INFORMATION TO USERS

This reproduction was made from a copy of a document sent to us for microfilming. While the most advanced technology has been used to photograph and reproduce this document, the quality of the reproduction is heavily dependent upon the quality of the material submitted.

The following explanation of techniques is provided to help clarify markings or notations which may appear on this reproduction.

1. The sign or "target" for pages apparently lacking from the document photographed is "Missing Page(s)". If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting through an image and duplicating adjacent pages to assure complete continuity.

2. When an image on the film is obliterated with a round black mark, it is an indication of either blurred copy because of movement during exposure, duplicate copy, or copyrighted materials that should not have been filmed. For blurred pages, a good image of the page can be found in the adjacent frame. If copyrighted materials were deleted, a target note will appear listing the pages in the adjacent frame.

3. When a map, drawing or chart, etc., is part of the material being photographed, a definite method of "sectioning" the material has been followed. It is customary to begin filming at the upper left hand corner of a large sheet and to continue from left to right in equal sections with small overlaps. If necessary, sectioning is continued again—beginning below the first row and continuing on until complete.

4. For illustrations that cannot be satisfactorily reproduced by xerographic means, photographic prints can be purchased at additional cost and inserted into your xerographic copy. These prints are available upon request from the Dissertations Customer Services Department.

5. Some pages in any document may have indistinct print. In all cases the best available copy has been filmed.

University Microfilms International
300 N. Zeeb Road
Ann Arbor, MI 48106

Baker, Eli

CONVICTION PRONENESS AS A PREDICTOR OF SWORN JUROR DECISIONS

City University of New York PH.D. 1984

PLEASE NOTE:

In all cases this material has been filmed in the best possible way from the available copy. Problems encountered with this document have been identified here with a check mark _√_.

1. Glossy photographs or pages _____
2. Colored illustrations, paper or print _____
3. Photographs with dark background _____
4. Illustrations are poor copy _____
5. Pages with black marks, not original copy _____
6. Print shows through as there is text on both sides of page _____
7. Indistinct, broken or small print on several pages _✓_
8. Print exceeds margin requirements _____
9. Tightly bound copy with print lost in spine _____
10. Computer printout pages with indistinct print _____
11. Page(s) _____ lacking when material received, and not available from school or author.
12. Page(s) _____ seem to be missing in numbering only as text follows.
13. Two pages numbered _____. Text follows.
14. Curling and wrinkled pages _____
15. Other _____

University
Microfilms
International

CONVICTION PRONENESS AS A PREDICTOR OF SWORN JUROR DECISIONS

by

ELI BAKER

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

1984

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

4/6/84
date

Robert Buckhout
Chairman of Examining Committee

April 11, 1984
date

Herbert D. Saltzstein
Executive Officer

Robert Buckhout

Morton Bard

Alan Gross

Supervisory Committee

The City University of New York

Abstract

CONVICTION PRONENESS AS A PREDICTOR OF SWORN JUROR DECISIONS

by

Eli Baker

Adviser: Professor Robert Buckhout

Scientific jury selection is the application of social science knowledge and technology to the courtroom trial. Seemingly successful, but never empirically tested, there are questions concerning the efficacy of this technique. The substantive content of jury research has been developed primarily through the use of the jury paradigm under experimental conditions. Poor replicas of actual trial conditions have generated conflicting and questionable conclusions. Moreover, sworn jurors are not randomly selected from the general population, and research conducted with samples of the general population produce results not applicable to actual sworn jurors who have been chosen through ordinary courtroom routines.

This study analyzed data from questionnaires returned by 365 sworn jurors who had actually sat on criminal trials immediately prior to their responses at the Federal Court, Eastern District of New York. These

jurors were randomly split into exploratory and validation samples. A conviction score was generated as a function of the respondent jurors' reported pre-deliberation vote and the vote breakdown of the juries of which they were members. A conviction proneness proxy was created from the attitude questions. Stepwise regressions and other exploratory procedures were used to derive a prediction equation of conviction proneness.

This study demonstrated that although a stable conviction proneness proxy can be derived from demographic data and attitude data not related to a specific case, it cannot reliably predict the conviction behavior of real jurors across samples, $r = .09$, one-tailed $p>.10$. Furthermore, the conviction behavior of the real jurors in the validation sample could not be predicted from the equation developed in the exploratory sample despite the inclusion of actual trial related variables, $r = .09$, one-tailed $p>.10$. Finally, a fundamental assumption of scientific jury selection that there is a predictable relationship between juror conviction proneness and juror conviction behavior remains uncorroborated by a non-significant correlation for the two measures used in this study, $r=.06$, $p>.10$.

The failure to predict the behavior of sworn jurors using scientific jury selection techniques suggests that these techniques do not help attorneys predict juror verdicts in ordinary criminal trials.

Acknowledgements

I'd like to thank Professor Robert Buckhout, my mentor, who has encouraged, prodded, and helped me in so many different ways.

To the members of my committee, Professors Morton Bard and Alan Gross, my gratitude for their patience, availability, encouragement, and help.

To the students and faculty of the Social/Personality program at the Graduate Center, CUNY, my sincere appreciation; working with them has enriched my life.

Most helpful have been the staff and students who have worked at the Center for Responsive Psychology at Brooklyn College these past years. Ungrudging of their time they have had a major impact on this study.

My special thanks to my friend and colleague, Larry Krasnoff. He has done yeoman work helping me to organize my sometimes scattered thoughts.

I want to thank my children, Michael, Leah, and Joey who each have contributed in some way to this work; and who have cheerfully done with less so that daddy could get an education.

And, mostly, my gratitude to my wife and lifelong partner, Rhoda Baker. It was she who 'dared' me to go to college, encouraged me to leave my job and attend graduate school, worked at a job she disliked while I enjoyed the academic life, and lived on a diminished income

while her friends and acquaintances reaped the benefits which accrue after years of hard work. To her I owe everything and dedicate this thesis.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

Introduction

The jury system is being challenged. Research has shown that non-evidential factors influence jury verdicts (Gerbasi, Zuckerman, & Reis, 1977), and the publicized successes of 'scientific jury selection' have reinforced this view of the manipulable juror. Scientific jury selection as recently developed and practiced involves conducting surveys and interviews of the potential juror population (registered voters in most cases) to create profiles of favorable and unfavorable jurors. These profiles are then used to modify in-court evaluations of prospective jurors by psychologists and attorneys; and have led to seemingly astonishing reversals of expected verdicts in such media celebrated cases as the Harrisburg seven (Shulman, Shaver, Colman, Emrich, & Christie, 1973), the Angela Davis trial (Sage, 1973), and the Mitchell-Stans trial (Zeisel & Diamond, 1975).

This marriage of law and social science methodology raises important questions: should social scientists evaluate questions of law? should social scientists intervene in actual trials and influence trial outcomes?

The successes of jury selectors led Etzioni to write a widely cited essay in the influential Washington Post deploring the intrusion of science into "that venerable traditional institution of being judged by a jury of one's peers." He argued that

> the impartiality of the jury is threatened because defense attorneys have recently discovered that they can manipulate the composition of the jury by the use of social science techniques, so as to significantly increase the likelihood that the defendants will be acquitted (1974: p. 28).

Scientific jury selection is an expensive and time consuming process. Survey experts, an army of interviewers, data collators, expensive computers, statisticians, and psychologists are needed to collect, interpret, and use the information. Etzioni continued:

> Clearly, the average defendant cannot avail himself of such aid and, therefore, the net effect of the new technique, as is so often the case with new technology, will be to give a leg up to the wealthy or those who command a dedicated following. This is hardly what the founders of the American judicial system had in mind (p. 30).

Etzioni warned the creators and perfectors of this technique that, instead of the democratization of the jury system which they seek, they are refining a two-edged tool which will subsequently be routinely used by the state in its prosecutions, and the wealthy in its defenses. Etzioni took the position that no knowledge is preferable to knowledge which can be used to subvert the criminal justice system (CJS); there is a precarious balance of interests involved here which social science should not disturb.

Investigations go forward, but there is a continuing debate between the jury researchers and the Etzioni viewpoint which is cited and answered in almost every study and review. The jury researchers emphasize the need to critically examine

all aspects of our society in order to ensure the fair distribution of resources (e.g. Van Dyke, 1978); they claim that they are doing systematically what lawyers have always been doing haphazardly (e.g. Berk, 1976; Christie, 1976); that the government has been using secret

sources of information and jury selectors are just trying to redress the imbalance (e.g. Moore, Jr., 1974); and, besides, jury selection doesn't really work (e.g. Saks, 1976a).

While the Etzioni-jury researcher debate may be decided on the basis of community priorities, there are more mundane but quite important scientific problems of the validity and generalizability (Campbell & Stanley, 1963) of jury research which can be investigated empirically. Is the extensive body of empirical knowledge we are now accumulating reliably accurate in its findings? Does jury selection really work? Jury deliberations are shielded from scrutiny by tradition and law, thereby undermining the research process and leading to questionable re-creations of jury trials for experiments (the mock jury trial) which may lead to incorrect conclusions. Are these results informing us about real life conditions or are they only theoretical estimates of general behavioral processes generated through the use of a convenient jury paradigm?

These are serious questions for jury researchers. Spurious results may lead to dubious reformations of traditional procedures to the detriment of society. The right to a jury trial is guaranteed, but its form is are mutable. It would be unfortunate if these legal procedures were modified in the light of poorly executed science and questionable results. Konecni and Ebbesen, two social scientists who have studied legal questions, feel that

> erroneous information obtained by scientific methods (and therefore having an aura of truth) is more harmful than no information at all, especially when issues as sensitive as legal ones are being dealt with, and people's futures are quite literally at stake (1979; p. 68).

Several Supreme Court decisions have cited and used social science studies, and there probably will be a continued and even closer relationship with legal opinions informed by empirical investigations. As a result of this interactive relationship between legal proscriptions and empirical investigations a higher proportion of females are now serving as jurors, more blacks have been empaneled, and the median age of the jurors is steadily approaching the population median in most jurisdictions; the description of the average jury as white, middle class, and middle aged is becoming less and less accurate (Van Dyke, 1977). As research findings become more acceptable and influential in the courts, the burden of responsibility for these findings will become correspondingly greater. The easy answers to complex questions which have proliferated in the recent past (see Ballew v Georgia, 1978 for a body of spurious research on an important legal question) should be tested under real conditions before their substance be considered 'scientific.'

The present study analyzes data from questionnaires administered to real jurors who have actually sat on criminal trials immediately prior to their responses at the Federal Court, Eastern District of New York, located in Brooklyn, New York. Our primary purpose was to evaluate jurors on a dimension of favorability to the defense using the usual

statistical techniques available to the jury selectors to answer the question: Can social and psychological survey information predict juror votes in mundane criminal cases? Our information was much more complete than that which is ordinarily available to jury selectors since we have extensive data on actual jurors rather than just a sampling of the population from which they are not randomly chosen (because of conditions explained below). We then compared the predicted votes with those actually delivered, weighted by the ambiguity of the cases. Secondly, we tested a scientific jury selection model which utilizes all the information contained in our questionnaire. And, finally, we tested the assumptions implied by the use of scientific jury selection.

