–Juror Similarity and Mock Juror Judgments*

Norbert L. Kerr,† Robert W. Hymes,‡
Alonzo B. Anderson,§ and James E. Weathers

It was hypothesized that juror–defendant similarity would lead to greater leniency toward a criminal defendant when the evidence against that defendant was weak or inconclusive; but when evidence was strong, it was expected that this relationship would be reversed. In Study 1, religious similarity was found to be simply and positively related to evaluation of the defendant and leniency, a relationship unaffected by the strength of evidence. This pattern of results was attributed to (a) insufficiently strong evidence against the defendant and (b) the lack of anticipated jury deliberation, problems addressed in Study 2. In that study, when evidence was strong against the defendant, juror–defendant racial similarity did increase the likelihood of conviction, but only when jurors anticipated being in the racial minority in their jury. Implications of the findings for psychological theory and for voir dire were discussed.

By exercising peremptory challenges, attorneys in civil and criminal jury trials may influence the composition of juries. They naturally have been interested in identifying any juror characteristics that may be reliably associated with a favorable or unfavorable verdict. Extant research suggests very few reliable and general relationships between juror demographic characteristics (e.g., Golash, 1992) or personalities (Bray & Noble, 1978; Fitzgerald & Ellsworth, 1984; Fulero & Penrod, 1990) and juror verdicts. However, many practitioners (Gobert & Jordan,

* The authors wish to thank Howard Klein and Jane Stanfel for their assistance in data collection and analysis in Experiment 2, and the Editor and two anonymous reviewers for their suggestions on an earlier draft. Correspondence should be sent to the first author at the Department of Psychology, Psychology Research Building, Michigan State University, East Lansing, MI 48824 (E-mail: Kerr@pilot.msu.edu).

† Michigan State University.

‡ University of Michigan—Dearborn.

§ University of Southern California.

1990; Jordan, 1980; Wishman, 1986) and scholars (Dane & Wrightsman, 1982; Hans, 1992; Hans & Vidmar, 1986; Monahan & Loftus, 1982; Wrightsman, 1987) have suggested that one cannot predict the effect of juror characteristics without also taking into account the characteristics of other actors in the courtroom drama.

In their classic review of early research on juror behavior, Davis, Bray, and Holt (1977) noted that the evidence from several studies was consistent with a *liking–leniency hypothesis*. According to Davis et al. (1977), “a well-liked defendant will receive more lenient treatment from jurors than one who is disliked” (p. 329). By coupling this liking–leniency hypothesis with the well-established direct relationship between similarity and attraction (e.g., Byrne, 1971; Byrne, Clore, & Smeaton, 1986; Smeaton, Byrne, & Murnen, 1989), one is led to the hypothesis that the more similar a juror and the defendant are, the more lenient the juror is likely to be; we term this the *similarity–leniency hypothesis*.

Judging from texts on trial tactics in voir dire (e.g., Bailey & Rothblatt, 1985; Keeton, 1973), the similarity–leniency hypothesis seems to be widely accepted by attorneys as a valid rule of thumb. For example, in their text on trial tactics and jury selection, Gobert and Jordan (1985) assert that attorneys should select jurors similar to the defendant because the jurors “are much more apt to have had life experiences similar to those of the accused. They are more likely to be understanding of the defendant’s actions, and perhaps more willing to exercise their powers of nullification [of the verdict]” (p. 47). Wishman (1986) suggests that “Black jurors will be less likely to return a conviction against a Black defendant because of a feeling of brotherhood” (p. 116). And Jordan (1980) suggests that “if the witnesses that counsel is going to attack and try to discredit are of a particular ethnic origin, counsel must make certain that persons with the same background do not dominate the jury” (p. 256).

Reviews of research on the similarity–leniency hypothesis seem to draw corroborative conclusions (Dane & Wrightsman, 1982; Wrightsman, 1987). Whether the dimension of juror–defendant similarity concerns language (Stephan & Stephan, 1986), general social attitudes (Gerbası & Zuckerman, 1975), political ideology (Amato, 1979; Griffitt & Jackson, 1973), or lifestyle (Griffitt & Jackson, 1973), jurors are often less punitive when defendants are similar. Likewise, studies examining similarity on such demographic factors as gender (Stephan, 1974) and race (Bernard, 1979; Feild, 1979; Klein & Creech, 1982; Ugwuegbu, 1979) find comparable leniency effects. Such evidence has led some legal scholars (e.g., Johnson, 1985; Sweeney & Haney, 1992) to conclude that the latter racial bias is pervasive and recommend that jury composition take into account the defendant’s race.

However, there are also reasons to suspect that there are important limits to the similarity–leniency relationship. For example, the strength of the similarity–leniency relationship may well depend upon juror personality factors (Kauffman & Ryckman, 1979; Mitchell & Byrne, 1973). Also, defense attorneys have occasionally cautioned their colleagues to avoid selecting jurors belonging to the same group as a defendant. For example, Belli (1963) has suggested that Jewish jurors are harsher on fellow Jews than on non-Jews accused of a crime. Owen (1973) and

Bailey and Rothblatt (1985) have suggested that women jurors are particularly harsh with women defendants. Wishman (1986) cites an example of a crime in which the defendant was Black and suggests that "middle-class Blacks might very well consider [the defendant] a 'disgrace to their race' and convict him more readily than would Whites" (p. 100). Indeed, Golash (1992) notes that, "the closer the jury is to the defendant's own social context, the more likely it is that the jurors will have a full understanding of the defendant's conduct, and the better they will be able to evaluate its excusability. Such full understanding will not always favor the minority defendant: The jury may correctly perceive the defendant's conduct as inexcusable, where a more demographically distant jury would have found it excusable" (p. 172).

More importantly, there is also a fairly sizable body of psychological theory and research that suggests that, at times, another's similarity can, indeed, actually repel and prompt harsh treatment. At least six separate lines of work come to such a conclusion:

1. The first (and most directly relevant to the current research) is work on social categorization. Consistent with the similarity-lenience hypothesis, this literature demonstrates a robust tendency for members of mutually exclusive groups to discriminate in favor of their ingroup and against the outgroup (e.g., Brewer, 1979; Gerard & Hoyt, 1974; Tajfel, Flament, Billig, & Bundy, 1971). According to social identity theory (Tajfel & Turner, 1986), such ingroup favorability results from a general need to maintain a positive self-image. By viewing the ingroup as more positive than the outgroup, individuals can maintain or even raise their self-esteem. However, because members are motivated to perceive the ingroup positively, they will tend to derogate any agent or event that threatens the ingroup image. An unlikable or deviant ingroup member represents a threat to this positive ingroup image, and should therefore be derogated more than an equally unlikable or deviant person who poses no such threat (e.g., an outgroup member). Marques and Yzerbyt (Marques, Yzerbyt, & Leyens, 1988; Marques & Yzerbyt, 1988; Marques, 1990) have termed this negative extremity toward ingroup members the *black sheep effect*. According to Marques (1990), group members are more polarized in their evaluation of ingroup than outgroup members. Thus, positively valued ingroup members will be perceived more favorably than positive outgroup members, but negatively valued ingroup members will receive greater disapproval than will negatively valued outgroup members. In support of this hypothesis, ingroup members with notably weak skills or undesirable traits are more likely to be rejected by fellow ingroupers than by outgroupers (Jackson & Hymes, 1985; Marques et al., 1988, Studies 1 and 3; Marques & Yzerbyt, 1988). And extremely negative evaluations are likely to be found for ingroup members who do not conform to relevant (but not irrelevant) ingroup norms (Marques et al., 1988, Study 2; Marques, 1990, Study 5).

2. A second basis for such black-sheep effects is research showing that negative behavior by an ingroup member can be especially salient to other group members. Studies of task performance in group settings (Jones & deCharms, 1957; Lauderdale, Smith-Cunniën, Parker, & Inverarity, 1984) have shown that where a single group member's output (or lack thereof) could potentially cause

harm to the entire group, the offending member is very likely to be rejected. In the same vein, deviates (e.g., criminals) may be perceived as dishonoring the entire group (Cooper & Jones, 1969; Jones & deCharms, 1957; Lauderdale et al., 1984; Marques, 1990). Those who share group membership with a defendant who looks very guilty may be concerned about the possibility of guilt-by-association. Such an injury could be reciprocated by harshness toward the defendant.

3. Third, early studies on the similarity–attraction hypothesis (Byrne & Wong, 1962; Smith, Williams, & Willis, 1967) indicated that disagreements with same-race others were associated with more negative evaluations than were disagreements with others of a different race. Later research demonstrated that information implying similarity to a negatively valued other led to various signs of repulsion, including an avoidance of the other (Novak & Lerner, 1968), a reduction in reported attraction (Senn, 1971), and a tendency to change responses on which the judged similarity was based (Cooper & Jones, 1969).

