Reviewing the Empirical Evidence on Jury Racism: Findings of Discrimination or Discriminatory Findings?

Jeffrey E. Pfeifer
University of Nebraska–Lincoln

Follow this and additional works at: https://digitalcommons.unl.edu/nlr

Recommended Citation
Available at: https://digitalcommons.unl.edu/nlr/vol69/iss1/8

This Article is brought to you for free and open access by the Law, College of at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Nebraska Law Review by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.
Comment*

Reviewing the Empirical Evidence on Jury Racism: Findings of Discrimination or Discriminatory Findings? ¹

TABLE OF CONTENTS

I. Introduction ............................................... 231

II. Empirical Evidence of Jury Racism 233
    A. Laboratory Research .................................. 233
    B. Archival Research .................................... 236

III. Evaluation of Evidence .................................... 241
    A. Limitations of Archival Research .................... 241
    B. Methodological Limitations of Laboratory Research . 243
    C. Predictive Limitations of Laboratory Research ....... 245
    D. Correspondence Limitations of Laboratory Research . 247

* The author wishes to thank James Ogloff and Jodi Wellnitz for their invaluable assistance in the preparation of this comment.

This work was supported by a Doctoral Fellowship awarded to the author by the Social Science and Humanities Research Council of Canada. Correspondence may be addressed to the author at the Law/Psychology Program, Department of Psychology, 209 Burnett Hall, University of Nebraska-Lincoln, Lincoln, Nebraska, 68588-0308.

1. Although the terms prejudice and discrimination are often used interchangeably within the context of racism, for the purposes of this comment, the terms racism, prejudice, and discrimination will be discussed according to the specific definitions put forth in 5 THE HANDBOOK OF SOCIAL PSYCHOLOGY (G. Lindzey & E. Aronson 2d ed. 1969). Racism is defined as any negative attitude or behavior which is directed at another based solely on their racial composition and some inferred knowledge (i.e., stereotypes), rather than on direct experience with the individual. Subsequently, prejudice may be defined as the holding of a negative attitude (i.e., the attitudinal component of racism) toward an individual based on his or her categorical membership (i.e., black, white, hispanic, etc.), while discrimination denotes an individual's negative behavior (i.e., the behavioral component of racism) toward another based on his or her categorical membership. See G. Allport, NATURE OF PREJUDICE (1954). This distinction between a prejudicial attitude and a discriminatory behavior is an important one which will be continuously referred to throughout this comment. See Harding, Proshansky, Kutner & Chein, Prejudice and Ethnic Relations, in 5 THE HANDBOOK OF SOCIAL PSYCHOLOGY 1 (G. Lindzey & E. Aronson 2d ed. 1969).
I. INTRODUCTION

The controversy surrounding the issue of juror competence has existed for almost as long as juries themselves.2 This controversy has been fueled by a number of critics who argue that jurors lack the necessary competence to skillfully assess the facts of a case and to set aside biases created by personal prejudices.3 Specifically, it has been suggested that juries, "fail to recall the evidence accurately, become confused by complex trials or those involving multiple defendants, and are often swayed by legally irrelevant information."4

Loh suggests that this critical view of juror competence and objectivity is rooted in a "historical distrust in the judgment of amateurs, untutored in legal subtleties and inexperienced in evaluating evidence."5 This distrust continues to be voiced by a number of scholars despite empirical studies which demonstrate that jurors tend to be relatively competent fact-finders.6 Wigmore, in one of the first empirical studies on jury decision-making, suggested that jurors were fairly adept at identifying the factual elements of a case, even in the face of testimonial errors by witnesses.7 Subjects in this study were asked to deliberate upon the testimony of witnesses who were reporting on a staged incident. The results indicated that during the course of deliberations, subjects identified more facts than would be expected from the testimonial errors.

Likewise, in their classic study on jury decision-making, Kalven and Zeisel found that there was a high degree of similarity between verdicts reached by juries and those a presiding judge would have reached.8 This study involved the analysis of data collected from approximately 500 judges across the United States. Judges were asked to complete a questionnaire following each jury trial they presided over and document the jury verdict as well as their personal verdict. Analysis of the subsequent 3,576 cases indicated a high degree of correspondence between verdicts reached by judges and juries. Specifically, it was found that judges agreed with 78% of the jury verdicts analyzed.

3. Id. at 4.
Additional analyses revealed that of the 22% of remaining cases where there were disagreements, trial evidence was not excessively difficult and judges did not point to that factor as a cause of the difference. These results suggest, therefore, that the behaviors of juries and judges are not that different in terms of case disposition.

Yet, despite the above findings, juror skill and objectivity continues to be an area of controversy for psycholegal scholars, especially regarding the issue of juror ability to overcome personal biases when considering the culpability of a defendant. As such, this comment will evaluate the empirical evidence employed by critics to support the contention that, "[j]urors are sentimentalists with bleeding hearts. They are gullible creatures, too often driven by emotion and too easily motivated by prejudice, anger, and pity." 9

This critical view of juror objectivity is premised upon the belief that jury verdicts are affected by a large number of extra-evidential trial factors which ultimately affect their decisions. 10 No single factor, however, has received as much attention as that of race and its effect on juror decisions. 11 One need only examine psychological, 12 sociological, 13 and legal literature 14 to see the extent to which the area has been investigated.