The Jury Trial

In jury research the emphasis is on the extraneous influences which may affect the juror; evidence is rarely mentioned. To appreciate the problems involved in doing meaningful jury research, one has to know the actual operation of the jury trial because it has a number of features which experimental work often ignores.

History, Tradition, and the Supreme Court

The jury trial has a long and somewhat hallowed tradition. Born in England as a check to official persecution and exported along with the colonists and English common law, it has flowered in the United States

with about 150,000 criminal jury trials tried in the U.S. yearly (Ellison & Buckhout, 1981).

The landmark case in English Common Law is commemorated by a plaque hanging in Old Bailey Court, London:

> Near this site William Penn and William Mead were tried in 1670 for preaching to an unlawful assembly in Grace Church Street. This tablet commemorates the courage and endurance of the Jury Thos Vere, Edward Bushness and ten others who refused to give a Verdict against them, although locked up without food for two nights and were fined for their final Verdict of Not Guilty.
>
> The case of these Jurymen was reviewed on a Writ of Habeas Corpus and Chief Justice Vaughan delivered the opinion of the Court which established 'The Right of Juries' to give their Verdict according to their Convictions (Moore, 1973).

And in our public schools, jury nullification (the jury does not have to follow the judge's interpretation of the law) is symbolically enshrined in the Zenger Trial: a jury overtly disregarded the law by declaring J. P. Zenger 'not guilty' even though he was technically guilty of publishing libelous statements (the jury believed the statements true, and saw no harm in declaring the truth).

Jury forms and procedures are legislated by state bodies and the U. S. Congress. The legislative intent is carried out by the administrative arms of the relevant courts, but the intent and actual operations are subject to constitutional constraints as interpreted by the Supreme Court. While there are wide variations in jury practice and usage (For instance, Kalven & Zeisel (1966) report that 74% of defendants waive

jury trial in Connecticut, while none waive in Montana.), it is the Supreme Court which decides what is considered within the bounds of proper jury procedure and its decisions frequently outline our areas of research. Supreme Court decisions are often based on assumptions on how jurors will act in particular situations; these assumptions can be accurate to differing degrees (Buckhout, 1978; Saks, 1974).

For example, Justice White in Apodaca v Oregon (1972) based the majority decision that unanimity is not required in felony cases upon the proposition that majority opinion jurors will listen to those in the minority and pay due considerations to these opinions. In his dissent, Justice Douglass questioned the validity of that proposition and appealed to the reality of juror prejudices and time pressures. Here is obviously a testable proposition. And, in fact, there are a number of studies that subsequently addressed this question with the inescapable conclusion that Justice Douglass' judicial intuition was more realistic: When sufficient votes were garnered to render a verdict, the minority opinions were substantially ignored (e.g. Buckhout et al, 1977; Davis, Kerr, Atkin, Holt, & Mark, 1975; Foss, 1981). The right to jury in criminal trials is guaranteed by the sixth amendment to the constitution; however, the rules under which these juries operate are negotiable and derived from common law and ongoing experience. There are, in the United States, many different judicial systems, each having its own set of procedures, each following the dictates of its own lawmakers, each

appealing to its own version of common-sense knowledge, and each conforming to the broad outlines set by the grand arbiter, the U.S. Supreme Court.

Present day procedures result from a series of slow and traditional reforms. Shapiro (1969) makes the argument that judicial decisions are not revolutionary but evolutionary. Not all alternatives are considered, only those most compatible with the status quo. Failure of the status quo leads to a search for alternatives. The first satisfactory alternative is usually chosen. Judges, theoretically, have freedom of choice (as do wielders of power in any organization); but in the real world the choices are narrowly confined. The Supreme Court never required random selection of jurors until congress passed the Jury Selection and Service act of 1968! It wasn't until 1930 (in Patton v U.S.) that the Court decided that a defendant could waive a jury trial in Federal Court. And it was only 16 years ago that Kalven and Zeisel (1966) could write: "It has never been decided whether the Constitution as a matter of due process, requires trial by jury for all criminal cases" (p.15).

The notion that the jury should be representative of the community was first mentioned in Smith v Texas (1940), and expanded in Glaser v U.S. (1942). While the majority held in Swain v Alabama (1965) that "an imperfect system [of selection] is not equivalent to purposeful discrimination" (Justice White, p. 208), it was Alexander v La (1972)

which laid down the "opportunity to discriminate" rule which further institutionalized the random jury selection that allows no such opportunity.

The Supreme Court frequently reflects the concerns of American society through its decisions on current contested cases. The fight for racial and social equality in the United States and the divisions engendered by the Vietnam war concerned social scientists who became involved in jury research and actual jury selection in celebrated court cases (Berk, 1977). This is not to imply that the Court had, until very recently, taken social science research seriously; Buckhout entitled his January, 1977 essay: The U. S. Supreme Court vs social science: the jury (emphasis added). Policy makers seem to use research to justify decisions they would make anyway (Weiss, 1970). Research is too often removed from reality, and too often testing some small effect while neglecting others equally important. Findings which conflict with what persons "know" in their day-to-day life (Meehl, 1971), do not enhance the research-policy relationship. Saks (1974), in reviewing some of the important decisions dealing with juries in the previous ten years, remarks that the Supreme Court's social science scholarship was of particularly poor quality. In Williams v Florida (1970), the Court decided that the size of the criminal jury can be determined by the state. Saks argues that the six studies cited by the Supreme Court majority were all methodologically unsound with one

particular study finding contrary to the Court's interpretaton. Referring to the Asch studies of conformity, Saks says,

> this is a classic and widely known finding in social psychological research on conformity. The presence of an ally is one of the most powerful known facilitators of minority resistance to conformity pressure (P. 19).

The study cited by the Court actually demonstrated, congruent with Asch's findings, that it is the absolute number of dissenters which makes likely a body of resistance; the Court held that it was the proportionate number of dissenters which was important, based its decision on the misinterpreted findings and tangentially generated an entire industry devoted to testing the differences between six and twelve person juries.

In Ballew v Georgia (1978) the Supreme Court reversed its neglect and distrust of jury research findings. On the basis of a rigorous review of social science research, the Court limited Williams by declaring that a minimum of six jurors is required when a defendant is subject to a substantial deprivation of liberty. Actually the weight of evidence was that 12 person juries were fairer than six person juries; but the Court used these findings to draw a line different from where the researchers thought the line belonged. The Court searched the social science literature for applicable jury size studies and found the entire field plagued by non-findings resulting from the one-sidedness of the trials presented to mock juries. Lack of verdict variability among the jurors led to support of the null hypotheses of no differences in verdicts due to the differences in the size of the juries.

On the basis of impressive critical reviews (Lempert, 1975; Saks, 1977) and well done purely statistical evaluations of jury size effects (Friedman, 1972; Nagel & Neef, 1975), the Court discarded the results of some poorly conceived research and decided that the interests of defendants (and especially minority defendants) were not sufficiently protected in juries with less than six members. In his concurring opinion, Mr. Justice Powell was less than impressed.

> ... I have reservations as to the wisdom -- as well as the necessity -- of MR. JUSTICE BLACKMUN'S heavy reliance on numerology derived from statistical studies. Moreover, neither the validity nor the methodology employed by the studies cited was subjected to the traditional testing mechanisms of the adversary process.

Despite the well defined problem in Ballew, the Court found it difficult to explicate matching, well defined answers in experiments expressly designed to test the difference between six and twelve person juries. It is ironic that both the appellant and respondent supported their case with empirical findings from published studies. Apart from its substantive import, this decision justified the torrent of research unleashed by previous decisions on jury size; and very strongly implied the continued dependence of the Court on scientific studies of jury and juror behavior.

The Supreme Court often poses the questions for our research. However, because of legislative constraints and generalization problems with mock jury research (Bray, 1976; Foss, 1975; Gerbasi, Zuckerman, & Reis, 1977), illuminating answers are not readily forthcoming.

Selecting the Jury Pool

In order to guarantee that jurypools (that body of potential jurors from whom the actual sitting jurors are chosen) represent a "fair cross-section of the community," The Jury Service and Selection Act of 1968 (hereinafter called the 'Federal Act') mandates random selection of prospective jurors from representative lists of citizens in each Federal District. The Federal Act requires the use of voter registration or voter lists, and suggests the use of other compilations (auto registration, auto licenses, utility records etc.) when necessary to ensure representativeness. In conformance with the Federal Act, the Federal Court, Eastern District of New York (our research setting) located in Brooklyn, New York, creates a master jury wheel for juror selection every four years using the registration lists of the presidential election.

A particular jury hearing a particular case need not be representative of the community, but the master wheel from which jurors are randomly chosen must be if the selection procedure is to conform with federal law.

Prior to 1968, selection procedures in the Federal as well as the state courts were more or less arbitrary with jurors often 'recommended' by someone of community standing (the Key Man system). These and similar selection procedures led to non-representative jurypools with the systematic exclusion of minorities, poor, young, old, women. Typi-

cally, the jurypools were male, white, middle-class with average ages of over 40 (Van Dyke, 1977).

Beiser (1973) compared federal and state courts in Rhode Island, both of which were using the same voter registration lists. The state's jury commissioner maintained investigators who visited each potential juror at home. Consequently there were systematic differences in age and sex variables (the two that Beiser examined) between jury panels: the state had a significantly higher mean age and significantly more males. Some states still use the key man or similar systems but most now follow the recommendations of the Uniform Jury Selection and Service Act, a suggested procedure for state courts which was modelled after the Federal Act. The two jury selection acts now provide for the most representative jurypools in jury history (Van Dyke, 1977) with demonstrated effects. For example, Ginger (1971) reports a substantial decrease in the conviction ratio in Baltimore since the replacement of elite 'Blue Ribbon' juries with 'Blue Collar' juries.

However the representativeness issue is far from settled (Simon, 1973; Van Dyke, 1977).

Beiser argues that voter registration lists are not necessarily representative because many groups do not register in the same proportion as their occurrence in the population: blacks, hispanics, the young, the old. Van Dyke (1977) reviews a number of studies which support Beiser's contention.

The selection of names for the master jury wheel is the first stage of jury selection. Other potential and actual sources of bias to fair representation occur at later points. In a study of a jury selection process essentially similar to that used in most federal courts, Alker, Hosticka, and Mitchel (1976) take census data for the area concerned, allow for all possible excuses and exceptions at each stage of the federal selection procedure and end with a final jurypool of 15,930 randomly selected venirepersons. This model of an ideal sample of randomized venirepersons is then compared to the actual jurypools selected by the Federal Court of Eastern Massachusetts during that period of time. Furthermore, comparisons are made at each stage of the selection process to discover the precise sources of variation.