4. Fourth, studies by Mettee and his colleagues (Mettee & Riskin, 1974; Mettee & Wilkins, 1972; Taylor & Mettee, 1971) have led to the conclusion that similar others are the most relevant sources of social comparison information. They suggest that when individuals are faced with perceived similarity to an other that carries negative implications about the self, they will tend to dissociate themselves from that other. This goal could be achieved by perceiving the similar other as not really sharing traits with the self or by acting so as to eliminate that perception of similarity by others. But individuals can also achieve a dissociation from undesirable similar others by derogating and giving them extreme negative evaluations.

5. Fifth, black-sheep effects may also be tied to a violation of expectations. According to Jussim, Coleman, and Lerch (1987), “when an individual’s characteristics violate stereotype-based expectations, evaluations should become more extreme in the direction of the expectancy violation” (p. 537). Thus, more negative evaluations are likely when a person acts contrary to societally prescribed roles. In extending this analysis to social category membership, it is clear that ingroup members are expected to be better task performers, possess more skills and positive personality traits, and engage in more favorable behaviors than their outgroup counterparts (Brewer, 1979; Duncan, 1976; Howard & Rothbart, 1980; Pettigrew, 1979; Taylor & Jaggi, 1974). When the behaviors of ingroup members violate these expectations negatively, it follows that they will be perceived more negatively than if the same behavior had been performed by an outgroup member (cf., Lauderdale et al., 1984; Marques et al., 1988; Schachter, 1951).

6. Sixth, black sheep effects are also suggested by work on defensive attribution. First noted by Walster (1966), and later refined by numerous researchers (Chaikin & Darley, 1973; Harvey, Harris, & Barnes, 1975; Lowe & Medway, 1976; Shaver, 1970), defensive attribution research indicates that as the consequences of an action become more dire, outside observers are more likely to hold the actor responsible or blameworthy. More to the current point, these defensive attributions tend to become even more pronounced when the actor’s behavior is of personal relevance to the observer (Lowe & Medway, 1976; Medway & Lowe,

1975; Shaw & Skolnick, 1971), as one might expect if the actor belonged to the same social group as the observer.¹

In summary, it is clear from several lines of basic, social psychological research that similarity can lead to harsher rather than more lenient treatment. It is important to note that this process is not simply a matter of perceiving a similar other as dissimilar. Rather, similarity appears to be associated with a polarization or extremization of feelings. Thus, when similar others are perceived as positive, they are seen as better than dissimilar others of equal positivity. On the other hand, similar negative others appear to be perceived more negatively than their negative-dissimilar counterparts.

With this social psychological research and with legal speculation about the possibilities of black sheep effects in mind, we hypothesized that if juror-defendant similarity were defined with respect to significant group memberships, criminal defendants who probably were guilty (i.e., those faced with a very strong prosecution case) would be treated more harshly by similar than dissimilar jurors—that is, an apparent “black sheep” would be judged more harshly by members of his own flock than by members of another flock. If, however, there was considerable uncertainty about the defendant’s guilt (e.g., the prosecution case was weak), we hypothesized that the similarity–leniency hypothesis would hold. Again, paralleling the defensive attribution research (Lowe & Medway, 1976; Walster, 1966), if there were not clear indications that the defendant was indeed a “black sheep,” we expected to see the usual favoritism displayed toward an ingroup member.

As noted earlier, there is considerable lab (Davis et al., 1977) and field (e.g., Bowers & Pierce, 1980) research documenting the latter, similarity–leniency effect. However, practically no prior research has simultaneously manipulated (or measured) both juror–defendant similarity and the strength of evidence to enable a clear test of our hypotheses. Two exceptions are the simulation studies by Feild (1979) and Gordon (1993). The designs of these studies are interesting. In both, White mock jurors considered cases in which race of defendant and strength of evidence were manipulated. And, for the nonstereotypic crime (i.e., one not seen as relatively more likely to be committed by members of one or the other racial group) in the Gordon study, all the dependent variables exhibited a black-sheep effect for the strong-evidence condition and most (guilt ratings being a notable exception) exhibited the predicted attenuation/reversal as evidence became weaker; however, reanalyses of Gordon’s data indicated that none of these interaction effects was statistically significant. It is important to note, however, that these studies were not designed to (nor did they) provide direct and unequivocal tests of our hypotheses. For example, in neither study is there clearly a contrast

¹ However, note that Shaver (1970) suggests that the opposite may happen when the similarity between observer and actor is so great that the observer cannot dissociate from the actor. Under these conditions, Shaver argues, an observer will hold the actor as less rather than more responsible, with increasingly severe consequences. A number of investigations support this line of reasoning (Chaikin & Darley, 1973; Shaver, 1970; McKillip & Posavac, 1975).

of cases with very strong vs. very weak evidence against the defendant (e.g., mean guilt ratings for Gordon's "weak evidence" condition was 5.72 on an 11-point scale, where 6 signifies maximal uncertainty). Our studies were designed to avoid such ambiguities.

STUDY 1

Religious affiliation was chosen as a significant group membership in this study. All participants were carefully selected to identify strongly with their religious group. Further, the nature of the trial evidence ensured that the defendant's membership in the ingroup could not be easily denied. As an ancillary question, we hoped to test speculation (e.g., Belli, 1963; Busch, 1959; Bailey & Rothblatt, 1985) that Jewish jurors are relatively good jurors for the defense.

Method

Participants

The participants were 66 male college students who were either devoted Christians or Jews. They were identified either as active members of campus religious groups (e.g., Campus Crusade for Christ, Jewish Student Association) by officers of those organizations, or recommended to us as strongly religious persons by other such members after the latter's participation. Potential participants were telephoned and invited to participate in a study of juror decision making at any convenient time. During subject recruitment and throughout the study itself, participants were never told that they had been recruited because of their strong religious convictions. Thirty-five Christian and 31 Jewish students participated.

Design and Stimulus Trials

The experimental design was a 2 (Subject's religion: Christian or Jewish) \times 2 (Defendant's religion: Christian or Jewish) \times 2 (Strength of evidence against the defendant: strong or weak) factorial. The first factor was a between-subjects factor; the latter two factors were manipulated within subjects. Thus, every subject considered four case summaries. Two were strong prosecution cases and two were weak cases; two of the cases involved Christian defendants and two involved Jewish defendants. Each case summary described a purportedly real child molestation case. This type of crime was chosen because it was strongly counternormative for both religious groups. To enhance the perceived similarity of juror and defendant, each defendant was a male student of college age. Pretesting had established that two of these cases (termed the "Harris" and "Burger" cases, after the defendants' last names) had strong evidence against the defendants (approximately 65% convictions among pilot participants) while the other two (the "Williams" and "Small" cases) had weak evidence (approximately 30% convictions). The defendants in each case were portrayed through testimony as active

members of a religious group; the group for any case could be either Christian or Jewish by changing a word or phrase (e.g., active in a church youth group vs. active in a synagogue youth group). Each case summary was 3–4 double-spaced pages long. Each began with a summary of the testimony of the victim followed by other prosecution witnesses (e.g., eyewitnesses, investigating/arresting officer). In each case, the defendant testified in his own behalf, along with other defense witnesses (e.g., eyewitnesses, character witnesses). Defense and prosecution summary statements completed each case.²

Participants considered weak and strong cases in an alternating order. The specifications of defendant's religion and the ordinal position for each case were also counterbalanced.

Procedure

Participants were scheduled and run individually. After completing a consent form that revealed the nature of the cases they would be considering (viz., cases of alleged child molestation), participants read a brief summary of a judge's instructions to jurors for such a case. This judge's charge offered a statutory definition of child molestation and reminded jurors of the prosecution's burden of proof and the requirement that guilt be established beyond a reasonable doubt. After reading each case summary, participants provided a verdict as well as some ancillary judgments. (The latter included certainty in verdict, probability that the defendant committed the crime, penalty recommendations, and ratings of the defendant on seven-point semantic-differential type scales [viz., likable/not likable, moral/immoral, similar/dissimilar, and good/bad].) After completing the last case and providing some demographic information (including religious affiliation), the subject was completely debriefed. Names of other male students with strong religious convictions were solicited and a promise of confidentiality was obtained.