Although the evidence from the above literature, concerning the effect of racism on jury decision-making, is far from conclusive, a number of legal scholars and psychologists contend that juries are in-

---

herently biased against minority defendants. Individual who argue that jurors are not able to put aside their prejudicial attitudes when deliberating on a case have suggested that the "historical distrust" in the judgment of jurors has been extended into the specific domain of racism. These arguments have been so persuasive that relatively little additional empirical work has been conducted recently in the area of jury racism.

However, the lack of recent empirical research regarding the role of racism in juror decision-making should not be taken as a sign that the issue has been resolved. In order to determine whether jury racism is in fact playing a decisive role in our courtrooms, it is necessary to evaluate the empirical data which is being used to support this contention.

This comment, therefore, will evaluate both laboratory and archival research in order to illustrate that the available data regarding the role of racism in jury decision-making does not support the contention that juror objectivity is adversely affected by the inability to set aside personal prejudices. Section II of this comment will describe both the laboratory and archival research which has been conducted in the area of jury racism. Following this description, section III will evaluate the limitations of this research.

II. EMPIRICAL EVIDENCE OF JURY RACISM

A. Laboratory Research

There have been a number of laboratory studies conducted which suggest that mock jurors invariably assign disproportionate outcomes to black defendants. These studies indicate that mock jurors tend to assign harsher dispositions to black defendants in comparison to white defendants, especially if the cases involved a white victim.


17. That is, in order to ascertain whether prejudicial attitudes on the part of jurors are negatively affecting the actual legal decisions they are making regarding the guilt of the defendant.

18. See, e.g., Bernard, Interaction Between the Race of the Defendant and That of the Jurors in Determining Verdicts, 5 LAW & PSYCHOLOGY REV. 103 (1979); Feild, supra note 10; Foley & Chamblin, The Effect of Race and Personality on Mock Jurors' Decisions, 112 J. OF PSYCHOLOGY 47 (1982); Gray & Ashmore, Biasing Influence on Defendants' Characteristics on Simulated Sentencing, 38 PSYCHOLOGICAL REP. 727 (1976); Klein & Creech, Race, Rape, and Bias: Distortion of Prior Odds and Meaning Changes, 3 BASIC & APPLIED SOC. PSYCHOLOGY 21 (1982); McGlynn, Megas & Benson, Sex and Race as Factors Affecting the Attribution of Insanity in a Murder Trial, 93 J. PSYCHOLOGY 93 (1976); Ugwegbu, Ra-
For the most part, the above studies employ a similar procedure. Subjects, usually white college students (mock jurors), are asked to either view a videotape or read transcript segments of a trial in which the race of the defendant and victim are varied. After the trial presentation, subjects are asked to rate the guilt of the defendant and/or recommend a sentence. The results from these studies are then discussed in terms of their implications for actual courtroom trials.\(^{19}\)

Included in the above research are a number of studies which investigated the effect of the defendant’s race on juror decisions. One such study was conducted by Gray and Ashmore\(^ {20}\) who had subjects read a description of a crime in which the defendant’s race was varied. Subjects were told that the defendant was already found guilty and as such were asked to give their opinion regarding the sentence which should be imposed. The results indicated that the race and religious beliefs of the defendant played a significant role in the sentencing decisions of the subjects. Specifically, it was found that black defendants were assigned longer sentences than white defendants.

McGlynn, Megas, and Benson\(^ {21}\) also found a significant difference in guilt attributions based on the race of the defendant. The authors asked 208 white subjects to rate the guilt of a defendant after reading a murder trial summary. Subjects were asked to judge whether he was guilty or not guilty by reason of insanity. Results indicated that black defendants were found guilty in 69% of the cases while white defendants were found guilty in only 54% of the cases.

Bernard\(^ {22}\) attempted to provide a sense of increased realism in his research on jury racism by employing a complex design which varied the racial composition of the juries as well as varying the race of the defendant. A total of ten juries were formed with subjects who were asked to view a videotape of a mock trial in which a white or black defendant was charged with assault and battery.\(^ {23}\) After viewing the “trial,” subjects were asked to cast individual votes regarding the guilt of the defendant both before and after deliberations. Although the results of the initial vote were not significant, they do indicate a definite trend for white jurors to find black defendants guilty more often than white defendants. Further analysis of the results indicated that


\(^{20}\) Gray & Ashmore, supra note 18, at 727.

\(^{21}\) McGlynn, Megas & Benson, supra note 18, at 93.

\(^{22}\) Bernard, supra note 18.

\(^{23}\) The juries were comprised of two sets of each of the following racial compositions: 100% white; 75% white/25% black; 50% white/50% black; 25% white/75% black; and 100% black. Id. at 107.
racially mixed juries were less likely to illustrate a racist trend than those wholly composed of blacks or whites.

In addition to the above research, a number of studies on juror racism have also been conducted by varying the race of both the defendant and the victim. Klein and Creech for example, asked subjects to view a videotape of a simulated rape trial in which the race of the defendant and victim were varied. During the videotape subjects were asked a number of times to rate whether the evidence being presented was pro-defense or pro-prosecution. Following the videotape, subjects were asked to rate the guilt of the defendant on a five-point scale as well as recommend a sentence. According to these authors, the results suggest that subjects tend to distort their evaluations of evidence in favor of white victims such that "neutral" evidence was significantly more likely to be perceived as being pro-prosecution when the victim was white than when she was black. The results also indicated that there was a significant difference in sentencing recommendations. Black defendants convicted of raping a white victim were given significantly longer sentences than white defendants.