Alker et al. summarize:

> Findings support the hypothesis that the juror selection process effectively discriminates against the poor, the young, racial minorities, women, and persons with low and high educational attainment. Sources of bias are found in the use of outdated voter registration lists, unreturned jury qualification questionnaires, and the excuse process. (emphasis added.)

Additionnally, Alker et al. and Van Dyke argue that such outdated voter registration lists systematically discriminate against the minorities and poor because of their greater mobility -- they change residence more often.

The Federal Act mandates a new master list randomly selected from voter registration every four years. During that four years the list is

not updated or modified. Of necessity, the youngest juror selected at the end of a four year term is 22 years old, compared to the 18 year old selected at the beginning. The members of the 18 to 22 year group are not only under-represented on the late jurypool, they are non-existent.

In a study undertaken to discover the effects of using voter registration lists for a term of four years, Baker (note 1) compared an actual jurypool of about 300 selected at the end of a four year term (the 'late' jurypool) to that of about 300 selected at the very beginning of a new four year term in Federal Court, Eastern District of New York (the early jurypool). Differential mobility in population groups should result in differential representation in these two jurypools selected at the same point in time but from different voter registration lists; one freshly chosen, the other 4 years old. The outdated list should supply an unrepresentative distribution of minority and blue-collar occupational groups.

These members of the jurypool filled out self-report questionnaires asking extensive demographic information during their first hour at court. With over 95% of the questionnaires returned and usable, Baker found that there were no significant differences (at the .05 level) for any of the demographic categories examined except age. After correcting for the differential age of the selection lists by adding four years to the age of the early jurypoolers, there were no significant

differences at all between the two jurypools. The similarity of the jurypools selected four years apart indicate that it is not mobility which leads to the discrepancy between jurypools and voter registration.

In a 1973 study of "The juror in New York City" conducted by a federal court committee (Simon, 1973) where 73% of the jurors were over 40 years of age (compared to the expected 57%) and 87.8% were males, the committee notes that "one-half or more of these persons summoned for jury service never come" (p. xvii).

Leo Levy, the Bronx County Clerk, has reported that "in Bronx County, we summon 1600 jurors for each two week cycle, of which 400 to 700 actually come" (personal communication, 1981).

The effect of the non-returned questionnaire or summons on the composition of the jury does not seem to have been systematically investigated. The most effective way of dealing with a jury summons when one does not want to serve is simply to ignore it. It seems likely that different demographic groups use this primitive but effective way of dealing with a jury notice differentially causing even more unrepresentative jurypool distributions and, to some extent, effectively negating procedures designed to achieve a more representative jury.

The final determinant of unrepresentative jurypools is the easy availability of excuses. Alker et al. (1976) and Van Dyke (1977) detail the excuse process which allows many professionals, small businessmen, and women caring for children under 16 automatic exemptions. More

important, perhaps, is the availability of excuses for anyone who does not wish to serve. One merely has to claim one of any number of likely events: sickness in family, rush time at job, school examinations, etc.

The non-returned jury notice and the easy excuse might very well be the major causes of non-representative jurypools when using random selection. There is a continuing disagreement concerning the ideal composition of the jury: some want a cross-section of the community; others a cross-section of responsible individuals. The argument has been presented that jurypools randomly chosen from voter lists at least represent those in the community who want to serve (Van Dyke, 1977), and that the excuse and unreturned notice problem further refine the list. The courts tacitly support this self-selection process by not prosecuting those who fail to appear when summoned and never attempting to verify tendered excuses. Society signifies its approval by not urging and financing such investigations.

Voir Dire: Selecting a Fair Jury

A defendant has a constitutional right in almost all instances to choose a jury trial (Baldwin v New York, 1970), but can request a bench trial (trial by judge only). In cases where the crime is a particularly heinous one, many defendants choose the bench trial. Additionally, trials are often avoided because the defendants plead guilty to a lesser charge with a promise of reduced sentence when the defendant feels his or her chances of conviction are high.

When a defendant does choose a jury trial, a panel of prospective jurors is brought into the courtroom for selection. These prospective jurors are then introduced to the principals involved in the case (defendants, prosecution and defense attorneys, judge) and given a skeletal outline of the alleged crime. If any of the prospective jurors knows one of the principals, or if for some other reason a prospective juror feels he or she cannot judge the case fairly, the juror is excused for "cause."

At this time the reduced panel undergoes voir dire (an expression usually translated as truth telling), which is the juror questioning process.

Voir dire is that aspect of the trial which is most diverse and negotiable throughout the country, sometimes differing within the same court. In fact, a particular judge can change the method he or she generally uses under petition from the attorneys because of some feature of the trial itself (e.g. Shulman et al., 1973). There are no constitutional proscriptions concerning voir dire and its form depends primarily on custom and judge discretion.

In this community (Brooklyn, N.Y.) the State Supreme Court is located one-half mile from our research setting. The Federal Court judges question prospective jurors during voir dire with suggestions as to emphasis and content from the opposing attorneys. At the State Court, the questioning is done by the attorneys themselves under the

watchful eye of the judge, and in civil cases the lawyers actually conduct the entire voir dire; the judges do not see the jurors until after they are selected.

It would be difficult to overstress the importance of voir dire. It has been said that "in England, the trial begins when the jury is picked; in the United States, the trial is over when the jury is picked" (Moore, 1973, p. 134). Various strategies have been developed to eliminate jurors thought prejudiced against one side or another. It is at this point that social scientists have recently had some inpact on jury trials.

The avowed purpose of voir dire is to help select a fair and impartial jury. If, under questioning, the prospective juror betrays an obvious bias, he or she is excused for cause. In addition, the prosecution and defense receive a number of peremptory challenges which vary depending on the seriousness of the charge. These peremptory challenges can be exercised to excuse any prospective juror without explanation.

Jurors are peremptorily excused for a variety of reasons: the prosecution may feel that the juror has an untoward sympathy for the accused; the defense that the prospective juror is unduly prejudiced. Whatever the reason, real or imagined, a peremptory challenge eliminates the prospective juror. Forming an accurate judgment of a prospective juror's inclination on the basis of his or her responses

during voir dire is a difficult task for which attorneys have little training. There is no shortage of advice; but this 'legal scuttlebutt' is notably lacking any empirical confirmation (Toch, 1961; Penrod, note 4).

However, it is a mistake to believe that selecting a jury is the only function of voir dire. The jurors are also given a 'feel' of court procedures. They learn a good deal about the pending case and the attorneys,and are immunized against opposing arguments. In a sense, jurors are socialized into the system by voir dire (Blunk & Sales, 1977).

Balch, Griffiths, Hall, and Winfree (1976) consider voir dire a rite of passage separating the ordinary man-in-the-street from the juror. This is a socialization process wherein the jurors learn what is expected of them and how to play their role.

> Ideally the juror should set aside most of the decision making rules of everyday life. He should delay judgment until all the evidence has been presented. He should not allow sympathy, personal prejudice, or personality conflicts to influence his decisions. In the course of deliberation he should avoid becoming so ego-involved in a particular point of view that he loses sight of the jury's principal goal -- the discovery of "truth" (pp. 16 & 17).

A content analysis of actual voir dires found that while less than 37% of the attorneys' and judges' statements concerned the jurors' personal-biographical characteristics, 43% were instructional in the sense that it informed the jury panel what was required and expected of them as jurors. Most of the instructional questions asked of the jurors (72%)

demanded a public committment by the juror to some standard of fairness and impartiality. The court with its formality, its impressiveness, its seriousness, its segregation from the mundane, lends itself to the acquisition of a demeanor appropriate to the role. Balch et al. claim that

> While it [the voir dire] is ostensibly a means of selecting unbiased jurors, it may also help create them....prospective jurors are taught the fundamentals of courtroom procedure and their role expectations are clarified (p. 19).

Plea Bargaining

In examining jury trials it is important to remember that we are merely studying the most publicly available feature of the criminal justice system. Nationally, less than ten percent of felony indictments reach trial (Ellison & Buckhout, 1981), and in some jurisdictions the figures are appreciably less (e.g., Blumberg, 1967).

Blumberg (1967) observed the workings of courts in or near metropolitan New York for five years and notes that:

> The rules of due process, as expanded and strengthened by the Supreme Court, are predicated on the existence of an adversary system of criminal justice.... Unfortunately, this model of criminal justice does not exist in fact. At each stage of the process a tacit but erroneous assumption is made. It is assumed that the accused will ultimately have 'his day in court.'... As a result, all the screening agencies in the system of criminal justice move a case along toward a trial--which seldom occurs (pp. 26-27).

Blumberg claims that, because of plea bargaining, less than four percent of grand jury indictments in the NYC courts over the fifteen year period 1950- 1964, for example, were disposed of by jury trial.

However, despite the jury trials rarity it has an extraordinary effect felt throughout the system since no one knows which defendant will eventually have his day in court.

While plea bargaining is constitutionally correct and seems almost necessary for the continued functioning of an overworked CJS, it leads to anomalies detrimental to society and to many of the actually innocent defendants (Finklestein, 1975). Those who are guilty receive lesser sentences as part of the bargain; those who are innocent but accede to the pressures to plead by promises of suspended or very short sentences are then labeled as criminals and subject to future harassment because of their resulting record.

Blumberg says there is

> an almost irreconcilable conflict: intense pressures to process large numbers of cases, on the one hand, and the stringent ideological and legal requirements of 'due process of law' (p. xi).

Blumberg reminds us that organizations and systems consist of individuals who learn the behavior appropriate to their particular position from their interactions with others in the system: the pattern of expectancies, rewards, and punishments. Blumberg makes no claim that these phenomena are unique to the CJS; they exist, although rarely seen with such clarity, in almost all organizations of work (Sofer, 1972). But it is important for others, familiar with such conditions in their own lives, to realize that these phenomena affect the dispensation of justice despite the formal and explicit goals of the CJS.

The Adversary System

While there are slight differences between trial procedures in the many jurisdictions, the criminal trial has a typical (and sometimes criticized) adversary format.

The psychologist-lawyer team of Thibaut and Walker (1975) report a series of interdisciplinary and cross-cultural experiments. Noting that "the continental judge exercises considerably more control over the [court] process than does his Anglo-American counterpart" (p. 23), they investigate the perceived fairness of adversary versus the inquisitory procedures they see as the obvious alternative.

In inquisitory procedures the third party (a judge or judges) has complete control over the proceeding and the investigation for information; the adversary procedure sets two parties against each other according to some prescribed set of rules; the control over the investigation of facts and the order and presentation of witnesses is almost entirely in the hands of the contending parties. Thibaut and Walker presented cases varying the degree of judge control, participant control, opportunities to present evidence, the order of their presentation, and the amount of pretrial bias to students at two locations in the United States, in Paris, and in Mannheim, Germany (to control for cultural differences).