Results

Checks on Experimental Manipulations

All participants recruited based on a particular religious affiliation reported that affiliation in the postexperimental questionnaire. Mock jurors' awareness of the defendants' religious affiliations was confirmed through analysis of participants' ratings of similarity to the defendant. Analysis of variance of these ratings produced the expected subject religion \times defendant religion interaction effect, $F(1,249) = 6.15, p < .02$; participants reported feeling more similar to coreligionist defendants. Finally, the strength of evidence manipulation was confirmed by a main effect of this factor within a log-linear analysis of verdict data, $G^2(1 \text{ df}) = 29.25, p < .001^3$; the conviction rate for the relatively strong cases was significantly higher than for the weaker cases (60% vs. 28% guilty, overall, respectively).

² Complete copies of these cases may be obtained from the authors on request.

³ The G^2 statistic is distributed approximately as the chi-square statistic on the indicated degrees of freedom.

Likewise, on the judgment of the likelihood that the defendant did indeed commit the crime, there was a comparable main effect for the strength-of-evidence factor, $F(1,255) = 4.50, p < .04$ (means, .71 vs. .57, respectively).

Verdicts and Evaluation of Defendants

As noted above, the participants' verdicts were analyzed within a 2 (Subject's religion) \times 2 (Defendant's religion) \times 2 (Strength-of-case) contingency table utilizing Feinberg's (1970) log-linear test procedures and Overall and Spiegel's (1969) Method 2 for ordering tests. Besides the strength-of-evidence main effect, two other effects emerged. Jewish participants were significantly less likely to convict than Christian participants (Jewish conviction rate = 37%, Christian conviction rate = 50%), $G^2(1 df) = 5.50, p < .025$. The other effect (plotted in Figure 1) was a marginally significant similarity–leniency relationship, $G^2(1 df) = 3.60, p < .06$. The figure reveals a pattern fully consistent with the similarity–leniency hypothesis; when the mock jurors were Jewish, conviction was less likely for a Jewish defendant than for a Christian defendant, whereas the opposite was true when the mock jurors were Christian. The predicted three-way interaction of strength-of-evidence \times subject's religion \times defendant's religion was not significant, $G^2(1 df) = 1.69$.

The remainder of the dependent variables (viz., penalty recommendations, ratings of defendant goodness, morality, and likability) were analyzed in 2 (Subject's religion) \times 2 (Defendant's religion) \times 2 (Strength-of-case) least-squares

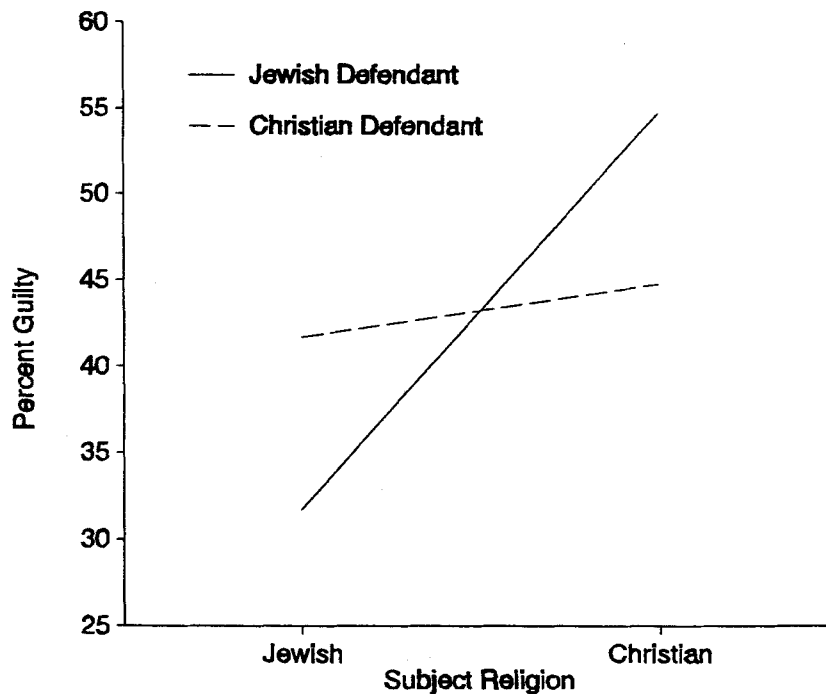


Fig. 1. Juror religion \times Defendant religion interaction effect on verdicts: Study 1.

repeated-measures analyses of variance. The only significant effect found was a subject-religion \times defendant-religion interaction on the good/bad rating of the defendant, $F(1,249) = 4.21, p < .04$; paralleling the verdict data, participants of both religions evaluated the religiously similar defendant more positively than the religiously dissimilar defendant. Similar but weaker and nonsignificant trends emerged for ratings of the defendant's morality ($p < .11$) and likability ($p < .14$).

Discussion

Consistent with much speculation to the same effect, Jewish participants were more lenient than the Christian participants. Of course, for this effect (as well as the similarity–leniency effect) we must remember that our participants were particularly religious and thus probably not highly representative of all those who profess a Christian or Jewish religious affiliation. Still, because juror religious belief is an item of information typically available or obtainable by attorneys in *voir dire* (Bailey & Rothblatt, 1985), this result deserves attempts at replication with different subject populations (particularly jury roll members), more representative samples, and different types of cases.

The data also tended to favor the direct similarity–leniency hypothesis. This effect was not moderated by the strength of evidence against the defendant, as was predicted by the black sheep hypothesis. Several reasons for the failure of the prediction (besides its invalidity) may be advanced. First, the evidence against the defendant in the strong case condition seems to have produced more of a moderate than strong certainty of the defendant's guilt (60% convictions). Of course, in developing our trial materials, we were caught between our desire for cases whose evidence was extreme, but not so extreme so as to result in floor/ceiling effects.

Another reason for the failure of the black sheep effect to obtain may derive from certain features of the present experimental procedure that could well have interfered with the mechanisms underlying the prediction. For example, since participants were run individually, were assured of the confidentiality and anonymity of their responses, and did not anticipate jury deliberation, there were no “others” who could associate these mock jurors with guilty-looking coreligionist defendants. The experimenter knew their verdicts, of course, but at the time those verdicts were expressed, participants had not yet revealed their own religious affiliation (and, as far as they knew, the experimenter did not presently and would not later possess this information). Finally, the nature of the verdict measures may have contributed to a lack of black sheep effects. The current dichotomous verdict measure may have lacked the sensitivity to detect the effect predicted by the black-sheep hypothesis.

Therefore, we conducted a second experiment in order to provide a stronger test of the black sheep hypothesis.

STUDY 2

Several improvements on the methods used in Study 1 were incorporated into Study 2. First, we attempted to intensify the strength-of-evidence manipulation.

This was accomplished by manipulating testimony so as to attain a higher conviction rate for the strong-evidence cases. A second improvement led participants to anticipate having to defend their individual juror preference before others in a jury deliberation. The theoretical logic underlying our "black sheep" prediction suggests that defending a black sheep from one's own flock is most threatening when one is presenting such a defense to members of the "out-flock." Indeed, Cooper and Jones (1969) have shown that similarity-harshness is most likely to appear when the individual anticipates that similarity will be negatively assessed by an outside observer. As suggested earlier, the act of dissociation from an obnoxious other may well stem from a concern for the individual's self-presentation (i.e., "what will another person think of me?"), and this concern should be greatest where there is potential of being perceived as a biased juror (i.e., exhibiting favoritism for similar defendants). Therefore, participants in Experiment 2 anticipated that most of the members of their jury would be drawn from the outgroup. (In addition, as noted in the following description of methods, it was also possible to run an additional control condition in which some jurors anticipated joining a jury in which the majority of jurors were in-group members.)

A revised verdict measure constituted the third difference between the two studies. Like Study 1, defendant guilt was measured using the standard dichotomous verdict question. However, subject jurors were also asked to report their confidence in the verdict rendered on a continuous scale. The latter confidence-in-verdict judgment was combined with the verdict to create a bipolar guilt rating scale, ranging from complete confidence in the defendant's innocence to complete confidence in his guilt. With respect to the latter's construct validity, Stasser and Davis (1981) have implicated changes in confidence as precursors to changes in verdict in jury deliberation.

Finally, group membership was defined through racial group membership rather than through religious group affiliation. The former is a visibly salient basis for group membership and identification (Crocker & Major, 1989). Furthermore, using race to define similarity allowed for denial of group membership less easily (e.g., as denominational differences might permit a juror to see a coreligionist as in an outgroup).