Using a similar manipulation, Foley and Chamblin presented subjects with an audiotape of a trial describing the case of a male defendant charged with sexual battery of a child. After listening to the trial, in which the race of the defendant and child were varied, subjects were asked to respond to a five-point guilt scale ranging from not guilty to guilty. Results indicate that white jurors are significantly more likely to find a black defendant guilty of battering a white victim than a white defendant who batters a white victim.

Like the above studies, Ugwuegbu suggested that his study, "showed very convincingly that the race of the defendant and the race of the victim inappropriately influenced the level of culpability the jurors ascribed to the defendant." In support of this observation, Ugwuegbu presented white mock jurors with a condensed transcript of an aggravated rape trial in which the race of the victim and defendant were varied. After reading the trial stimulus, subjects were asked to judge the responsibility of the defendant and whether the defendant acted intentionally. Subjects were also asked to rate the guilt of the defendant as well as to recommend a suitable sentence. These four responses were then combined to give an overall measure of defendant's culpability. The results indicated that black defendants were significantly more likely to be held culpable for their actions than white defendants, especially if the victim was white.

25. Foley & Chamblin, supra note 18, at 47.
26. Ugwuegbu, supra note 18, at 143.
27. Id. at 139.
Feild\textsuperscript{28} also found that the race of the victim and the defendant affected the decisions of mock jurors. Feild collected data from 896 white subjects who read a transcript of a mock rape trial in which he varied the racial composition of the defendant and victim. After reading the transcript, subjects were asked to rate the guilt of the defendant and suggest a sentence. Results of this study also indicated that black defendants are assigned significantly lengthier sentences than white defendants, especially if the victim is portrayed as white.

The results of the preceding laboratory studies have been employed by a number of individuals to suggest that white jurors cannot overcome their biases toward black defendants when deliberating on a case.\textsuperscript{29} A number of limitations regarding these studies, however, must be considered before a final appraisal of the research can be made. These limitations will be discussed in section III.

\textbf{B. Archival Research}

In addition to the preceding laboratory studies, there have been a number of archival analyses completed which also suggest that black defendants are more likely to be convicted and receive harsher sentences than white defendants in the criminal system.\textsuperscript{30}

\begin{table}
\centering
\begin{tabular}{lcc}
\hline
\textbf{DEFENDANT} & \textbf{WHITE} & \textbf{BLACK} \\
\hline
\textbf{WHITE} & 13.8a (60) & 17.2b (62) \\
\textbf{BLACK} & 13.6a (60) & 14.5ab (62) \\
\hline
\end{tabular}
\caption{Mean Ratings of Culpability (White Subjects) Race of Victim X Race of Defendant}
\end{table}

Note: Mean values with different superscripts differ from each other at the .05 level or better by the Duncan Multiple-Range test. Numbers in parentheses indicate the number of subjects in each experimental group.

\textsuperscript{28} Feild, supra note 10.
\textsuperscript{29} Weiten \& Diamond, supra note 15, at 72.
In an early study, Johnson examined the data on 660 convicted capital offenders sentenced to death row in North Carolina in order to determine whether there were any selective factors operating regarding sentencing. His data indicated a large disparity between the number of black and white offenders who were executed in North Carolina between 1909 and 1954. Based on these results, Johnson concluded that "capital punishment operates differentially on the basis of the type of crime and type of offender."

In a similar study, Partington examined the incidence of death penalty sentences for rape cases in Virginia from 1908-1963. Partington found that,

"[t]here is, however, at least the permissible inference, that rape by a Negro man is a crime in Virginia which has been treated as meriting an extreme punishment. It is also evident from the facts that Virginia judges and juries have been less reluctant to impose the death penalty upon Negro defendants than white."

In support of his argument, Partington pointed to the fact that between 1908 and 1963 all of the fifty four men executed for rape or attempted rape in Virginia have been black. In addition to this, Partington also noted that there was a racial disparity among the rape cases decided by the Court of Appeals during this period. Finally, Partington suggested that an analysis of sentencing patterns for rape cases in Virginia between 1908 and 1963 indicated that black defendants were disproportionally sentenced to longer prison sentences. According to Partington, during this period, "Negroes committed 58% of the rapes and received 62% of the sentences over 20 years; 58% of the sentences over 10 and under 20 years; and 56% of the sentences under 10 years. Whites committed 42% of the rapes, and received 38% of the sentences over 20 years; 42% of the sentences over 10 and under 20 years; and 44% of the sentences under 10 years."