Their finding generally support the Anglo-American adversary system as the procedure perceived fairer by participants and

observers. The adversary structure favored the party at a disadvantage early in the proceeding by encouraging the attorney to 'even it up' through further diligent investigation. Furthermore, it is well known that premature hypothesis forming distorts the perception of future events and inhibits a search for supporting facts (e.g. Bruner & Potter, 1964); the adversary structure seems to compensate for this by some sort of 'reactance' which causes the jurors or judges to search for contradictory information.

Thibaut and Walker comment:

> The adversary role structure therefore seems most congruent with a public policy requiring overwhelming proof before a verdict can be rendered. In contrast to the inquisitorial model, the adversary system systematically compensates for possible sampling error[in the discovery of facts] (p. 40).

There is a strong recency effect in the presentation of evidence. The usual primacy effect observed in other circumstances is controlled by warning jurors of evidence to come and the aforementioned reactance effect. The ordering of weak to strong in the presentation of evidence for each side is the most effective and is usually presented that way by the opposing attorneys, the plaintiff opens and closes the arguments and counter-balancing is prevented. However,

> the traditional Anglo-American adversary trial is remarkably well arranged to neutralize the ultimate effects of order and thus to maintain the decision making process relatively free of this powerful irrelevant influence [recency effects and guilt emphasis] (p. 66),

and, at the same time, produces the usual effect in small group research of greater input and control. Thibaut and Walker report that these findings hold across cross-cultural boundaries and led to greater satisfaction of participants and observers in each of their student populations.

Jury Research

The seminal work in the jury research area is The American Jury by Kalven and Zeisel (1966) in the sense that, along with the more recent efforts of the National Jury Project, it is chiefly responsible for the lately renewed interest in the jury trial by the social sciences. This book, despite serious methodological difficulties, is a monumental achievement of the University of Chicago Law School's Jury Project. It is a non-ending source of hypotheses to be tested and its conclusions have been cited by numerous judges including several Supreme Court justices, sometimes on both sides of the case. For these reasons, let us review this book in some detail.

The Chicago Jury Project received completed questionnaires from 555 judges concerning 3576 jury trials over which they presided. Agreements and disagreements between judge and jury were noted (judges were asked how they would have voted if they had had the responsibility of making the decision). Of the 3576 trials, 1063 disagreements between the judge and the jury were recorded. The judges

were asked to report the reasons for disagreement, and of the 1063, 101 were left unexplained. It is the remaining 962 disagreements, for which Kalven and Zeisel have a judge's explanation, which form the universe of cases and supply the data. This book, then, is an attempted explanation, using the techniques of cross-tabulation and 'reason assessment,' of these 962 disagreements. Kalven and Zeisel make the point that the only reasonable alternative to the jury trial is the 'judge only' trial and this book is a comparison between the two.

In 90% of the disagreements investigated plausible reasons are available for the reported differences and in only one case does a judge advance jurors' misunderstanding as a reason for disagreement. A cross-tabulation of easy and difficult cases shows no differences in the proportion of disagreements and Kalven and Zeisel conclude that the jury generally understands the case and what's expected of them, and that the reasons for disagreement lie elsewhere. This is an important finding since there has been much written about the purported difficulty of uneducated and amateur jurors understanding the intricacies of many criminal cases (e.g. Emerson cited in Moore, 1973; Frank, 1949).

In 45% of the cases the jury makes its decision on the evidence plus one or more other factors. This leads Kalven and Zeisel to three hypotheses with which they explain the judge-jury disagreements.

1. The liberation hypothesis: the most basic thesis explaining the findings of this study is that "the jury does not consciously

and explicitly yield to sentiment in the apparent process of resolving doubts as to evidence" (p. 165). The closeness of the evidence 'liberates' the jury to decide on other bases, it justifies decisions on these other bases, a process which the jury itself (and, incidentally, the judge too) does not realize is taking place but which becomes clear when one assesses the universe of disagreements as a whole.

2. The credibility hypothesis: "Disagreement arises because one decider accepts and the other rejects a given item of proof" (p. 166).

3. The reasonable doubt hypothesis: "Disagreement arises because the jury will tolerate less doubt in convicting than will the judge" (p. 167).

From which follows Kalven and Zeisel's general conclusion:

> when the jury reaches a different conclusion from the judge on the same evidence, it does so not because it is a sloppy or inaccurate finder of facts, but because it gives recognition to values which fall outside the official rules.

In his 1970 review of jury research, Erlanger argues that because of methodological problems "the precision of the Kalven and Zeisel data ought not to be taken too seriously". Despite this warning the American Jury remains the most influential work in the jury research area.

Juror Recall and Understanding

Herbert Spencer called the jury "twelve men with average ignorance," but let's see what the research reports.

Hunter (1935) questioned jurors after trials and concluded that the 'typical' juror does not understand the law relevent to the case being tried. In 1941, Harvey (as reported in Erlanger, 1970) sent questionnaires to jurors who had completed service. 185 were returned, and about 40% admitted that they did not understand the judges' instructions.

More recently (Sigworth & Henze, note 5), seven judges tested jurors' understanding of the uniform jury instructions in Pima City, Arizona. 185 jurors had a mean score of 52.06 out of a possible 100; "Jurors don't understand half of what they're told" (p. 5). Those without a high school education scored the lowest in the sample, but those with a university degree scored only seven or eight percent higher. Having written instructions available during deliberation increased understanding, and rewritten instructions for criteria of simplicity and understandability improved scores substantially.

It seems clear that when individuals are tested for recall and understanding they fail to do well. And yet, Kalven and Zeisel conclude strongly that jurors <u>do</u> understand the cases and Simon (1968) reports that despite personal preference for the Durham rule (regarding the legal meaning of insanity) jurors understand when instructed to

follow the M'Naghten rule. More recent empirical work on juries rather than jurors (Bridgeman & Marlowe, 1979) does not support the hypothesis that jurors fail to understand the relevent legal issues. Group recall is more efficient than that of any individual juror (Kaplan & Schershing, 1978), and interactions during deliberation probably helps further the recall and understanding of the jury. Kaplan and Schersing stress one of the important points supporting the idea of having a jury decide a case, and of having as large a jury as practicable: different jurors pay attention to different parts of the evidence and therefore the collective memory of the jury is better than any individual member, and more influential.

Race and Ethnicity

Percy Foreman is one of the "super lawyers" who have given the lesser talented some pointers on how to conduct a criminal trial and how to pick jurors (Dorman, 1969). For example:

> You should never allow the defendant to be tried. Try someone else -- the husband, the lover, the police or, if the case has social implications, society generally. But never the defendant (p. 1).

Foreman wants jurors that work with people rather than things; a truck driver rather than an engineer, a salesman rather than an accountant.

> If a man has never been in a bar -- or was brought up in a home where a deck of cards wasn't allowed -- he may not be able to understand some bully who gets full of beer and then gets killed when he tries to whip the biggest man in the place (p. 135).

When complex testimony is to be presented, Foreman prefers educated jurors. He never allows anyone who has had any contact with law-enforcement to become a juror. Where long terms are possible he prefers older jurors who understand the passage of time and how horrible it could be. He likes his juries to consist liberally of oppressed minorities: Jews are best; next are Blacks except when Blacks are on trial; the Irish and Latins complete the list. He avoids the Nordic types.

Other famous defense counsel have made similar assertions (e.g. Darrow, 1936; Goldstein, 1967); generally, they advise defense attorneys to choose jurors of Latin descent, Blacks, and Jews; and avoid the non-latin Anglo-Saxon types who are more likely to convict. The research literature seems to support this position but not without some dissenting voices.

Reed (1965) sent questionnaires to 432 jurors who had sat on 36 trials which came to a verdict in East Baton Rouge, Louisiana -- the entire population of deciding jurors during a two year period. After phone calls and personal visits when needed, about 145 usable responses concerning criminal cases were obtained from the group.

Southern Louisiana is populated almost exclusively by Cajuns of French descent (Latins); while Northern Louisiana is populated by Anglo-Saxons. The court used both of these populations as their jurypool and provided a naturally occurring between groups comparison (Cajuns vs. Anglos).

TABLE 1

Birthplace and Voting of Responding Petit Jurors[1]

Birthplace	Voting[2] Guilty	Not guilty
South Louisiana	39	44 .
Not South Louisiana	40	17 .

[1]As reported by Reed, 1965
[2]X^2 (1) = 7.4 $p<.01$

Examining Table 1 for which Reed has supplied the numbers, we see that here is strong empirical support for the attorneys' thesis. Furthermore, Reed reports that prior jury service and Anglo-Saxon descent were the two most important predictors of a guilty vote: Anglo-Saxon jurors with previous service were most likely to convict; Latins with no previous service were the least likely to convict. Unfortunately, he does not supply this additional breakdown of the data nor report any supporting statistics.

Field studies by the Chicago Jury Project (Broeder, 1965; Simon, 1967) report similar findings. But laboratory studies present conflicting results. Mock jurors are usually from similar social classes and educational backgrounds, a lack of variability which makes differences difficult to detect. Gleason and Harris (1975) as well as Nemith and Sosis (1973) support the null hypothesis of no difference between

black and white racial groups; whereas Ugwuebi (1976) and Miller and Hewitt (1978) find that each racial group identifies with the alleged victim of similar race in the production of sentences and verdicts.

Personality Characteristics

Researchers have long been fascinated by the possible connection between personality and non-evidential factors. A person posited to have a predisposition to convict is called 'conviction prone,' and the personality correlates of conviction proneness have been extensively investigated.

Mills and Bohanon (1980) hypothesized that a person's character structure influenced his or her juror voting. Character structure was conceived as "largely unconscious, typified ways of selecting, using, justifying, and enforcing rules" (p. 662). Mills and Bohanon, with the permission and cooperation of the City of Baltimore, sent questionnaires to former jurors who had actually sat on cases. Matching these self-reported data with court records, they found character structure significantly related to juror decisions as well as perceived participation and effectiveness. Moreover, race, age, education, and especially sex were found to modify the personality effects (but see Penrod, note 4 for completely different results). In sum, all of the tested factors were significant, supporting the view of a large majority of jury watchers who believe the jury affected by various non-evidential factors.

Authoritarianism. Authoritarianism is a personality construct hypothesized to accompany undue deference to authority. Adorno, Frenkel-Brusnswik, Levinson, and Sanford (1950) originally created the F scale (F for fascism) to measure the construct which has subsequently been found to correlate highly with a host of undesireable personality characteristics: xenophobia, anti-semitism, racism, intolerance and emnity towards low-status groups. Despite the construct's undeniable shortcomings, Christie has claimed authoritarianism as the single most important predictor of conviction proneness (personal communication, 1978), and there is substantial empirical evidence that high authoritarians are more likely to convict.