An additional motive for making race the basis of juror-defendant similarity/dissimilarity was to contribute to the important literature examining the role of racial factors in jury trials. Legal institutions, like every other significant institution of American culture, have wrestled with the problem of identifying and combatting racial prejudice. For example, one controversy in case law has been whether near or total exclusion of racial minorities from juries, either historically through jury panel composition (Fukurai, Butler, & Krooth, 1992), or, more commonly today, through the exercise of peremptory challenges (Johnson, 1985; *Batson v. Kentucky*, 1986), represents a denial of an unbiased jury of one's peers. Public perception of the integrity of the jury system has clearly been affected by trials with racial overtones (e.g., the trial of the white Los Angeles policemen charged with beating Rodney King; the O. J. Simpson trial).

There are a number of laboratory and field studies that have examined the effects of juror and defendant race (for reviews, see Dane & Wrightsman, 1982;

Dodge, 1990; Foley, 1993; Johnson, 1985). The results of these studies are complex, but they do tend to reinforce the belief that juries are not always color blind. One of the complexities to emerge from this literature is that the effects of the race of the defendant are sometimes moderated by the race of the victim of the crime (e.g., Klein & Creech, 1982; Pfeifer & Ogloff, 1991; Ugwuegbu, 1979; Baldus et al., 1990; Bowers, Pierce, & McDevitt, 1984; Gross & Mauro, 1989; Foley & Rasche, 1979). For example, Ugwuegbu (1979) found White mock jurors punished a Black defendant more severely when his victim was White rather than Black. Black mock jurors demonstrated a parallel pattern. Such findings led us also to include a manipulation of victim race in Study 2.

Method

Participants

The participants were 168 (52 males, 116 female) students attending community colleges in the Detroit Metropolitan area. Arrangements were made with instructors to conduct the study as a classroom demonstration. The average age of the participants was 25.35 years ($SD = 8.63$).

Design and Stimulus Trials

The plan for the basic experimental design was a 2 (Race of juror) \times 2 (Race of defendant) \times 2 (Race of victim) \times 2 (Strength of evidence) factorial. Drawing on the pilot work that preceded Study 1, elements of the same four child-molestation cases were altered⁴ in order to make the apparent guilt of the defendants more (but not too) extreme in the strong-evidence condition (in which the key, "black sheep" defendant was presented). New pilot data were encouraging—the mean conviction rate of the two strong cases was 71.4% (81% for the Harris case; 62% for the Burger case); the corresponding conviction rate for the two weak cases was 35.6% (38% for the Williams case; 33% for the Small case).

Participants in Study 2 were not able to consider all four cases as in the preceding experiment because of time limitations. Therefore each subject considered only two cases. One of those cases was a strong case and the other was a weak case. Two pairings of strong and weak case were used; approximately half of the participants considered one pair of cases, the rest considered the other. Race of defendant was manipulated between subjects; that is, the defendants for both cases considered by a subject were Black or were White. We also counter-balanced across participants the order in which a subject considered the strong vs. the weak case. The between-subjects race-of-victim factor was crossed with the race-of-defendant factor. The alleged victim always had the same race across the

⁴ These alterations were small changes in testimony, which were unrelated to the independent variables of defendant or victim race. These alterations all tended to weaken the defendants' alibis (e.g., a defendant explicitly admitted that no one could directly corroborate his alibi; an eyewitness expressed greater certainty about an fact which contradicted an alibi). Again, copies of these cases may be obtained from the authors on request.

pair of crimes; the victims were both White for approximately half of the participants and both Black for the other half.

For reasons noted above, we expected the "black sheep" effect to be most likely to occur when there were many outgroup members in the jury. For that reason, our initial plan was to lead all participants to expect to deliberate in a jury in which their own race would constitute a small minority. (Of course, for White mock jurors, this represents an atypical arrangement, at least in the usual American courtroom, Fukurai et al., 1992.) However, it was also of interest to determine directly whether the composition of the jury might moderate such an effect. Unfortunately, there were not enough students (particularly White students) available to us in our target subject population to completely cross the initially planned design with a new experimental factor of jury composition (i.e., majority of the jury will be Black vs. majority of the jury will be White). However, there were enough Black students available to add an external control condition in which Black jurors anticipated deliberating in a jury that would be predominantly Black.

Procedure

The study was run during regular class meeting times in the classrooms. A stack of sealed envelopes containing stimulus materials had been shuffled randomly prior to each experimental session. After passing these out,⁵ the experimenter introduced the study as concerned with how jurors reach a verdict. The cases were described as summaries of two trials tried in New York City within the last 36 months. The experimental cover story suggested that comparisons of New York and Michigan jurors' verdicts was of interest. The nature of the charges in these cases (viz., sexual molestation of a minor) was then described. It was noted that the trial summaries did not contain any graphic or sensational descriptions and that none of the victims had been hurt physically. However, because of the potential for such a crime to upset some jurors, participants' freedom to discontinue participation at any time was reiterated. Participants were told that at the present time, only individual juror decisions would be made, but that at a later time, 12-person mock-juries would meet and deliberate the cases. These juries were allegedly to be composed to be as similar as possible to the juries that had decided the original cases. Before opening the envelopes and completing the materials, participants were reminded not to interact with one another. The experimenter carefully monitored compliance with these instructions.

Within each envelope there was a booklet. In order to simulate voir dire, the cover sheet of this booklet asked for background information on the juror including sex, age, race, prior jury experience, and acquaintance or relation with law enforcement personnel. The instructions indicated that this information would be used to "compose juries that are representative of the New York district where the actual trials took place." The manipulation of jury composition was achieved

⁵ Because of the race-of-juror/jury manipulation, there were actually two stacks, one for White mock jurors and another one (placed on top of it) for Black jurors. Black jurors received envelopes off the top of the complete stack, whereas White jurors received envelopes off the bottom of the stack.

at this point. All participants were told that there would be 6 males and 6 females on their jury. All White participants and approximately half of the Black participants (viz., participants in the external control conditions) were told that “in that district the population is about 18% White and 82% Black, so your jury will have 2 White jurors and 10 Black jurors.” The other half of the Black participants had these values reversed, so that they expected their jury to be composed of 2 Black and 10 White jurors.

After reading a brief judge’s charge, similar to that used in Study 1, participants proceeded to read and decide each of the two cases they had received. Three judgments were requested of the subject for each case—a dichotomous verdict (guilty or not guilty beyond a reasonable doubt), a certainty rating for that verdict (on a 9-point bipolar scale), and a sentence recommendation (which could vary from 0 to 600 months in prison). The last sheet of the booklet contained questions ostensibly checking participants’ understanding of instructions and memory about the details of the last (i.e., second) case considered. The key questions for our purposes were those that asked about the race of the defendant and the racial composition of the jury in which they would later deliberate.

After all students in the class had completed the entire booklet, the class was fully debriefed as to the true purposes of the study and all deceptions were revealed (e.g., that no deliberation would, in fact, be occurring). After explaining and stressing the importance of not discussing the study with others, the participants were thanked and excused.

Results

Manipulation Checks

Participants were asked to recall the race of the defendant and of the victim for the last case they considered. Analysis of the proportion of correct responses indicated reasonable recall accuracy. Defendant race was correctly recalled 81.3% of the time; victim race, 89.0% of the time. Participants were also asked about the racial composition of their jury. Once again, most participants were accurate. In all, 75.1% of participants remembered correctly whether or not theirs was the majority or minority race on the jury.⁶

Primary Analyses: Jurors Anticipating a Minority of Ingroup Members in the Jury

Guilt scale scores (GS) were created by compounding the participants’ verdicts with their verdict confidence ratings. This procedure yielded a 18-point measure with the endpoints “–9” (*complete certainty of defendant’s innocence*)

⁶ It was possible that participants who did not correctly recall this information could have responded differently from those who could recall it accurately to the primary dependent variables. However, analyses that excluded the former participants yielded a pattern of results identical to the pattern obtained when they were included. Therefore, analyses reported in the text include all participants.

Table 1. Mean Guilt Scale Scores (GS) by Subject Race, Defendant and Victim In/Outgroup Status, and Strength of Evidence

Jurors Anticipating Being in a Racial <i>Minority</i> in their Jury				
Strength of evidence	Black victim		White victim	
	Black defendant	White defendant	Black defendant	White defendant
<i>Black mock jurors</i>				
Weak	-3.90	-2.70	-3.29	-5.88
Strong	+6.62	+4.85	-0.14	+3.78
<i>White mock jurors</i>				
Weak	-2.21	-6.30	-4.38	-4.15
Strong	+2.00	+4.90	+2.81	+6.54
Jurors Anticipating Being in a Racial <i>Majority</i> in their Jury				
<i>Black mock jurors</i>				
Weak	-3.80	-5.00	-5.14	-3.33
Strong	+3.00	+8.67	+1.86	+1.00

to “+9” (*complete certainty of defendant’s guilt*). A mean GS score near “0” indicated maximum subject uncertainty about which verdict to favor.