Racial differences in sentencing were also alluded to by Gerard and

32. Id. at 169. According to Johnson, data regarding the percentage of capital offenders committed to death row in North Carolina who were executed indicate the following racial disparities: 62% of blacks committed to death row for murder were executed, while only 43.8% of whites committed to death row for murder were executed. A similar result is obtained regarding rape (56.4% of blacks executed and 42.9% of whites executed) and burglary (26.3% of blacks executed and 0% of whites executed).
33. Id. at 169.
34. Partington, supra note 30, at 62.
35. Id. at 43.
36. "In sum, the Virginia Supreme Court of Appeals has written opinions in seventy-two rape cases during the period with which this article is concerned. Twenty-three of the cases involved Negroes: seventeen were affirmed, six reversed. Forty-nine of the cases involved white men: eighteen were affirmed, thirty-one reversed." Id. at 49.
37. Id. at 53-54.
Terry in an analysis of data collected in Missouri for the American Bar Foundation's 1962 *Nationwide Survey of the Defense of Indigents Accused of Crime*. These authors compared jury verdicts as well as sentencing decisions for white and black defendants across a number of criminal offenses and concluded that, "the data clearly showed that Negroes were treated less favorably than whites. In support of their conclusion, these authors referred to the fact that an examination of 19 cases tried by juries illustrated that 33% of the white defendants were found guilty, while 77% of the black defendants were found guilty. In addition, the data suggested that blacks were more likely to receive prison sentences (as opposed to fines or probation) than whites for offenses such as assault, auto theft, forgery, larceny, and robbery.

Although it has been argued that the above studies are irrelevant for any modern analysis of racial disparities because they only reflect data prior to the early 1960's, Zeisel suggests that the past trends continue and can be illustrated. In defense of his argument Zeisel compared the statistical evidence presented in two capital cases, *Maxwell v. Bishop* and *Spinkellink v. Wainwright*, in order to illustrate that racial discrimination in the imposition of the death penalty has not decreased since *Furman v. Georgia*. According to Zeisel, a re-

39. *Id.* at 436.
40. *Id.* at 430.
41. *Id.* at 432.
42. We see here the beginning of what has since become a pattern — government officials admitting, after it no longer matters legally, that discrimination has affected capital sentencing and executions, but professing that such discrimination is all a matter of the past and that the current data are too scanty to support conclusions of continuing racial discrimination. Zeisel, *supra* note 30, at 458.
43. 398 F.2d 138 (8th Cir. 1968), vacated on other grounds, 398 U.S. 262 (1970)(petitioner offered data to the court suggesting that convicted black defendants were more likely to receive the death penalty for raping a white victim than convicted white defendants).
44. 578 F.2d 582 (5th Cir. 1978), cert. denied, 440 U.S. 976 (1979)(petitioner presented data suggesting that an analysis of the ratio of current offenders on death row in Florida to the number of Florida arrests for murder during a felony illustrates a distinct racial disparity. Specifically, it was suggested that 47% of the black defendants arrested for felony murder of a white victim were sentenced to death. In comparison, only 24% of white victims arrested for felony murder of a white victim were sentenced to death).
45. 408 U.S. 238 (1972). In *Furman*, the Supreme Court held that death sentence statutes in the United States were in violation of the eighth amendment, prohibiting cruel and unusual punishment. Although the rationales of the Justices varied (nine separate opinions were written), a number of the opinions pointed to evidence suggesting that the arbitrary infliction of the death penalty invited racial discrimination. Subsequent to this decision, a number of states legislated death penalty statutes in which an attempt was made to reduce the arbitrariness of the death penalty process. Typically, states introduced more specific criteria which
view of the Florida data presented in *Spinellink* indicates that a black defendant is 43% more likely to receive the death penalty, for a murder of a white victim during a felony, than a white defendant.\textsuperscript{46} As such, it is argued that differential sentencing based on race still continues in capital cases.

Evidence regarding the continuation of differential sentencing was also presented by Baldus and his colleagues who employed a large data base and more sophisticated statistical analysis to conduct two studies which examined the archival data from 2,500 cases in Georgia. These studies also illustrated a trend toward gross discrimination in death penalty decisions.\textsuperscript{47} After controlling approximately 270 factors (i.e., holding constant any variables aside from race which might explain the disparity in sentencing), these authors concluded that the race of the defendant and victim played a significant role in determining whether the death sentence was imposed. Specifically, defendants who were found guilty of murdering a white victim received the death penalty in 22% of the cases examined. In comparison, white defendants found guilty of murdering a black victim were sentenced to death in only 1% of the cases.

Bowers and Pierce\textsuperscript{48} also conducted an extensive study of post-\textit{Furman} capital sentencing decisions and found evidence for arbitrary impositions of the death sentence for homicide in Florida, Georgia, Texas, and Ohio.\textsuperscript{49} These results led the authors to conclude that, "[i]n these four states, which accounted for approximately 70 per cent of the nation's death sentences in the first five years after Furman, race of both offender and victim had a tremendous impact on the

\textsuperscript{46} Zeisel, supra note 30, at 461.
\textsuperscript{47} Baldus, Challenge, supra note 14; Baldus, Lessons, supra note 14.
\textsuperscript{48} Bowers & Pierce, supra note 30.
\textsuperscript{49} Probability of Receiving the Death Sentence in Florida, Georgia, Texas, and Ohio for Criminal Homicide, by Race of Offender and Victim (from effective dates of respective post-\textit{Furman} capital statutes through 1977).
chances that a death sentence would be handed down." These racially biased results were found even when the type of homicide involved was held constant.\textsuperscript{51}

Similarly, Wolfgang and Riede\textsuperscript{52} examined 361 rape cases in Georgia between 1945 and 1965 in an attempt to examine the effect of both legal and non-legal factors on decisions to impose the death sentence. Employing a sophisticated analysis technique to examine the data, these authors, like the above studies, found that the black defendant/white victim racial combination was more likely to evoke a death sentence than any other defendant/victim racial combination.