In order to discover whether juror decisions differ along a liberal-conservative continuum, and to test the rigidity of extreme political views, Boehm (1968) created a new scale, The Legal Attitudes Questionnaire (LAQ), of ten items, each with three possible answers: authoritarian, anti-authoritarian, and egalitarian; reflecting right wing, left wing, and liberal attitudes respectively. Subjects chose the answers they agreed with most and those they agreed with least, supplying a score for each of the three political positions: a three was scored for agree, two for a blank, and a one for disagreement. Student subjects took the LAQ, heard a case either heavily weighted towards innocence or guilt, then suggested a sentence for guilt. There was a significant difference for severity of sentence by type of case

with authoritarians giving longer sentences when the defendant was obviously innocent and the anti-authoritarians giving shorter sentences in the obviously guilty condition. In addition, the authoritarians were more confident of their verdict.

Jurow (1971) used a number of personality measures to plumb the connection between belief in capital punishment and verdict preferences for a group of mock jurors who listened to two tape recorded recreations of murder trials. He concluded that the LAQ was the best predictor of 'conviction proneness.'

Mitchell and Byrne conducted two studies investigating the relationship between the jurors' irrelevent evaluation of a defendant's attractiveness, their legal-judicial decisions, and the relationship to authoritarianism. In the first experiment (1973) 139 college students were measured on an acquiescence free F scale prior to the reading of a case where an attractive or unattractive fellow student purportedly stole an examination. As predicted there was an interaction for authoritarianism and the attractiveness of the defendant: low authoritarians did not allow the attractiveness of the defendant to influence their judgments of guilt or recommended punishment; high authoritarians gave more severe judgments to the unattractive defendant and felt less certain of the guilt of the attractive defendant.

In their next reported experiment, Mitchell and Byrne (note 3) then investigated the possibility that high authoritarians do not ignore

biasing information when instructed to do so by a presiding judge, while low authoritarians do. The theoretical prediction is not so clear when one considers that high authoritarians are more subject to experimental demand characteristics (Ruben & Moore, 1971) and more submissive to persons of authority (Adorno et. al., 1950). This second study tests the alternate hypotheses generated: high authoritarians can not separate their subjective feeling about a defendant despite a high status person's (the judge) clear instruction; high authoritarians would follow the instructions of the high status person, and the low authoritarians would not. Once again about 140 college students constitute the sample for this experiment. They read a modified version of a negligent homicide case originally used by Landy and Aronson (1963), and then read either biasing negative or positive testimony; with no instructions, or instructions by the judge to ignore, or instructions by the judge to pay special attention; generating a 2x2x3 factorial design: high or low authoritarians, positive or negative biasing, no instructions, or instructions to ignore or instructions to attend. Subjects supplied the dependent measures by rating certainty of guilt on a seven point scale, recommended length of sentence, decided when defendant should be eligible for parole on a ten-point scale (it is much easier to analyze interval data in these factorial designs); and then rated the defendant on 13 evaluative bipolar adjectives. Authoritarianism was measured on the same type of scale as in the previous experiment.

The authors note that

> since completion of the simulated trial experiment took less than the students' required hour of participation for this experiment, they were asked to complete a general opinion survey for another research project (p. 8, emphasis added).

Despite the basic construct of authoritarianism being the deference to persons of authority, the high authoritarians did not do as instructed by the high-status judge (apparently assigned status in a written summary can not simulate the power of attending a high status person in the full majesty of his or her courtroom). High authoritarians recommended more severe sentences, paid more attention to negative biasing information, and could not separate their evaluative judgments from their "legal-judicial decisions."

The authors conclude by suggesting the possibility of real trials with real judges influencing people differently from such a simulated trial since "verbally simulated trials cannot make salient many situational influences which might be operative in actual trial" (p. 12). And then, in the almost obligatory coda this research seems to require, do not hesitate to suggest we change our trial procedures to minimize "the apparent authoritarian proclivity to be influenced by irrelevant and biasing information during a trial" (p. 13).

Berg and Vidmar (1975) hypothesized that high and low authoritarians would focus on different aspects of the evidence and the subsequent recall of evidence would reflect this focus. High authoritarians are source oriented, low authoritarians are message oriented. The

student subjects were median split using scores of Boehm's LAQ and then read a case of student cheating with status and situation information varied (similar to Mitchell & Byrne, 1973). The hypotheses were confirmed; additionally, high authoritarians were especially severe with low status defendants. The authors conclude by suggesting more research attention to the personality characteristics of jurors.

Kalven and Zeisel (1968) and Kaplan and Schersing (1978) have suggested that group decisions polarize towards the initial majority sentiment. Bray and Noble (1978) investigated the effect of authoritarianism on individual jurors, juries, verdict changes, and the polarization thesis. An acquiescence free F scale (as Mitchell & Byrne, 1973) was used to separate 280 college students into high and low authoritarianism for their two hour experimental session. Subjects then listened to a 30 minute recording of a murder trial, judged guilt (yes or no), and recommended sentence presuming guilt (14 to 99 years). Answering questions as individuals before and after deliberation, they came to verdicts as members of high or low authoritarian juries.

Consistent with other studies, high authoritarians judged more guilty verdicts and more severe sentences both in the individual and group conditions. As expected, juries polarized towards majority sentiment; and, finally, high authoritarians changed their verdicts more often than low authoritarians.

Bray and Noble admirably restrain themselves and comment that "results reported here are more properly viewed as suggestive rather than definitive" (p. 1430).

Age

Attorneys feel that older jurors are more conviction-prone, but the empirical evidence is not clear. Faust and Carlson (1977) found strong relationships between age and attitudes on capital punishment. Scroggs (1976) found older jurors more conviction-prone. However, Simon (1967) found no such significant differences in her measures for the different age groups, and Reed (1965) tested the age groups in his survey of real jurors in Louisiana and found no age differences.

The issue is undecided at this point, but recently the courts have taken steps toward recognizing the importance of age by declaring the young juror (18-34 years old) a cognizable class for purposes of deciding whether or not a 'fair section of the community' is represented (e.g. Commonwealth v Bastarch, 1980).

Social Status: Occupation and Education

Attorneys probably receive more detailed advice about the effects of jurors' social status on their verdicts than all the other posited influences combined. For example, Clarence Darrow has said that the wealthy consider the jail a building second in importance only to the

board of trade. Rita James Simon (1967) notes that trial lawyers consider lower status jurors as more likely to favor defendants; that most criminal attorneys make the same ethnic distinctions as Foreman above; and that there is ample evidence that higher status jurors are more conviction prone.

Using subjects selected from an actual jurypool to judge a simulated trial under differing defenses of insanity, Simon reports that about ten percent of the juries that reached consensus were convinced by jurors who were originally in the minority, but

> there was no instance in which one juror or even two held out against the other ten or eleven and finally persuaded them to their position. In those minority juries higher status jurors were overrepresented.

Opinion pollsters, attitude researchers, and social scientists generally, believe that one's position in society is a pervasive influence in the perception and valuation of various issues. Simon concludes that

> status distinctions as occupation and education are meaningful predictors of jurors' attitudes and behaviors. In other words, the jurors' behavior contains the same general regularities that we have learned to attribute to all human behavior in our society.

Furthermore, she confirms the widely held notion that lower status jurors are more defendant prone.

Gleason and Harris (1975) investigated the hypotheses that socio-economic status rather than race similarity is the major predictor of guilt verdicts, and that blame judgments and guilt verdicts are not comparable. Four different cases as summarized by police files were

presented to subjects in a 2x2 design varying race (black or white) with either low or middle class status. Subjects then used a seven point bipolar scale to estimate:

1. potential guilt of defendant,
2. if defendant is guilty then length of sentence,
3. defendant blame,
4. how harshly should defendant be treated,
5. likelihood of self being in such a position,

The experimentors found that high socio-economic defendants were seen as less guilty ($p<.005$), received lower sentences, and received more lenient treatment. When the subject viewed himself as not likely to be in such a position, he assigned more blame; but, in general, middle SES blacks and lower SES whites received more blame. Blame and juridic decisions did not show the same pattern; and, as many studies have shown, those who favor death sentences were more conviction prone and gave more severe verdicts.

Reed (1965) found that the higher his or her education, the more likely the juror would vote guilty (X^2 (2) = 9.846, $p.<.01$); and the higher the social class as determined through juror occupation, the more likely the juror would render a guilty verdict (X^2 (2) = 16.760, $p.<.001$). Adler (1973) reported essentially the same results.

Attractiveness

In what has become an oft cited classic in the jury research literature, Landy and Aronson (1969) report two experiments. In the first experiment subjects read cases which vary the attractiveness of the victim in terms of personality profiles, then sentenced defendant to jail for from one to twenty years. The mean sentence assigned in the attractive victim condition was 15.77; for the unattractive victim, 12.90 -- not significant at the .05 level.

The second experiment incorporated a 2x2 design varying victim and defendant attractiveness. Sentence differences once again were not significant for the attractive victim condition, but were in the same direction as in the first experiment. The range of possible sentences were different in the two experiments necessitating a transformation to Z scores for purposes of comparison across experiments. An ANOVA performed on these transformed and combined scores resulted in a significant difference for the main effect due to defendant's character and attractiveness, $F(2,11) = 3.27$, $p.<.05$.

In three studies designed to examine the effects of real versus hypothetical consequences of juror verdicts, Wilson and Donnerstein (1977) found an interesting pattern of verdict differences. Student jurors either thought their decision would have real consequences for a student accused of cheating on an exam or that the case was hypothetical without any real consequences. In the first experiment guilt and

sentence judgments were dissimilar, with manipulated character attractiveness not resulting in differences in guilt but doing so for punishment. The second experiment showed no differences on any measures. The third experiment found that "real" jurors recalled more evidence than the "hypothetical" jurors. The authors state:

> In all three studies, more guilty verdicts occurred in the real consequences condition than in the hypothetical consequences condition. It was concluded that much current research on simulated juries may be misleading and that more attention should be given in the future to the variable of real versus hypothetical consequences.

Validity Problems with Jury Research

Both prior to and during trials there are evidential hearings which determine which evidence can be presented to the jury without creating unfair prejudice. At every point in the defendant's 'career,' there is an extraordinary effort to shield him or her in the traditional 'cloak of innocence' in the eyes of the jury so that its search for facts is not biased. But it is difficult to evaluate the effectiveness of these efforts. Jury deliberations are shielded from scrutiny by custom and the force of law necessitating research techniques which do not examine the deliberations directly. There have been some retrospective studies of real jurors, but it is the mock jury experiment that dominates the field. The very structure of the trial lends itself to controlled experimentation.

Mock Jury Research

The jury trial is an adversary procedure but the trial is an ideal vehicle for psychological inquiry since its structure parallels that of the laboratory experiment in so many ways and it occurs naturally and frequently in American society. Similarities to an experiment, especially the search for facts and the protection against possible bias, strongly present themselves. All the relevant information is presented in an orderly fashion during its tenure (as with the experiment), both are standardized with procedural safeguards against explicit and implicit confounds, both try to ascertain the truth of a statement (the guilt of the defendant, the truth of the competing hypotheses), both are weighted against the undesireable and more damaging Type I error, the false positive. Both the jury trial and the experiment search for 'truth' in a strictly controlled environment, both use randomization to control for error and prejudice, and both use decision rules which admit of uncertainty.