In our initial analyses, we considered our initial basic design. That is, we focused on those participants led to believe that their race had *minority* status on the jury. This initial analysis on GS scores⁷ employed a 2 (Subject race) × 2 (Defendant race) × 2 (Victim race) × 2 (Strength of evidence) factorial design with repeated measures on the final factor. The relevant mean GS scores are displayed in the top two panels of Table 1.

As expected, the defendants were seen as guiltier when the evidence against them was strong (strength-of-evidence main effect $F(1,126) = 10.164, p < .0001$). However, this effect was not unmoderated, as revealed by the significant four-way interaction effect, $F(1,126) = 4.21, p < .05$. Inspection of Table 1 reveals one clear exception to the effectiveness of the strength-of-evidence manipulation—Black jurors trying a Black defendant accused of molesting a White child.

For this particular combination of juror, defendant, and victim race, the strength-of-evidence manipulation did not have a significant simple effect on GS scores, $F(1,126) = 2.15, ns$. Even with what was generally perceived as strong evidence against the defendant, these jurors were actually on the acquittal side of the symmetric bipolar guilt scale (mean GS = $-.14$). In our discussion below we speculate on why this may have occurred. The more important point, at present, is that it is not meaningful to test the “black sheep” hypothesis unless the manipulation of strength of evidence is clearly effective. That is, we do not expect

⁷ The 18-point guilt scale measure was utilized in the hope that it would provide a more sensitive measure of juror’s willingness to convict than an analysis of the simple dichotomous verdict. Nonetheless, a parallel ANOVA was also conducted on the dichotomous verdicts (see Lunney, 1970, for a justification of ANOVA on dichotomous dependent variables). This analysis showed that the pattern of results for the two measures was identical. Only the results for the guilt scale are presented in the text.

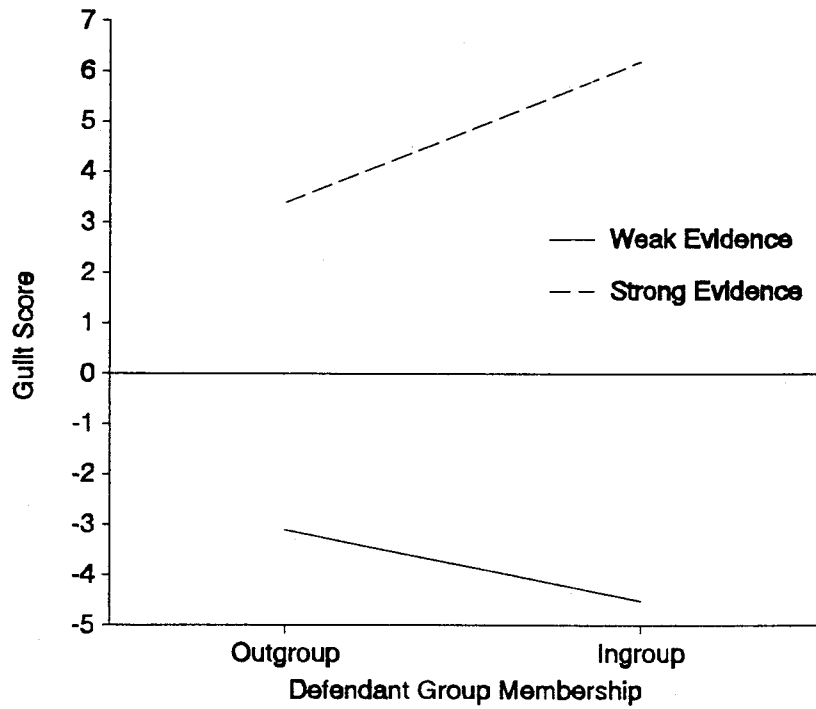


Fig. 2. Strength of evidence \times Defendant group membership interaction effect on guilt scores: Study 2.

in-group members to be harsher toward a defendant from the ingroup *unless* that defendant looks guilty. Clearly, that was *not* the case in the Black juror/Black defendant/White victim condition.

However, the strength-of-evidence manipulation was clearly effective elsewhere. Hence, it is possible and meaningful to test the black-sheep prediction (i.e., to compare reactions to same vs. opposite race defendants) under the remaining juror race/victim race combinations. And the pattern predicted by the black sheep hypothesis did obtain in those conditions. This can be seen most easily by recoding the recombination of juror race and defendant race as a dichotomous "Racial similarity" factor (1 = juror and defendant are of the same race; 2 = juror and defendant are of different races). The resulting ANOVA⁸ yielded a significant racial similarity \times strength of evidence interaction effect, $F(1,126) = 6.26, p < .02$, which is plotted in Figure 2. As Figure 2 shows, when the evidence was weak, there was the usual favoritism for the ingroup member. But when the evidence against the defendant was strong, the opposite occurred; the ingroup "black sheep" was judged more harshly than the corresponding out-

⁸ That is, a 2 (Racial similarity) \times 2 (Strength of evidence) \times 3 (Juror-Victim combination: Black juror/Black victim, White juror/Black victim, White juror/White victim) on GS scores. (Again, the Black juror/White victim combination is not included in this analysis because of the evident disruption of the strength-of-evidence treatment within this combination.)

group defendant. This interaction effect was uniform across the three combinations of juror and victim race examined in this analysis, $F(2,126) < 1$. That is, both Black jurors (when the victim was Black) and White jurors (whether the victim was Black or White) were comparably harsher toward a same-race defendant than toward an other-race defendant.

Secondary Analyses: Jurors Anticipating a Minority of Ingroup Members in the Jury

The preceding analysis confirmed that mock jurors will indeed judge another ingroup member more harshly than an outgroup member under certain conditions. Those conditions were (a) the evidence against the ingroup defendant must be seen as strong, and (b) jurors must anticipate that most of the jury will be composed of outgroup members. But is the latter condition critical? Is it the prospect of trying to defend the indefensible behavior of a fellow ingroup member *in a jury containing at most a single ingroup ally* that leads to rejection of the ingroup's "black sheep"? The behavior of participants in the external control condition (where jurors believed that most of the rest of the jury would be *ingroup* members) could speak to this question. A 2 (Racial similarity) \times 2 (Race of victim) \times 2 (Strength of evidence) ANOVA was performed in this external control condition. The relevant means are presented in the bottom panel of Table 1. The expected main effect for strength of evidence was significant, $F(1,126) = 23.04$, $p < .001$. Although there was some hint of the same neutralization of this manipulation for these control Black jurors as was observed among the experimental Black jurors (viz., when they considered a crime involving a white victim), victim race did not enter into a significant interaction effect (either the two-way or the three-way effect) with the strength of evidence manipulation. Thus, the necessary condition of an effective strength-of-evidence manipulation seems to have been met for all the control mock jurors. However, these control jurors did *not* exhibit the key racial similarity \times strength of evidence interaction effect observed for the experimental jurors (and exhibited in Figure 2), $F(1,126) < 1$. (In fact, the nonsignificant trend for the simple interaction effect in the Black victim condition is in the opposite direction $F(1,126) = 1.98$, $p < .17$.) These analyses suggest that jury composition may indeed be a relevant moderating variable for the "black sheep" effect.⁹

Ancillary Analyses: Sentencing

In addition to rendering verdict judgments, participants were asked to imagine for the moment that the defendant had been found guilty and to recommend a

⁹ A more direct comparison between experimental and control participants reinforced this conclusion. When the experimental participants whose data are plotted in Figure 2 were contrasted directly with the control participants in a 2 (Jury composition: own race in majority vs. own race in minority) \times 2 (Racial similarity: defendant in ingroup vs. defendant in outgroup) \times 2 (Strength of evidence) analysis of variance, the three-way interaction effect, $F(1,126) = 3.52$, $p < .06$, confirmed that the pattern of Figure 2 was *not* replicated among the control participants.

sentence. Sentences were converted to years and submitted to an omnibus repeated-measures ANOVA, which included all participants. Analyses indicated only one significant effect—a main effect for defendant race, $F(1,125) = 3.96$, $p < .05$. White defendants received significantly longer sentences ($M = 12.00$ years) than did Black defendants ($M = 7.11$ years).