Finally, Zimring, Eigen, and O'Malley\textsuperscript{53} conducted an analysis of 204 homicide cases in Philadelphia since 1970 and concluded that, "[b]lack defendants who kill white victims receive the life or death sentence more than twice as often as black felony suspects who kill black victims."\textsuperscript{54}

The results of the archival studies, like those of the laboratory studies, have been used to reinforce the suggestion that juries are in-

\begin{center}
\begin{tabular}{lll}
\hline
Offender/Victim Racial Combinations & (1) Estimated Number of Offenders & (2) Persons Sentenced to Death & (3) Overall Probability of Death Sentence \\
\hline
\textbf{Florida} & & \\
Black kills White & 240 & 53 & .221 \\
White kills White & 1768 & 82 & .046 \\
Black kills Black & 1922 & 12 & .006 \\
White kills Black & 80 & 0 & .000 \\
\textbf{Georgia} & & \\
Black kills White & 258 & 43 & .167 \\
White kills White & 1006 & 42 & .042 \\
Black kills Black & 2458 & 12 & .005 \\
White kills Black & 71 & 2 & .028 \\
\textbf{Texas} & & \\
Black kills White & 344 & 30 & .087 \\
White kills White & 3616 & 56 & .015 \\
Black kills Black & 2597 & 2 & .001 \\
White kills Black & 143 & 1 & .007 \\
\textbf{Ohio} & & \\
Black kills White & 173 & 44 & .254 \\
White kills White & 803 & 57 & .046 \\
Black kills Black & 1170 & 20 & .017 \\
White kills Black & 47 & 0 & .000 \\
\hline
\end{tabular}
\end{center}

\textit{Id.} at 594.
50. \textit{Id.} at 596-97.
51. \textit{Id.} at 598-99.
52. Wolfgang & Reidel, \textit{supra} note 30.
54. \textit{Id.} at 232.
JURY RACISM

HERENTLY prejudiced toward black defendants and, as such, act in a discriminatory fashion when deliberating upon the guilt and sentencing of the non-white defendant. However, as with the laboratory studies, the archival evidence of jury discrimination must also be scrutinized before such a conclusion can be drawn.

III. EVALUATION OF EVIDENCE

Although the large amount of laboratory and archival evidence cited above suggests that there is indeed a prejudicial (i.e., attitudinal) element to jury decision-making, there are a number of methodological and conceptual limitations which beg consideration before one may argue that actual jurors are engaging in racist behavior. Indeed, we should follow the advice of Justice Frankfurter who suggested that "we have to be constantly on our guard lest psychology be more unequivocal in her wisdom when she speaks to lawyers than when she speaks to psychologists." Psychologists, therefore, should be no less critical when analyzing the outcome of psycholegal research than they are in analyzing research in other areas of psychology. Thus, the remainder of this comment will examine the empirical evidence regarding jury racism, according to a number of psychological arguments, to illustrate that evidence of discriminatory racist behavior in jury decision-making is weaker than has been previously suggested.

A. Limitations of Archival Research

As noted above, evidence regarding jury racism can be separated into two distinct domains: laboratory research and archival research. The laboratory research on jury decision-making is aimed at the individual level of analysis (i.e., the decisions of individual mock jurors), while the archival research is aimed at the institutional level of analysis (i.e., discrimination in the criminal justice system generally). Therefore, although both of these domains may suggest that blacks and whites are treated differently in the legal system, the results from one domain cannot be used to support the results from the other domain. In other words, even though a number of authors cite archival data as evidence in support of their laboratory research, the comparison is not valid because the archival data suggest only that the differential treatment is occurring "somewhere in the system," it does not pinpoint the jury as the cause. Baldus himself suggests that there are

55. See, e.g., Ugwuegbu, supra note 18, at 134.
56. F. FRANKFURTER, LAW AND POLITICS 297 (1939).
58. See, e.g., Ugwuegbu, supra note 18, at 134. "Many of these studies were surveys or archival investigations. However, these data generally indicate that the race of the defendant and the victim may influence juror judgement [sic]." Id.
at least four distinct decision points in the judicial system during which discrimination can play a role in determining whether the death penalty is imposed.\(^5\)

A second limitation of archival research lies in the fact that a number of studies have produced conflicting evidence regarding the contention that discriminatory sentencing practices exist.\(^6\) Kleck, for example, reviewed fifty-seven published studies on racial discrimination in criminal sentencing for capital and non-capital offenses. Results of this analysis indicated that only fifteen of the studies presented any strong evidence of racial bias. An additional sixteen studies presented "mixed" findings, and the remaining twenty-six studies found no evidence at all of racially motivated sentencing disparity.\(^61\)

Similarly, a recent report by Klein, Petersilia, and Turner suggests that knowledge of a defendant's race does not increase the accuracy of disposition predictions.\(^62\) Specifically, these authors analyzed 11,553 California offenders convicted of assault, robbery, burglary, theft, forgery, or drug crimes, and developed a model which allowed them to predict the disposition of the defendant (i.e., probation or prison sentence) with approximately 80% accuracy. Knowledge regarding the race of the defendant did not enhance the predictive power of the model. That is, the authors suggest that knowledge of factors such as the defendant's prior record (e.g., previous adult and juvenile convictions), specific elements of the crime (e.g., use of a weapon, vulnerability of the victim), and demographics (e.g., education, mental or alcohol-related problems) are much more useful in predicting a resultant sentence than knowledge about race.