It is these similarities which attract experimenters and have contributed to the recent profusion and confusion of jury research. This seeming fitness for investigation has also generated jury research's greatest problem: the equating of experimental findings with real jury effects. There are certain essential basic features of a jury trial which must exist before a trial can be said to have taken place. For example, the selection of jurors during voir dire, the viewing of witness testi-

mony and evidence presentation, and the unrestricted deliberation of jurors should surely be considered the absolute minimum features of a decently simulated jury trial. Some features such as the responsibility for another's life can not be duplicated, but it is fair to say that the closer the real thing is approximated, the more confidence we can have in the results.

In 1975, Foss reviewed the procedures of 23 jury simulation studies and asserted that their external validity was questionable; nearly all the studies were of individual judgment. But jury decisions are not averages of individual judgments; most jury studies collect continuous data, but the jury operates under a simple guilty-not guilty dichotomy. It seems that most of the jury studies examined were actually using the jury paradigm to study psychological questions rather than studies of juries qua juries. The jury and the trial were not the objects but the tools of study.

In their thorough and skeptical review of jury research for a legal audience, Elwork and Sales (1977) question much of jury research design. For example, they note that most mock jury studies collect sentencing data instead of the guilty-not guilty decisions jurors ordinarily make, and then base their inferences on these data. But in criminal cases jurors judge the facts; it is their exclusive province. Upon delivery of a guilty verdict it becomes the judge's task to assign sentences. Elwork and Sales feel that it is appropriate for different

people committing an identical crime to receive different sentences; more dangerous individuals are, and should be, dealt with more severely than the less dangerous. Sentences are discretionary simply because the judge should take matters extraneous to the present offense into consideration: character, past record, possibility of repeating offense, danger to the community, and so forth. When, as is usually the case, authoritarians mete out more severe sentences, inferences from these data are irrelevant to their performance as jurors. It is likely that authoritarian jurors function differently from non-authoritarians, but it does not necessarily follow from these data.

When character information is used to differentiate attractiveness researchers usually find that subjects assign severer sentences to the unattractive defendant (e. g. Landy & Aronson, 1969). It is inappropriate to conclude that attractiveness affects juror decisions, since sentencing is the function of the judge. When physical attractiveness is the manipulated independent variable, researcher inferences are even less appropriate: It may be true that good looks are regarded favorably, but "in a real courtroom situation ... a jury may get to know a defendant beyond his physical characteristics," because of his or her actions during the trial.

Elwork and Sales suggest that the plethora of poor jury research results from researchers' ignorance of the law and actual trial conditions. Echoing Foss, they feel that some jury studies seem designed to

test the "basic tenets of human behavior" instead of actual juror behavior.

As noted above, the voir dire is an integral part of any jury trial. Bray (note 2) surveyed 45 reported jury experiments. In only one of these studies did the juror undergo anything resembling a voir dire. Most attorneys who conduct voir dire use it to socialize the juror and extract promises of fairness (Balch et al., 1976). To the extent that this voir dire socialization is effective, mock jury research must be misleading!

Erlanger (1970) reviewed the jury research of the previous ten years. He warned then that student subjects had contributed all they can and suggested the use of quasi-experiments with different populations for further study. But most reported jury studies still use students as subjects (Ellison & Buckhout, 1981) with obvious generalization problems: students are a largely homogenous group who are not the repository of community common sense; and are not representative of the juror population.

Now the jury paradigm -- in the sense that a group of peers make judgments about a particular issue -- can be useful in other settings and for other purposes (e.g. Kantrowitz, 1977; Wolf & Arnstein, 1975). Levine (1974) has used juries to evaluate the effectiveness of graduate programs, and has suggested the use of the adversary model and juries to decide between competing scientific hypotheses when the empirical

evidence is not completely clear. While juries make decisions in these instances, in no sense can they be considered examples of criminal trial juries; they are missing one or more of the necessary ingredients.

Actually, as Kaplan and his colleagues have shown (Kaplan & Miller, 1974; Kaplan & Schersing, 1978), the experimental jury paradigm has been useful in the study of human behavior; paradoxically, it's not at all clear that it is very useful in the study of real jury behavior.

Most of the published research deals with the juror as an individual because of the relative simplicity of its data collection and analysis. Jurors are first presented with different versions of a trial and then differences in opinions and verdicts (more often, since they are on an interval scale, sentences) are related to various socioeconomic categories or personality traits. Regrettably, much of the data are not comparable because of differing criteria and measuring instruments. More regrettably, most of the results come from poor simulations of what a trial is like and what a juror must do. The results can not be generalized to any population of trials with any confidence. The jury trial is used as an instrument of person research rather than an ecologically valid simulation which might tell us something specifically about jurors. To the extent that the jury process is simulated in all its complexity and all its constraints, the experiments generate confidence in their generalizability to real jury phenomena; but faithful adherence is the exception rather than the rule (Bray, 1976; Foss, 1975). Faithful as the

reproduction might be, there is no way that the consequences of the decisions can be approximated. The responsibility of deciding another's fate is a heavy one, generating feelings which are impossible to recreate within the limits of ethical experimentation.

Jurors undergo intense socialization for their roles in which they make decisions which have awesome consequences. Buckhout (1978) reports that, after the final jury vote in U.S. v Swinton, an informal poll by the jurors found that the majority thought Swinton guilty of 'something,' despite the fact that they acquitted the defendant. As jurors their vote was recored as twelve to zero for acquittal; as persons their vote was four to eight. Would it be possible to simulate such a result in the laboratory? Balch et al. (1976) have provided us with an intuitively reasonable explanation of why mock jury research is not generalizable. Even under the best of experimental conditions, the jurors are not sufficiently impressed or prepared for their new roles.

In addition to those previously mentioned, there are other features a trial must have before it may be truly characterized as a "jury trial." For example, the order of argument and evidence presentation is fairy rigid in almost all jurisdictions; procedural safeguards are intrinsic to a proper trial. in all jurisdictions juries are the finders of fact, a function they fulfill by viewing evidence and witnesses and judging their relative credibility and truthfulness. Most importantly, the jury reaches its decision in its own time and is not answerable for its findings.

To be more specific, jurors

1. do not view trials without sufficient, detailed, and often interminable instructions regarding their duties,
2. do not judge guilt on any continuous scale, but get involved explanations as to the meaning of 'reasonable doubt,'
3. do not recommend punishment or sentences for innocent or even guilty defendants,
4. do not have a limited time for deliberation,
5. do not make repeated judgments or view repeated trials,
6. do not undergo written tests before judging a defendant (In point of fact, our understanding with the judges concerning this research specifically barred any questionning, oral or written, of the jurors until they had completed their jury service because of possible improper influence.),
7. do not come from homogenous educational strata with the same or similar experiences,
8. do not have to justify their verdicts outside of the jury deliberation room (recently, the Supreme Court reversed a guilty verdict because the presiding judge asked questions of the jurors pertaining to their verdict.),
9. are read the "Allen Charge" when the jury is split in an attempt to force a decision (often several times) [The "Allen Charge" is a standardized attempt by the judge to break voting deadlocks

by noting that any other group of jurors would do no better in deciding the case and appealing to the jurors to re-examine their reasons for disagreeing with their fellow jurors.],

10. and, do have a real responsibility with possibly awesome consequences.

Gerbasi et al. (1977) in their fairly recent review of jury research are careful to restrict their comments to mock jury studies since

> the extent of comparability between responses of mock jurors, whose judgments have no real impact, and verdicts of real jurors is unknown (p. 324).

In their summary abstract they conclude that

> data have indicated that many extra-evidential factors contribute to mock jurors' verdicts. However, the applicability of these results is limited by the many methodological and sampling problems inherent in the studies (p. 323).

The Law vs. Social Science

Jury studies, of interdisciplinary interest, are published in different types of journals. Many of these journals do not employ peer review, have editors and readers not trained in research methodology, and cater largely to the legal profession. One of the most difficult tasks facing a researcher who has both a technical and a lay audience is to have enough technical material in the article to support his or her conclusions and, at the same time, write interestingly enough to hold the attention of the layman. One would expect different emphases to emerge according to the targeted audience. The lawyer and the social

scientist have been trained differently and have different concerns. Social scientists are members of academe and primarily theoretical; lawyers are practitioners who would like practical rather than theoretical conclusions. With respect to jury research, lawyers would like a menu of important factors they can use during voir dire. Lawyers express their beliefs in homilies (often contradictory) and have faith in their intuitive use (Penrod, note 4); social scientists define hypotheses and are often skeptical even of their own empirical findings. The pressure to publish contributes to the problem of muddled jury research; when one has executed a lengthy program of research and can't find an audience in one's own discipline, one publishes where he or she can. This is not to denigrate these publications, but for most academics they are a secondary resource, to be used when the primary resource is not available.

The situation is such that poorly grounded but scientific sounding research is published in lay journals and mistakenly seen as representative of the latest scientific findings on the subject. Furthermore, the demand characteristics are such that even careful researchers are inspired to draw unwarranted conclusions and make gratuitous recommendations.

This problem is noted here not to demonstrate a theoretical point or to erect more methodological barriers for researchers to hurdle. But before social scientists present findings which may be used to change

society, they had better be sure these finding are real and to the point. Changing the criminal justice system is not some idle intellectual exercise. Real people have their lives affected in important ways when legal systems change.

It is difficult to do meaningful applied studies in the courts. During my first project in a court an undergraduate student, who actually supplied most of the contacts we needed, came to some misunderstanding with a judge whose cooperation she was trying to secure. We were then called on the carpet, lectured, and literally banished from the courtrooms by the chief judge. We were unable to complete our project, almost six months of daily court monitoring was lost, and even more regrettably we lost the cooperation -- I should call it collaboration -- of an intelligent and interested group of judges. As experimenters and project directors we have the power to choose our subjects and our environment. No such opportunities are out there in the real world of powerful persons and powerful interests. Tedeschi and O'Donovan (1971) have noted that psychologists have studied only the powerless (they are the ones who come into our laboratories), ending up with a psychology of the powerless. In the courts the researcher is least powerful, subject to the whims of guards and clerks as well as an assortment of judges, an atmosphere not conducive to attracting academics.

In an essay directed to an interdisciplinary audience, Meehl (1971) suggests that one should be suspicious of social science studies that refute widely held common sense notions. "We psychologists should be cautious where an alleged principle of modern behavioral science appears to conflict with the fireside inductions" (p. 85).

"Fireside induction" is what everybody, except the skeptical social scientist, believes about human conduct. Meehl claims that the typical social scientist is skeptical of the law and common sense but accepting of experimental evidence, while the legal worker often questions the value of experimentation: if research agrees with common sense then it is a waste of time; if it doesn't, it's not to be believed.