Discussion

Taken together, Studies 1 and 2 present the first empirical evidence (to our knowledge) that, under certain conditions, juror-defendant similarity can lead to a type of black sheep effect. Consistent with the ingroup-favorability bias, when evidence favoring conviction was weak or even moderately strong (as in Study 1), jurors reached a “not guilty” verdict more readily for ingroup than outgroup defendants. However, in Study 2, when the evidence was perceived as overwhelmingly in favor of a guilty verdict, and jurors (whether Black or White) anticipated their group to be in the minority in the jury, ingroup (i.e., same race) defendants were judged as more guilty than outgroup (i.e., other race) defendants.

The ingroup favorability found in weak or moderately strong evidence conditions replicates the similarity-lenieny effects reported elsewhere. According to Kalven and Zeisel (1966; see also Baumeister & Darley, 1982), jurors' biases will most likely influence their judgments when the evidence presented in the courtroom is ambiguous. Under these circumstances, jurors will make use of any and all information available to them, including extralegal information. For example, given inconclusive evidence, jurors have been found to make use of inadmissible evidence (Sue, Smith, & Caldwell, 1973), and to rely on sexist stereotypes (Reskin & Visher, 1986) in rendering judgments. More relevant to the present study, investigators of factors influencing rape cases (e.g., Klein & Crech, 1982; Ugwuegbu, 1979) have shown strong juror lenieny toward same-race defendants when given marginal evidence.

The black sheep effect observed in Study 2 is not consistent with Kalven and Zeisel's hypothesis, however. According to their liberation hypothesis, unambiguous admissible evidence that argues strongly in favor of conviction should result in the suspension of juror biases and result in a verdict based simply on that evidence. Yet, the relative harshness with which jurors in the present study evaluated the ingroup defendant in the strong evidence conditions still reflects an extralegal bias. According to Marques et al. (1988), the dissociation of the self from a negatively valued ingroup member “should be considered a ‘sophisticated’ form of ingroup favoritism” (p. 8). Thus, in this sense, jurors exhibited a form of ingroup “favorability” in both the weak and strong evidence cases.

It is worth noting that while evidentiary strength influenced verdict, it did not affect sentencing. Bray and Kerr (1979) point out that such discrepancies between verdict and sentencing are not unusual, and, under certain circumstances (e.g., Kerr, 1978), may even be expected. In the present case, although strength of evidence differed across conditions, the crime of child molestation itself was likely to be perceived as heinous by everyone, regardless of group membership. As a result, strength of evidence could have affected jurors' confidence in verdict

without affecting their desire to penalize a convicted defendant. Future research might well investigate whether the black sheep effect is also manifest in penalties for crimes of lesser severity.

It is not clear why White defendants received uniformly longer sentences than Blacks. It has been shown (Sunnafrank & Fontes, 1983) that child molestation and child sexual crimes are associated stereotypically more with Whites as a social group than with Blacks. Perhaps, as a result, molestation by a Black might be more readily attributed to situational or idiosyncratic personal factors (e.g., high stress), whereas for Whites it might be attributed to more permanent dispositional factors. Such stereotypic beliefs about group dispositions have been shown (e.g., Bodenhausen & Wyer, 1985; Gordon, 1993) to affect juror judgment.

The findings from Study 2 implicate the importance of self-presentational concerns for the black sheep effect. Subject-jurors (of both races) were harsher toward similar defendants only when they expected to be in a racial minority in the jury. When one is surrounded by other ingroup members, an ingroup deviant can be viewed as a black sheep who, although base, is not a threat to the entire group's identity. When, however, one is surrounded by outgroup members who might be unfamiliar with or even hostile toward the ingroup, ingroup members will probably want to put on the best face possible. This can mean, in the words of Cooper and Jones (1969), adopting a "strategy to avoid being miscast" as similar to the ingroup deviant.

It is therefore interesting that when they constituted a minority in the jury, Black subject-jurors did not exhibit a black sheep effect when the victim was White (see the Black juror/Black defendant/White victim condition vs. the corresponding White defendant condition in Table 1). It is noteworthy that both the victim *and* the majority of other jurors were White in this condition. Black jurors might well have been concerned about a Black defendant's chances for a fair trial under such conditions. In addition, race may also have been especially salient in this condition; considerable research has shown that minority status strengthens one's identification with the ingroup (Gerard & Hoyt, 1974; Mummendey & Simon, 1989). Black jurors (already in the minority in their community) may have felt a particularly strong identification with their racial group and with the Black defendant in this condition, and felt obliged to protect the defendant's interests when everything seemed stacked against him.¹⁰

Although the present findings are certainly preliminary and need replication (especially in field settings) before confident policy recommendations could be made, the present findings do carry interesting implications for courtroom procedures, especially *voir dire*. First and foremost, they suggest that apart from con-

¹⁰ One might also interpret such "bending over backwards" (to give a similar and apparently disadvantaged defendant the benefit of any possible doubt) as another instance of the similarity-leniency effect. However, such an interpretation implies the same similarity-leniency effect in the corresponding weak-evidence condition (*viz.*, Black jurors/White victim/weak-evidence), and this did not occur. For this and the reasons advanced in the Results section, we think it is more reasonable simply to conclude that the boundary conditions required to test the black-sheep hypothesis were not met in these conditions.

stitutional barriers to basing peremptory challenges upon juror group membership (e.g., *Batson v. Kentucky*, 1986), attorneys should be cautious in applying the similarity–leniency rule to jury selection in all trials. When the evidence against a defendant is very strong and the defendant and potential jurors are members of the same minority group, a defense attorney may, in fact, end up hurting his/her prospects for acquittal by selecting jurors based on their similarity to the defendant. (Of course, the black sheep effect has precisely the opposite implications for a prosecutor exercising peremptory challenges.)

Several questions regarding the black sheep effect ought to be addressed further. First, will the present effects based on race generalize to other social categories? For example, another reason for the failure to obtain a black sheep effect in Study 1 may be that religion may not be as important a basis of social identification as race in our culture. Second, in Study 2, participants only anticipated jury deliberation but the primary data remain jurors' pre-deliberation verdict preferences. While there is considerable evidence (e.g., Stasser et al., 1982, 1989) that such juror preferences are highly predictive of jury verdict, it is also plausible that the psychological processes presumed to underlie the black sheep effect (e.g., self-presentational concerns, social identification concerns) would change in significant ways in a group context (i.e., within deliberating juries). For example, it has been argued (Golash, 1992; Lipton, 1983) that the presence of minority jurors can inhibit majority members from making arguments that could be construed as prejudiced. Third, future research should consider additional limiting factors on a juridic black sheep effect. For example, jurors might be less willing to be harsh toward an ingroup defendant if the penalties of the crime were high (e.g., death penalty) or (as speculated above) if the prospects for a verdict biased by defendant's group membership (e.g. racially biased) were particularly high. Finally, we began this article by noting that the black sheep effect can be predicted from several different theoretical perspectives, each suggesting a somewhat different underlying causal mechanism. Some of these mechanisms emphasize motivational processes (e.g., the explanations based on social identity theory or defensive attribution), whereas others stress more cognitive processes (e.g., explanations based on social comparison or expectancy violation). It would obviously contribute much to both understanding and application of this effect if there were more direct evidence on the key mediating process(es).

Although some encouraging research (e.g., Johnson, 1985; LaFree, 1980; Pfeifer & Ogloff, 1991) suggests that racial bias by jurors is not inevitable, our research also contributes to a less sanguine literature (e.g., Bullock, 1961; Gerard & Terry, 1970; LaFree, Reskin, & Visser, 1985) that suggests that justice is not color-blind. Consistent with some other work employing both mock juries (Klein & Creech, 1982; Sweeney & Haney, 1992; Ugwuegbu, 1979) and actual juries (e.g., Baldus et al., 1990; Bowers et al., 1984), our results corroborate that race-related juror biases may depend heavily upon the specific racial combination of defendant and victim and the specific racial composition of the jury.

It should be reiterated that the use of jury simulation in the present study may limit its generalizability to actual jury behavior. Jury simulations are often criticized because of their artificiality (see Weiten & Diamond, 1979 for a review). On

the other hand, we would argue that (a) even a fairly artificial simulation can be important for testing and developing psychological theory (cf. Bray & Kerr, 1982), and (b) as Lewin (1951) pointed out, there is nothing so practical as a good theory. While admittedly not a perfect simulation of courtroom reality, the jury simulation used here did permit an empirical first test within a juror decision-making setting of an interesting prediction suggested by several social psychological theories. Moreover, the validity or invalidity of this prediction carries implications not only for attorney behavior in voir dire but for case law and legal policy. A number of legal scholars (Johnson, 1985; Sweeney & Haney, 1992) have called for a legal policy dictating that juries trying a minority defendant always include jurors of similar racial or ethnic background. As we have shown, given the complex nature of juror-defendant similarity (even in simulated situations), such a policy may not always serve best the interests of justice.