Therefore, although a number of studies argue that discrimination does exist within the judicial system, the evidence supporting this claim is unclear at best. What is clear is that, even when sentencing

---

59. Baldus, *Challenge*, supra, note 14, at 139-40. Baldus and his colleagues suggest that most jurisdictions have at least four "decision points" at which a responsible authority makes a judgment which "will either keep the defendant eligible for a death sentence or effectively insulate him from that possibility." The decision points are as follows:
   1. Decision by prosecutor (or grand jury) to charge defendant with capital murder or a lesser offense.
   2. Possibility of plea bargaining in which prosecutor may reduce the charge in exchange for a guilty plea.
   3. Jury decides on guilt of the defendant.
   4. Jury (or Judge) decides on sentence.

disparities are documented, there is no conclusive archival evidence which suggests that juries are specifically responsible for these disparities. It seems, therefore, that archival research in and of itself does not present any evidence of jury discrimination, although it does illustrate at least some support for the contention that racism does exist "somewhere" in the justice system.

B. Methodological Limitations of Laboratory Research

Perhaps the most widely held criticism of laboratory research on jury racism, and jury decision-making in general, revolves around the external validity limitations which are inherent in these studies.\(^{63}\) This criticism is generally based upon the methodological inadequacies of the research that arise because most studies do not attempt to replicate all aspects of an actual trial when investigating jury decision-making issues.

The major experimental procedures which are cited by authors as detracting from the external validity of results include: 1) the primary use of college students as subjects;\(^{64}\) 2) concentration on aspects of the juror as opposed to the jury;\(^{65}\) 3) the realism of the situation;\(^{66}\) 4) the use of measures which do not reflect actual trials (i.e., using scales of guilt, or sentencing as a measure of prejudice);\(^{67}\) and, 5) the lack of some important aspect of the trial situation (i.e., instructions, opening

63. External validity can be defined as the degree to which laboratory findings can be extended or generalized to the real world. D. COOK & D. CAMPBELL, QUASI-EXPERIMENTATION: DESIGN AND ANALYSIS ISSUES FOR FIELD SETTINGS 39 (1979).


65. It has been argued that concentration on aspects of individual juror decision-making tells us very little about the decisions which a jury, as a whole, would make. Davis & Bray, supra note 64, at 341; Loh, supra note 5, at 344.

66. These researchers suggest that all subjects are aware that they are in an experiment, no matter how realistic the situation, and as such may not act the way they would if faced with a real judicial decision-making task. Bermant, McGuire, McKinley & Salo, supra note 64, at 228; Davis, Bray & Holt, supra note 64, at 350; Gerbasi, Zuckerman & Reis, supra note 64, at 342-43; Loh, supra note 5, at 344; Weiten & Diamond, supra note 15, at 81-83.

67. It has been suggested that the use of dependent measures, such as rating guilt on a seven-point scale, are not reflective of legal procedure (i.e., guilty or not guilty) and therefore provide little useful information. Bray & Kerr, Use of the Simulation Method in the Study of Jury Behavior, 3 LAW & HUM. BEHAV. 107, 111-12 (1979); Loh, supra note 5, at 344; Weiten & Diamond, supra note 15, at 79.
Lipton, however, attempted to overcome many of the above criticisms by trying to replicate an actual trial situation while varying the race of the defendant. Subjects were led to believe that they were participating in an actual trial to determine the guilt or innocence of a student accused of cheating. Accordingly, the subjects were told that they were selected to be part of an “Intercampus Grievance Committee” (ICGC). As part of their ICGC responsibilities, the subjects were asked to read a number of verbatim transcripts of student disciplinary hearings and suggest an appropriate disposition which would be adhered to by the university. By varying the race of the accused student, Lipton found that Hispanics were treated more harshly than whites with regard to disposition. Although Lipton did find a race effect for juror decisions, his results are somewhat suspect because of the lack of jury instructions.

The issue of external validity is also alluded to by Bermant, McGuire, McKinley and Salo who suggest that results from laboratory research on juries may not be a useful indicator of actual jury functioning because these studies concentrate on functional verisimilitude as opposed to structural verisimilitude. These authors suggest that researchers are attaching so much importance to the output, or function, of subjects (i.e., guilt ratings, sentencing decisions, etc.) that they are neglecting to attend to the importance of the input variables, or process, of the phenomenon being studied (i.e., the elements of an actual trial such as instructions and deliberations) resulting in non-generalizable findings.

Given the above methodological limitations, it appears that although many studies on jury racism may illustrate internally valid

---

68. Some researchers argue that the results of studies which do not incorporate all aspects of a trial, such as jury instructions and deliberations, provide us with little information about how a “real” jury would behave. Bermant, McGuire, McKinley & Salo, supra note 64, at 230; Davis, Bray & Holt, supra note 64, at 341; Gerbasi, Zuckerman & Reis, supra note 64, at 342; Loh, supra note 5, at 341; Pfeifer & Ogloff, supra note 19, at 10; Weiten & Diamond, supra note 15, at 77-78.

69. Lipton, Racism in the Jury Box: The Hispanic Defendant, 5 Hispanic J. Behav. Sci. 275 (1983).

70. For a review on the importance of jury instructions in jury decision-making studies on racism, see Pfeifer & Ogloff, supra note 19.