Meehl presents the differing views regarding the value of punishment as a case in point. Most social scientists doubt the efficacy and/or deterrent effect of punishment, yet in real life deterence often works: What social scientist would park his or her car in a "no parking" zone while a police officer is watching? Petain took over a mutinous French army in World War I and quelled disturbances by the wholesale execution of soldiers. During WW II those OPA districts that punished infractions of the law had higher rates of compliance. Perhaps the punishment studies have neglected to test an important variable; it might be that the certainty of punishment is more important than the severity. The union of Law and Social Science is not easily implemented. Most legal scholars are beginning to learn about signifi-

cance tests; Meehl imagines that it will take a long time before they learn what is wrong with them. Meehl concludes this perceptive essay by suggesting that the best way to study the criminal justice system is in real life settings using quasi-experimental designs.

Ceteris Paribus

The differences being tested in simulated jury studies probably do not have any substantial effect on juror decisions except in ambiguous cases because they are dwarfed by such important factors as evidence and witness credibility (Saks, 1976a); but when reporting experiments and advising the legal profession, researchers often neglect to stress the tacit 'all thing being equal' which they take for granted but others forget. As researchers the ceteris paribus clause figures prominently in our work and in the interpretations we make. It is part of our 'background of shared meanings.' It is understood, almost intuitively, that when we discuss the effect of, say, authoritarianism, we mean over and above the effects integral to the criminal trial itself; however, this emphasis on marginal issues serves to becloud the ever present, strong determinants of real verdicts -- such as evidence -- which we 'hold equal.' Saks' (1976a) reminder of the importance of the evidence in the trial may not seem too profound to researchers, but it was a well-timed, necessary corrective to uncalled for flights of fancy and speculation in the 'pop' publications. The layperson, those who have not shared this

training which allows us to talk to each other in our formal shorthand, has naturally misunderstood our emphasis and believes we are saying that these peripheral causes are the strong deciding factors in a criminal trial; most probably do not believe we are correct, they just believe we are claiming it.

Generalizations should be limited to occasions such as those which provided the data (Campbell & Stanley, 1963), but jury research almost demands justification for its very existence from social scientists who seem to be straying from their own professional fields. Jury researchers who claim isomorphic relationships between their studies and the real world quite often do so out of ignorance of actual trial conditions (Elwork & Sales, 1974); yet, justification for their work is often furnished by words of advice to professionals who are practicing in what to them are unfamiliar, highly specialized legal disciplines. In general, individual effects are tested through the experimental control of other perhaps more important variables. We equalize. Through random assignment the biasing effects, known and unknown, are minimized. But social scientists invariably fail to stress this ceteris paribus clause, thus misleading laymen, who are then subjected to unrealistic generalizations to a world where things are rarely equal. In the real world nothing seems to be held equal; each trial is a unique event with a unique blend of ingredients.

Scientific Jury Selection

In this section we review studies of "scientific jury selection." These accounts of social scientists successfully intervening in courtroom procedure have promoted the image of 'jury packing' psychologists using their science to alter trial outcomes. We then look at some critical commentary on jury selection, and examine some recent experimental replications.

The Selectors

In an article in Psychology Today which probably helped initiate the avalanche of jury research in the past few years, Shulman, Shaver, Colman, Emrich, and Christie (1973) report an attempt to select a fair jury in the Harrisburg trial for conspiracy. The social scientists conducted a survey of voters in the area and then helped the defendants and their lawyers rate a panel of 46 potential jurors on their estimated favorability for the defense. Survey results were modified by defendant and lawyer intuitions, personal information, psychological estimates of jury behavior, and the performance of potentials in voir dire. On a one to five scale(one being preferred), eight jurors were rated one, five rated two, 15 rated as three, and 18 as four and five. The prosecution challenged six of the eight best potentials. It seems that the defense, with its huge team of advisors, and the prosecution, working normally, had a similar profile of the favorably disposed juror:

- under 30; the younger the better,
- blacks,
- counter-culture way of life,
- opposition to Vietnam war,
- having a close male relative near draft age.

The trial resulted in a hung jury with ten voting for acquittal and two for conviction. The two holding out for conviction against the overwhelming majority were original second choices.

Shulman et al. commented that it is impossible to predict the behavior of an individual from group data with any confidence, and that individual differences and interpersonal problems which arise as an ambiguous trial unfolds make the task more difficult. The authors further noted that

> most of the jurors took their role as fair decision makers more seriously than we anticipated. Most of them tried earnestly to consider the defendants "innocent until proven guilty." It seems wise to consider a person's conception of the role of juror as a distinct characteristic when assessing jurors (P. 83).

Christie (1976) presents a systematic history of his experience as a jury selector. One of the originators of scientific jury selection, Christie had worked on many of the successful cases which have given scientific jury selection its reputation. Social scientists and attorneys receive different training with differing emphases and Christie believes that while "most criminal lawyers are not interested in jury selection;" attorneys have an enormous advantage over social scientists because of

the experience they acquire in their ordinary day to day work. "When they are interested in problems of jury selection, they can develop an uncanny ability to psych out potentially good and bad jurors" (p.267). However, the relative efficiency of attorneys and social scientists is still an unanswered question. Christie reports that in the Camden 17 trial, 15 of the defendants participated in ratings of the selected jurors. Post trial comparisons of ratings and performance revealed an exactly chance distribution of predictions.

Christie claims that, usually, all the information available, stereotypic and personal, agree, and the very favorable and unfavorable jurors are struck by each side. He says

> that the central problem in jury selection is the careful evaluation of those who are so biased that they cannot be fair and impartial jurors (p. 8).

After rating the individual jurors, the jury selectors try to estimate the probable juror interactions: friendships, cliques, influential leaders, status considerations. Quite often not unfavorable jurors must be eliminated in order to facilitate these jury interactions. For example, in a trial when a very favorable woman of high status was selected, high status men were struck in order so that they might not interfere with her effectiveness. Christie has said (personal communication) that authoritarianism is the single most important predictor of conviction proneness; yet the jury selectors in this same case purposely chose two men they felt were highly authoritarian because experimental evidence

predicts that authoritarians are more easily swayed by group pressure: the combination of low status and high authoritarian seems particularly suited for proposed followers. In general, the methods used by social scientists and lawyers are similar. But the social scentists make the assumptions of the selectors explicit and then attempt to verify these assumptions with collected data such as they would with any research project. Post trial verification is careful and used to generate better hypotheses in following cases. In other words, there is an attempt on the part of the social scientits to systematically arrange their experiences so that future projects can use improved assumptions. Christie stresses the use of many different procedures and the importance of not depending on any single one.

The jury selection techniques were originally developed by Jay Shulman to assist anti-war radicals being persecuted by a government trying to stifle growing dissent concerning the Vietnam War; but, as Etzioni has noted, once formulated the techniques are available to all.

In 1974 John Mitchell and Maurice Stans went to trial in Federal Court accused of conspiricy to impede a governmental investigation. This trial was one outgrowth of the Watergate scandals during the Nixon administration and received vast media coverage: two wealthy insiders (both were members of President Nixon's Cabinet) being prosecuted for exerting improper influence. Here was the case of wealthy men buying techniques designed to protect the powerless from excessive

government zeal. Drawing on court records, newspaper accounts, and their previous research, Zeisel and Diamond (1975) have furnished a penetrating analysis of the Mitchell-Stans voir dire with an evaluation of the effectiveness of scientific jury selection in winning acquittal when the defendants were apparently guilty to most spectators.

The defense was allowed twenty-three peremptory challenges, and the prosecution eleven, to be exercised against a final jury panel of fifty-two prospective jurors. The defense used all its challenges; the prosecution, only six.

With the defense challenges guided by a survey conducted by a private research firm, the jury compostion changed from the complete panel where 46% had some college education, to eight percent college educated in the final selected jury. From a jurypool where 32% were well informed about Watergate to a jury of only eight percent well informed. From 46% of the jurypool who read the NY Times or New York Post to 33% of the final jury who read these newspapers.

Even so, the initial jury vote was eight to four for conviction. "The likelihood that the defendant will be acquitted [after such a vote] is about five percent, or one in twenty" (p. 162).

One of the jurors fell ill and was replaced by an alternate; a vice-president of the First National City Bank, and a Republican with strongly held conservative opinions. This alternate juror was able to influence the rest of the jurors to acquit the defendants despite what

seemed like overwhelming evidence. It seems the high-status person was able to persuade the low status jury to his view.

Scientific jury selection resulted in an eight to four vote for conviction and can be said to have failed; but it also resulted in a jury easily manipulated by a high status person with strong opinions; in this respect it can be said to have succeeded.

In the same year that Shulman et al. reported their careful account of the Berrigan trial, William Sage (1973) presented an emotional account of the 'inner workings' of the team that won an acquittal for Angela Davis, a black communist brought to trial in California by the Federal Government in a media celebrated case. Sage claims that "every trial is fundamentally an emotional event," but

> fortunately for Angela Davis, her performance turned out to be flawless, played to the hearts by way of the minds of a white, middle-class jury, psychologically tailored to be a captive audience.

At no point does Sage mention the charges brought against Davis; the evidence or lack of evidence is irrelevant to the story and, one feels, to the trial. As Sage tells the tale Davis was aquitted because she was beautiful, intelligent, and presented a good story -- her guilt or innocence had nothing to contribute to the outcome of the trial -- and, of course, she had psychology, in the persons of five <u>black</u> psychologists, on her side. The four lawyers who helped Davis act in her own defense tried to hire a <u>white</u> psychologist whose specialty, according to Sage, was in packing juries for the wealthy; a Dr.

Bryant, who "like most specialists turned out to be a man who knows his own worth," refused to contribute his services for free.

As Sage tells the tale these five black psychologists, in effect, packed the jury with fair minded white persons (it took some doing), and orchestrated a well-presented stage play for their consideration.

Articles such as this written by an assistant editor of a "pop" psychology magazine celebrating the magic of psychology in overcoming the intrinsic biases of the 'system' have helped to foster the image of the 'jury packing' abilities of psychologists.

Critique

As noted above, the basic concern of jury researchers is that non-evidential factors affect jury decisions; but it is almost inconceivable that this would not be the case. The use of a jury rather than an individual judge implicitly suggests what we all know to be true -- human beings are biased in different ways. Not only are jurors, as individual human beings, no less biased than their assessors, but there is some question as to just how much bias a juror should have (Moore, 1973). In fact, "the counterbalancing of various biases is critical to the accurate appliction of the common sense of the community to the facts of any given case" (Ballew, p. 10).

The central question regards the effect of this individual bias on jury verdicts. We cannot discount the importance of the individual

juror; after all, there is strong evidence that the pre-deliberation majority generally prevails (Kalven & Zeisel, 1966). The question naturally revolves around biased jurors. Do they counterbalance? Once again we have an expression of belief and hope which, theoretically at least, is capable of empirical testing. The central issue is the fairness of jury decisions. Are they based on the evidence as the law and tradition demand, or do they result from the staging and prejudices as some interpreters of jury research maintain (e.g. Sage, 1973)?