REFERENCES

- Amato, P. R. (1979). Juror-defendant similarity and the assessment of guilt in politically motivated crimes. *Australian Journal of Psychology*, *31*, 79–88.
- Bailey, F. L., & Rothblatt, H. B. (1985). *Successful techniques for criminal trials*. Rochester, NY: Lawyers Co-Operative.
- Baldus, D. C., Woodworth, G., & Pulaski, C. A., Jr. (1990). *Equal justice and the death penalty: A legal and empirical analysis*. Boston: Northeastern University Press.
- Batson v. Kentucky (1986). *United States Supreme Court Reports*, 476, 79.
- Belli, M. (1963). *Modern trials* (abridged ed.). Indianapolis: Bobbs-Merrill.
- Bernard, J. L. (1979). Interaction between the race of the defendant and that of the jurors in determining verdicts. *Law and Psychology Review*, *5*, 103–111.
- Bodenhausen, G. V., & Wyer, R. S. (1985). Effects of stereotypes on decision making and information-processing strategies. *Journal of Personality and Social Psychology*, *48*, 267–282.
- Bowers, W., & Pierce, G. (1980). Arbitrariness and discrimination under post-Furman Statutes. *Crime and Delinquency*, *26*, 563–635.
- Bowers, W. J., Pierce, G. L., & McDevitt, J. F. (1984). *Legal homicide: Death as punishment in America, 1864–1982*. Boston: Northeastern University Press.
- Bray, R. M., & Kerr, N. L. (1979). Use of the simulation method in the study of jury behavior: Some methodological considerations. *Law and Human Behavior*, *3*, 107–119.
- Bray, R. M., & Kerr, N. L. (1982). Methodological considerations in the study of the psychology of the courtroom. In N. L. Kerr & R. M. Bray (Eds.), *The psychology of the courtroom* (pp. 287–324). New York: Academic Press.
- Bray, R. M., & Noble, A. (1978). Authoritarianism and decisions of mock jurors: Evidence of jury bias and group polarization. *Journal of Personality and Social Psychology*, *36*, 1424–1430.
- Brewer, M. B. (1979). In-group bias in the minimal intergroup situation: A cognitive-motivational analysis. *Psychological Bulletin*, *86*, 397–324.
- Bullock, H. A. (1961). Significance of the racial factor in the length of prison sentences. *Journal of Criminal Law, Criminology, and Police Science*, *52*, 411–417.
- Busch, F. (1959). *Law and tactics in jury trials* (Vol. 1). Indianapolis: Bobbs-Merrill.
- Byrne, D. (1971). *The attraction paradigm*. New York: Academic Press.
- Byrne, D., Clore, G. L., & Smeaton, G. (1986). The attraction hypothesis: Do similar attitudes affect anything? *Journal of Personality and Social Psychology*, *51*, 1167–1170.
- Byrne, D., & Wong, J. T. (1962). Racial prejudice, interpersonal attraction, and assumed similarity of attitudes. *Journal of Abnormal and Social Psychology*, *65*, 246–253.

- Chaikin, A. L., & Darley, J. M. (1973). Victim or perpetrator: Defensive attribution of responsibility and the need for order and justice. *Journal of Personality and Social Psychology*, 25, 268–273.
- Cooper, J., & Jones, E. E. (1969). Opinion divergence as a strategy to avoid being miscast. *Journal of Personality and Social Psychology*, 13, 23–30.
- Crocker, J., & Major, B. (1989). Social stigma and self-esteem: The self-protective properties of stigma. *Psychological Review*, 96, 608–630.
- Dane, F. C., & Wrightsman, L. S. (1982). Effects of defendants' and victims' characteristics on jurors' verdicts. In N. L. Kerr, & R. M. Bray (Eds.), *The psychology of the courtroom* (pp. 83–118). New York: Academic Press.
- Davis, J. H., Bray, R. M., & Holt, R. (1977). The empirical study of decision processes in juries: A critical review. In J. Tapp & F. Levine (Eds.), *Law, justice, and the individual in society* (pp. 326–361). New York: Holt.
- Dodge, L. (1990). *Death penalty sentencing* (Report No. B-236876). Washington, DC: U.S. General Accounting Office.
- Duncan, B. L. (1976). Differential social perception and attribution of intergroup violence: Testing the lower limits of stereotyping of Blacks. *Journal of Personality and Social Psychology*, 37, 621–634.
- Feild, H. S. (1979). Rape trials and jurors' decisions: A psycholegal analysis of the effects of victim, defendant, and case characteristics. *Law and Human Behavior*, 3, 261–284.
- Feinberg, S. (1970). The analysis of multidimensional contingency tables. *Ecology*, 51, 419–433.
- Fitzgerald, R., & Ellsworth, P. (1984). Due process vs. crime control: Death qualification and jury attitudes. *Law and Human Behavior*, 8, 31–51.
- Foley, L. A. (1993). *A psychological view of the legal system*. Madison, WI: Brown & Benchmark.
- Foley, L. A., & Rasche, C. E. (1979). The effect of race on sentence, actual time served, and final disposition of female offenders. In J. A. Conley (Ed.), *Theory and research in criminal justice: Current perspectives*. Cincinnati: Anderson.
- Fukurai, H., Butler, E., & Krooth, R. (1992). *Race and the jury*. New York: Plenum.
- Fulero, S. M., & Penrod, S. D. (1990). Attorney jury selection folklore: What do they think and how can psychologists help? *Forensic Reports*, 3, 233–259.
- Gerard, H. B., & Hoyt, M. F. (1974). Distinctiveness of social categorization and attitude toward ingroup members. *Journal of Personality and Social Psychology*, 29, 836–842.
- Gerard, J. B., & Terry, T. R. (1970). Discrimination against Negroes in the administration of criminal law in Missouri. *Washington State University Law Quarterly*, 415–437.
- Gerbasi, K. C., & Zuckerman, M. (April, 1975). An experimental investigation of jury biasing factors. Paper presented at the meetings of the Eastern Psychological Association, New York City.
- Gobert, J. J., & Jordan, W. E. (1990). *Jury selection: The law, art, and science of selecting a jury* (2nd ed.). New York: McGraw-Hill.
- Golash, D. (1992). Race, fairness, and jury selection. *Behavioral Sciences and the Law*, 10, 155–177.
- Gordon, R. A. (1993). The effect of strong versus weak evidence on the assessment of race stereotypic and race nonstereotypic crimes. *Journal of Applied Social Psychology*, 23, 734–749.
- Griffitt, W., & Jackson, T. (1973). Simulated jury decisions: The influence of jury-defendant attitude similarity-dissimilarity. *Social Behavior and Personality*, 1, 1–7.
- Gross, S. R., & Mauro, R. (1989). *Death and discrimination: Racial disparities in criminal sentencing*. Boston: Northeastern University Press.
- Hans, V. P. (1992). Jury decision making. In D. K. Kagehiro & W. S. Laufer (Eds.), *Handbook of psychology and law* (pp. 56–76). New York: Springer-Verlag.
- Hans, V. P., & Vidmar, N. (1986). *Judging the jury*. New York: Plenum.
- Harvey, J. H., Harris, B., & Barnes, R. (1975). Actor-observer differences in the perceptions of responsibility and freedom. *Journal of Personality and Social Psychology*, 32, 22–28.
- Howard, J. A., & Rothbart, M. (1980). Social categorization and memory for ingroup and outgroup behavior. *Journal of Personality and Social Psychology*, 38, 301–310.
- Jackson, L. A., & Hymes, R. W. (1985). Gender and social categorization: Familiarity and ingroup polarization in recall and evaluation. *Journal of Social Psychology*, 125, 81–88.
- Johnson, S. L. (1985). Black innocence and the white jury. *Michigan Law Review*, 83, 1611–1708.
- Jones, E. E., & deCharms, R. (1957). Changes in social perception as a function of the personal relevance of behavior. *Sociometry*, 20, 75–85.