71. Bermant, McGuire, McKinley & Salo, supra note 64.

72. In terms of a computer analogy, functional verisimilitude can be described as the throughput and output elements of data processing. In jury research, therefore, functional verisimilitude would revolve around the concern over whether the results obtained in the study are reflective of actual jury behavior.

73. Structural verisimilitude is based on a concern for input variables. In jury research, this concern would be reflected by the importance of testing subjects in situations which exactly mimic a real trial.
findings of prejudice, their techniques of investigation limit their externalization to the legal forum.

C. Predictive Limitations of Laboratory Research

As stated earlier, Harding, Proshansky, Kutner and Chein suggest that racism can be divided into two distinct forms: an attitudinal component labeled prejudice and a behavioral component labeled discrimination.\(^7\) Given this division, it has been suggested that the procedures being used in laboratory studies on jury racism are primarily tapping into the concept of prejudice as opposed to discrimination.\(^8\) Specifically, subjects in these laboratory studies are being assessed on their attitudes, or prejudices, as opposed to their behaviors. Indeed, Nickerson, Mayo, and Smith refer to the stimulus being employed in these studies as "veiled attitude questionnaires."\(^7\) Therefore, the fact that these studies suggest differential treatment of defendants based on race simply reflects the subjects' differential attitudes toward racial minorities.

Given these attitudinal findings, it is a large and perilous theoretical leap to assume that there are behavioral consequences which will naturally follow.\(^7\) One cannot simply assume that discriminatory behaviors can be reliably predicted in cases where prejudicial attitudes have been identified,\(^8\) even though a number of authors have implicitly or explicitly made these assumptions.\(^9\) The reliability of the attitude/behavior prediction paradigm has led a number of authors to suggest that there is in fact very little relationship between "attitudinal predictors" and subsequent behavior.\(^8\) In fact, one recent study has found that attitudes become less reliable as predictors of behavior if the behavior has "significant and far-reaching effects."\(^8\) In this study, the subjects' attitudes toward capital punishment were measured. Following this, the subjects were asked to make a decision in a mock jury trial involving capital punishment. Before giving their verdict, subjects were reminded to either respond according to their atti-
tudes (high attitude relevance) or according to the facts of the case (low attitude relevance). In addition, subjects were informed that their decisions would either be used to influence the real jury in this case (high behavior relevance) or would not influence the real jury at all (low behavior relevance). The results indicated that the lowest correlation (i.e., degree to which the attitude and behavior variables were related) was discovered in the high attitude relevance/high behavior relevance condition. In contrast, when subjects were led to believe that their attitudes were relevant but would not play a significant role in the trial outcome, the correlation between the two variables increased. Therefore, the argument that these attitudinally-oriented laboratory studies are solid predictors of actual juror behavior is not a strong one.

Interestingly enough, it has been suggested that the law is more likely to make a distinction between “bias and action” (i.e., a distinction between attitudes and behaviors) in the realm of jury decision-making than some psycholegal researchers who are more likely to equate the two, even though the attitude/behavior question is far from resolved in psychology.\(^8\) For example, in *McCain v. Employment Division*,\(^8\) the Oregon Court of Appeals held that there was a distinct difference between an employer's sexist attitudes and any subsequent behavior which that attitude may produce. This opinion seems to provide a clear indication that at least one court has made an attitude/behavior distinction in defining discrimination. It seems that in the realm of attitude-behavior predictability, some cases wholeheartedly

82. Loh, *supra* note 5, at 335.

Otherwise stated, psychological writings equate potential bias (general fallibility) with actual bias (mistaken verdict). The law distinguishes these two elements. . . . The law's reluctance to assume categorically a one-to-one correspondence between the potential unreliability (which is conceded) and its actual impact in a particular trial (which must be proven) is one reason why the courts have not rushed to embrace expert psychological testimony.

*Id.*

83. 17 Or. App. 442, 445, 522 P.2d 1208, 1210 (1974). The court drew a clear distinction between an individual's attitude and behavior in terms of defining sexism. According to the opinion in this case, even though “[d]iscrimination on the basis of sex is an unlawful employment practice. . . . This does not mean, however, that an employer's 'sexist' attitude, by itself, is an unlawful employment practice or such other cause as would constitute 'good cause' for a female employe [sic] to quit. 'Good cause' would exist only if this 'sexist' attitude produced some actual discrimination, undue harassment, or other grievous cause of reasonable foundation, evidence of which must appear in the record” [citations omitted]. Although this case deals with sexism, as opposed to racism, one can argue that the same reasoning would apply since the opinion was based on the Oregon Fair Employment Practices Act which deals with unlawful employment activities due to race, religion, color and sex. See Or. Rev. Stat. § 659.030 (1987).
adhere to Justice Frankfurter’s advice by rejecting psycholegal research data that has yet to be fully accepted in the field of psychology.

D. Correspondence Limitations of Laboratory Research

Even were one to agree that attitudes are a direct predictor of behavior, there is evidence suggesting that the relationship between the attitudes being tapped and the subsequent behavioral predictions will be very low due to the lack of correspondence between key elements. Ajzen and Fishbein suggest that attitudes and behaviors will have a higher relationship if the two correspond in terms of four elements: target, action, context and time. In other words, the probability that one can make a reliable behavioral prediction, based on knowledge of an individual’s attitude, increases in direct proportion to the elemental correspondence between the two. For example, the probability of correctly predicting an individual’s voting behavior is increased by probing his or her attitudes toward a specific candidate (target), about specific issues (context), during a specific election (time). Likewise, behavioral predictability would decrease if the attitudes were probed about a generic candidate, regarding general issues, between election years.