- Jordan, W. E. (1980). *Jury selection*. New York: McGraw-Hill.
- Jussim, L., Coleman, L. M., & Lerch, L. (1987). The nature of stereotypes: A comparison and integration of three theories. *Journal of Personality and Social Psychology*, *52*, 536–546.
- Kalven, H., & Zeisel, H. *The American jury*. Boston: Little Brown.
- Kauffman, R. A., & Ryckman, R. M. (1979). Effects of locus-of-control, outcome severity, and attitudinal similarity of defendant on attributions of criminal responsibilities. *Personality and Social Psychology Bulletin*, *5*, 340–343.
- Keeton, R. E. (1973). *Trial tactics and methods* (2nd ed.). Boston: Little, Brown.
- Kerr, N. L. (1978). Severity of prescribed penalty and mock jurors' verdicts. *Journal of Personality and Social Psychology*, *36*, 1431–1442.
- Klein, K., & Creech, B. (1982). Race, rape and bias: Distortion of prior odds and meanings changes. *Basic and applied social psychology*, *3*, 21–33.
- LaFree, G. D. (1980). The effect of sexual stratification by race on official reactions to rape. *American Sociological Review*, *45*, 842–854.
- LaFree, G. D., Reskin, B. F., & Visser, C. A. (1985). Jurors' responses to victims' behavior and legal issues in sexual assault trials. *Social Problems*, *32*, 389–407.
- Lauderdale, P., Smith-Cunnie, P., Parker, J., & Inverarity, J. (1984). External threat and the definition of deviance. *Journal of Personality and Social Psychology*, *46*, 1058–1068.
- Lewin, K. (1951). Problems of research in psychology. In *Field theory in social science: Selected theoretical papers* (pp. 155–169). New York: Harper & Brothers.
- Lipton, J. P. (1983). Racism in the jury box: The Hispanic defendant. *Hispanic Journal of Behavioral Sciences*, *5*, 275–290.
- Lowe, C. A., & Medway, F. J. (1976). Effects of valence, severity, and relevance on responsibility and dispositional attribution. *Journal of Personality*, *44*, 518–538.
- Lunney, G. H. (1970). Using analysis of variance with a dichotomous dependent variable: An empirical study. *Journal of Educational Measurement*, *7*, 263–269.
- Marques, J. M. (1990). The black sheep effect: Outgroup homogeneity in social comparison settings. In D. Abrams & M. Hogg (Eds.), *Social identity theory: Constructive and critical advances* (pp. 131–151). London: Harvester Wheatsheaf.
- Marques, J. M., & Yzerbyt, V. Y. (1988). The black sheep effect: Judgmental extremity towards ingroup members in inter- and intragroup situations. *European Journal of Social Psychology*, *18*, 287–292.
- Marques, J. M., Yzerbyt, V. Y., & Leyens, J. P. (1988). The black sheep effect: Extremity of judgments towards in-group members as a function of group identification. *European Journal of Social Psychology*, *18*, 1–16.
- McKillip, J., & Posavac, E. J. (1975). Judgments of responsibility for an accident. *Journal of Personality*, *43*, 248–265.
- Medway, F. J., & Lowe, C. A. (1975). Effects of outcome valence and severity on attribution of responsibility. *Psychological Reports*, *36*, 239–246.
- Mettee, D. R., & Riskin, J. (1974). Size of defeat and liking for superior and similar ability competitors. *Journal of Experimental Social Psychology*, *10*, 333–351.
- Mettee, D. R., & Wilkins, P. C. (1972). When similarity 'hurts': Effects of perceived ability and a humorous blunder upon interpersonal attractiveness. *Journal of Personality and Social Psychology*, *22*, 246–258.
- Mitchell, H. E., & Byrne, D. (1973). The defendant's dilemma: Effects of jurors' attitudes and authoritarian judicial decisions. *Journal of Personality and Social Psychology*, *25*, 123–129.
- Monahan, J., & Loftus, E. F. (1982). The psychology of law. *American Review of Psychology*, *33*, 441–475.
- Mummendey, A., & Simon, B. (1989). Better or just different? III: The impact of comparison dimension and relative ingroup size upon intergroup discrimination. *British Journal of Social Psychology*, *28*, 1–16.
- Novak, D. W., & Lerner, M. J. (1968). Rejection as a consequence of perceived similarity. *Journal of Personality and Social Psychology*, *9*, 147–152.
- Overall, J., & Spiegel, D. (1969). Concerning least squares analysis of experimental data. *Psychological Bulletin*, *72*, 311–322.

- Owen, I. (1973). *Defending criminal cases before juries*. Englewood Cliffs, NJ: Prentice-Hall.
- Pettigrew, T. F. (1979). The ultimate attribution error: Extending Allport's analysis of prejudice. *Personality and Social Psychology Bulletin*, 5, 461-476.
- Pfeifer, J. E., & Ogloff, J. R. (1991). Ambiguity and guilt determinations: A modern racism perspective. *Journal of Applied Social Psychology*, 21, 1713-1725.
- Reskin, B. F., & Visser, C. A. (1986). The impacts of evidence and extralegal factors in jurors' decisions. *Law and Society Review*, 20, 423-438.
- Schachter, S. (1951). Deviation, rejection, and communication. *Journal of Abnormal and Social Psychology*, 46, 190-207.
- Senn, D. J. (1971). Attraction as a function of similarity-dissimilarity in task performance. *Journal of Personality and Social Psychology*, 18, 120-123.
- Shaver, K. G. (1970). Defensive attribution: Effects of severity and relevance on the responsibility assigned for an accident. *Journal of Personality and Social Psychology*, 14, 101-113.
- Shaw, J. L., & Skolnick, P. (1971). Attribution of responsibility for a happy accident. *Journal of Personality and Social Psychology*, 18, 380-383.
- Smeaton, G., Byrne, D., & Murnen, S. K. (1989). The repulsion hypothesis revisited: Similarity irrelevance or dissimilarity bias? *Journal of Personality and Social Psychology*, 56, 54-59.
- Smith, C. R., Williams, L., & Willis, R. H. (1967). Race, sex, and belief as determinants of friendship acceptance. *Journal of Personality and Social Psychology*, 5, 127-137.
- Stasser, G., Kerr, N. L., & Bray, R. M. (1982). The social psychology of jury deliberations: Structure, process, and product. In N. L. Kerr & R. M. Bray (Eds.), *The psychology of the courtroom* (pp. 221-256). New York: Academic Press.
- Stasser, G., Kerr, N. L., & Davis, J. H. (1989). Influence processes and consensus models in decision-making groups. In P. Paulus (Ed.), *Psychology of group influence* (2nd Ed.) (pp. 279-326). Hillsdale, NJ: Erlbaum.
- Stasser, G., & Davis, J. H. (1981). Group decision making and social influence: A social interaction sequence model. *Psychological Review*, 88, 523-551.
- Stephan, C. W. (1974). Sex prejudice in jury stimulation. *Journal of Psychology*, 88, 305-312.
- Stephan, C. W., & Stephan, W. G. (1986). Habla Ingles? The effects of language translation on simulated juror decisions. *Journal of Applied Social Psychology*, 16, 577-589.
- Sue, S., Smith, R. E., & Caldwell, C. (1973). Effects of inadmissible evidence on the decision of simulated jurors: A moral dilemma. *Journal of Applied Social Psychology*, 3, 344-353.
- Sunnafrank, M., & Fontes, N. E. (1983). General and crime related racial stereotypes and influences on juridic decisions. *Cornell Journal of Social Relations*, 17, 1-15.
- Sweeney, L. T., & Haney, C. (1992). The influence of race on sentencing: A metaanalytic review of experimental studies. *Behavioral Sciences and the Law*, 10, 179-195.
- Tajfel, H., Flament, C., Billig, M., & Bundy, R. (1971). Social categorization and intergroup behavior. *European Journal of Social Psychology*, 1, 149-177.
- Tajfel, H., & Turner, J. C. (1986). The social identity theory of intergroup behavior. In S. Worchel & W. G. Austin (Eds.), *The social psychology of intergroup relations* (pp. 7-24). Chicago: Nelson-Hall.
- Taylor, D. M., & Jaggi, V. (1974). Ethnocentrism and causal attributions in a South Indian context. *Journal of Cross-Cultural Psychology*, 5, 162-171.
- Taylor, S. E., & Mettee, D. R. (1971). When similarity breeds contempt. *Journal of Personality and Social Psychology*, 20, 75-81.
- Ugwuegbu, D. C. (1979). Racial and evidential factors in juror attributions of legal responsibility. *Journal of Experimental Social Psychology*, 15, 133-146.
- Walster, E. (1966). Assignment of responsibility for an accident. *Journal of Personality and Social Psychology*, 3, 73-79.
- Weiten, W., & Diamond, S. S. (1979). A critical review of the jury simulation paradigm: The case of defendant characteristics. *Law and Human Behavior*, 3, 71-93.
- Wishman, S. (1986). *Anatomy of a jury: The system on trial*. New York: Times Books.
- Wrightsmann, L. S. (1987). *Psychology and the legal system*. Monterey, CA: Brooks-Cole.