Given this concept of element correspondence, the contention that laboratory studies on racism using mock jurors are evidence of actual jury discrimination is also weakened because these studies lack correspondence with actual trials. The specific details of each study on jury racism with regard to target (i.e., who the defendant is) and action (i.e., facts of the case), as well as context and time, do not correspond to the elements of an actual trial and, as such, are only valuable as predictors of behavior within the confines of the study. This argument is especially important since McCleskey v. Kemp established the legal standard for proving discrimination in a capital case as being dependent on the presentation of evidence which specifically establishes that the defendant was discriminated against.

It seems, therefore, that the law requires specific evidence of discrimination in capital cases while psychology can only provide general trends. While these trends are interesting psychological findings, their utility as specific indicators of discriminatory behavior by jurors in a specific case is questionable.

E. Additional Laboratory Research

Finally, there is laboratory evidence suggesting that mock jurors

---

84. Ajzen & Fishbein, supra note 80.
86. Loh, supra note 5, at 336.
may not be as inherently biased as previously believed. Pfeifer and Ogloff, for example, found that mock jurors in a rape trial tend to be less discriminatory when they are evaluated according to a stricter legal standard. These authors asked subjects to read a transcript of a rape trial in which the race of the defendant and victim were varied. After reading the transcript, half of the subjects were asked to rate the guilt of the defendant on a dichotomous scale (guilty/not guilty) while the other half were asked to make their judgment on a seven-point bi-polar scale (from not guilty to extremely guilty). In addition, half of the subjects were also given jury instructions which specified the elements of the crime and noted that in order to find the defendant guilty, the State had to prove each element beyond a reasonable doubt. Results replicated earlier studies in which subjects overwhelmingly rated the defendant as guiltier, regardless of scale employed, when he was portrayed as being black as opposed to white. These differences disappeared, however, when subjects were given jury instructions. This study suggests that the addition of basic trial procedures (i.e., instructions to the jury) has a vast effect on findings of prejudice in psycholegal research on jury decision-making. Thus, this research forces one to question the validity of previous laboratory research which did not employ jury instructions.

Stephan and Stephan also found that the addition of jury instructions tended to dissipate prejudicial findings in their study on juror reactions to non-English speaking defendants. These authors found that English subjects tended to rate Hispanic and Thai defendants as guiltier if their testimony was presented through an interpreter rather than in English. The discrepancy in ratings, however, disappeared when subjects were instructed by the judge to ignore the fact that the testimony was translated. According to Pfeifer and Ogloff, this dissipation of racism in jury studies when instructions are added can be


88. Pfeifer & Ogloff, supra note 19.
89. Stephan & Stephan, supra note 87.
90. Pfeifer & Ogloff, supra note 19.
explained by modern racism theory. They suggest that when faced with an ambiguous case involving a black defendant, mock jurors will tend to fall back on their prejudicial attitudes unless instructions are issued which disperse the ambiguity of the situation forcing mock jurors to focus on the specific facts of the case as opposed to the race of the defendant.

Similarly, Bermant, McGuire, McKinley and Salo suggest that subjects tend to "fill in the gaps" with their own images of persons and events in a trial when they are not presented with all components of a trial (i.e., opening statements, instructions, deliberations, etc.). These studies certainly substantiate earlier jury research which suggested that jurors are competent fact-finders who take their job seriously and generally adhere to the "rules of law." 

Finally, there are studies which suggest that a number of variables play a more important role in juror decision-making than race. These studies, for the most part, suggest that variables such as defendant attractiveness, socio-economic status, and attitude similarity play a larger role in jurors decisions than the race of the defendant.

It seems, therefore, that like archival research, laboratory research on jury racism does not provide any positive proof of discriminatory actions by jurors. The methodological limitations of the research, the lack of strong external validity, the predictive and correspondence limitations, and the large number of studies that refute findings of jury racism, suggest that the previous findings in this area are far from conclusive.

IV. CONCLUSION

Although the research presented above is limited in its external validity, it seems to suggest that juries are not particularly discriminatory towards non-white defendants, even though there is some evidence of juror prejudice. Accordingly, it seems that the evidence of jury racism, thus far, is less reflective of a finding of discrimination than a misinterpretation of data resulting in a discriminatory finding. That is, the findings themselves have been incorrectly applied to support a concept which is far from established. If anything, jury decision-making research on racism supports only the contention that

91. Bermant, McGuire, McKinley, & Salo, supra note 64, at 232.
mock jurors may hold prejudicial attitudes toward non-white defendants when they are not given jury instructions. There is little evidence, however, to suggest that real jurors are adhering to these attitudes when they determine the guilt or innocence of defendants.

The information discussed in this comment, however, is not meant to imply that racism does not exist, or is an unimportant question in the judicial system. Rather, this comment investigates the extent to which current laboratory and archival data on jury racism may be used to support the contention that juries engage in biased behaviors when asked to assess the guilt of non-white defendants. It may be that future research will add to our knowledge in this area. At this time, however, it seems that the evidence regarding the effects of racism on jury verdicts does not warrant any need for major changes in the jury system.

Jeffrey E. Pfeifer