In an essay addressed primarily to lawyers (1976a) and another addressed to the interested public (1976b), Saks disputes the efficacy of scientific jury selection. He claims that most of the variance in jury verdicts is accounted for by differences in evidence. The early jury consultants helped lawyers to win their cases because the charge of conspiracy is notoriously difficult to prove. Large numbers of professionals had been helping the defense teams in these political cases. Perhaps their successes were due to the number and effort rather than the then newly developed techniques.

> No lawyer would be harming his client by taking advantage of scientific jury selection. But if he wanted to have an even greater influence over the outcome of the trial, he ought to hire social scientists to help build and structure the evidence to be presented (p. 22).

Saks believes that social science jury selection techniques are not as effective as knowledgeable lawyers. He also feels that Etzioni and others who are concerned about the intrusion of science into selection

have been unrealistic in their assumption of social science efficacy. Furthermore, Saks disagrees with Etzioni's ethical objections. He argues that the complaints about the use of science to do what lawyers have always tried to do result from a success that seems around the corner, just as behavior modification arouses complaints simply because it is being used successfully to do what society always wanted to do. "Apparently, many of our society's goals are acceptable only as long as they cannot be achieved" (p. 11).

Richard Berk has been involved in using scientific selection techniques in several jury trials, but remains skeptical about their usefulness. In 1969 the Chicago police raided the Black Panther headquarters killing their leader and one other and wounding several members. Not receiving 'justice' in the criminal courts, the Black Panther survivors instituted a civil case for damages along with the assistance of several interested social scientists. Berk (1976) discusses some of the problems faced by the social scientists who have come together in the type of case which developed the jury selection techniques. There are the touchy relationships of volunteers with different motives, the different degrees of research sophistication, and the intrinsic difficulties with survey and interview data. The scientists were involved for different reasons:

1. Those who were radical leftists, did not believe a fair trial was possible under this system, and were hoping to 'pack' the jury;

2. Those who believed that it was some powerful groups that precluded the fair trial rather than the system itself;
3. And those who believed in the system and notions of 'justice and fair play' and who were trying to present the best possible case.

And, of course, just as there were different motives for joining the 'team' there were differing expectations as to the effectiveness of the jury selection technique. As a methodologist Berk is doubtful but says he is willing to be convinced. He feels that the key to jury selection is prediction. Causal explanations are useful only if they enable the researcher to find a better predictor: "Understanding per se is an unnecessary luxury." He says that trivial differences are probably the result of sampling error and are not helpful; jury selectors need huge differences that can be easily identified in the courtroom. The primary function of the survey analysis is to legitimize the intuitive insights of the attorneys that are correct and to question those which do not stand up well to scrutiny. The data are crude, and

> to slap fancy statistical techniques on crude data only serves to obscure errors. Statisticians are not alchemists despite the wishes of many social scientists. Complicated statistical techniques will not salvage weak data and can give the impression of more rigor than actually exists.

In a more general vein, Berk, Hennessy, and Swan (1977) evaluate the status of "scientific" jury selection noting that the selectors themselves claim little science and much hard work while their supporters

are impressed with the obvious results and make the claims which have prompted Etzioni and others to warn researchers of future dangers to the jury system. They claim that, despite the contrary allegations, attorneys have always tried to select jurors who have predispositions in their favor. The intervention of social scientists is a promising but unproven innovation. It is the statistical trappings that bestow the scientific label on their efforts, and evaluation techniques regarding the intervention as treatments are required to estimate their effectiveness. In jury research we find all the problems that inhere in survey data itself. Category answers to attitude questions assume that the researchers know the proper range of answers and the questions that should be asked. The survey questions usually define three types of dependent variables:

1. the facts of the particular case being tried,
2. the respondents' more general feelings about relevant issues,
3. and deep-seated personality characteristics that might help predict predispositions.

There is a practical problem that arises in this type of research with respect to the sampling procedure. Do we take a random sample of the jurypool population, or do we stratify in order to assure large enough subsamples of unusual categories that might prove important for selection purposes? For example, black professionals might be rare in the population and will turn up in numbers too small to properly

analyze, yet we might very well want to choose one or two for our jury.

Problems abound with factor analysis, the statistical tool of choice in jury surveys. There are numerous dependent variables and "the weight of the data reduction typically falls on some version of factor analysis" (p. 150). We can factor analyze our data, but how do we know which are the dimensions that predict decisions in our particular case?

Selectors have used the Automatic Interaction Detector (AID) (Sonquist, Baker, & Morgan, 1973) and multiple regression analysis (MRA) in their analyses. AID is a stepwise computer program for use with ordinal and nominal independent variables. The program evaluates all possible dichotimizations of the independent variables to find that variable which explains the most variance. The 'interaction' in the title refers to the program's exclusive use of interactions with the initial main effect in subsequent equations. Berk et al. report that they compared an AID and an MRA of randomly generated data. The AID accounted for 14% of the variance while the MRA accounted for an understandable two percent; therefore it would seem that AID capitalizes on sampling error to an extent which makes it unreliable. Berk et al. suggest that AID not be used in this type of real world research.

With regard to multiple regression, "once a model ... has been developed, the goal is to use the regression coefficients and intercept

to generate a predicted score for any potential juror"(p. 152). Berk et. al. suggest that even the best results have a wide range of error generating huge confidence intervals; perhaps as much as 40 on a scale of 100. Only very gross differences count in data such as these since a predicted score of 70 could vary from 50 to 90. Then again prediction errors do not center around zero in this type of procedure. The jurors are picked on the same criteria and the errors tend to be correlated and in the same direction. Berk et. al. point to some serious problems and then wind up the article vindicating scientific jury selection with the claim that it surely will not do any worse than most attorneys and might very well do better.

Experimental Studies

Recently there have been several studies investigating the efficacy of jury selection strategies. Horowitz (note 9) conducted an experiment comparing the effectiveness of attorneys' conventional jury selection methods with systematic social science methods. He found that neither method was intrinsically superior: when the relationships posited by systematic jury research were strong, the social science methods worked; otherwise the attorney selections were more effective.

Zeisel and Diamond (1976) report a shadow jury experiment where juries of rejected jurors and others composed of randomly selected jurors viewed and judged actual criminal jury trials. Comparing the

verdicts delivered by the three juries concerned with each trial, Zeisel and Diamond concluded that prosecutor challenges did not affect verdicts significantly while defense challenges did. Despite the inability of attorneys to articulate their selection criteria they were able to use their intuition and experience in their clients' interest.

More germaine to our concerns, Penrod (note 4) reports an experiment in his dissertation which, despite a number of serious research problems, won a major award for social science research. As noted earlier, Mills and Bohanon found that almost all posited non-evidential factors seemed to systematically affect real juror verdicts; in this experiment, purporting to replicate and test the social science jury selection model, Penrod found otherwise.

Penrod used 367 members of actual jurypools with an average age of 43 who had volunteered at their induction. These jurors were contacted during their second week of service in order to guarantee a body of experienced jurors since "the study was not an exact duplication of real trial procedures" (p. 122); and, as it turned out, 90% had served in at least one case while 63% had served in at least one criminal trial. Subjects were then given written summaries of three trials (2 criminal, and 1 civil), and listened to audiotape recreations of the three trials. After each taped trial each juror rendered an individual verdict. Upon completion of the last trial the jurors supplied demographic information and an ad hoc twelve (12) question scale "assessed jurors' attitudes on a range of issues."

To test the social science model which Penrod feels he has replicated,

> the criterion variable would be something other than the jurors' actual votes. The most plausible candidate would be to perform a factor analysis on the attitudinal measures employed in the study and select the factor that most plausibly could be interpreted to reflect authoritarianism or dogmatism. Individual juror factor scores would be computed on this factor and employed as the criterion variable in regression and AID analysis.

One could then do a regression analysis predicting factor scores from the demographic data. The social science model that Penrod describes views the estimated factor score as an indicator of conviction proneness and a predictor of potential juror voting behavior. The test of the model would be a comparison between predicted factorscores and actual juror votes.

As it turned out, the scale would not factor into any worthwhile pattern and "efforts to construct and test a social science model were abandoned." Penrod then tested what he termed "the optimal empirical model" using the four sets of verdict preferences as dependent variables in stepwise MRA and AID procedures. An idiosyncratic AID effect disappeared when the sample was split for retesting; and the variance explained using stepwise regression was negligible despite that procedures' well known capitalization on chance in exploratory studies (Cohen & Cohen, 1975).

The question addressed by this experiment was:

> Given the best possible information (jurors' verdict preferences and plentiful information about the jurors) can we or

> other social scientists achieve a usable degree of prediction accuracy? (p. 120).

Penrod supplies the following answer:

> The jury selection process remains an art, not a science. The somewhat paranoid critics of scientific jury selection have not been vindicated. Indeed, the more circumspect practitioners of scientific jury selection are placed on even weaker ground than before.

The programmatic jury research of Hastie and his colleagues at Harvard, of which Penrod's thesis is part, is an ambitious attempt to make explicit and test assumptions and theories in the field of jury research. However, the research reported by Penrod surely does not support his conclusions. In no way can Penrod be said to have given scientific jury selection a fair test. Most of the criticisms of mock jury research detailed above can be applied here. While these mock juror subjects were chosen from actual jurypools in a regular courtroom, they did not undergo voir dire; they had to judge 3 trials and fill out questionnaires in 4 hours; knew beforhand they would have to explain and justify their verdict; did not deliberate their verdicts which were individual rather than group decisions; were experienced despite jury selectors' wariness of the experienced juror; judged both civil and criminal cases despite different rules of evidence and standards of proof (criminal lawyers often hesitate to empanel jurors who have sat on civil trials in their present term); and, finally, knew that their decisions did not have any real consequences.

A criticism more specific to this study concerns the factor analysis which didn't work. Jury selectors almost always use factor analysis and they almost always factor. Can this study be considered a fair replication of the social science model when the basic methodology is not replicated? Why use an unproved ad hoc scale of 12 attitudinal items when there are published versions of past scales used in actual jury selections with estimated reliabilities and factor structures? We can imagine no better test than the use of the jury selectors' own instrumentation.

Penrod is not unaware of these weaknesses in design and dangers to external validity; and, in fact, goes into lengthy explanation and justification of the experimental procedures. He even goes so far as to say:

> Though it is a rather backhanded endorsement of the simulation methods used in the present study, the fact that Wilson and Donnerstein obtained offender attractiveness effects only in the hypothetical consequences condition suggests the possibility that a simulation study has a better chance of detecting individual differences in actual jurors that might predict juror verdicts than would a study using real consequence jurors (if such a study were possible) (pp. 131 & 132).

Perhaps it is true that scientific jury selection does not alter probabilities of conviction and that critics of jury research are really "paranoid," but Penrod's study does not add to the perception of that truth because of its obvious external validity problems. However, the Penrod study represents a significant advance in the scientific jury

selection because, for the first time, the survey portion of the scientific jury selection model is specified and applied to empirical data in full view of interested readers.

The Research Questions

Attorneys use varying criteria when choosing jurors with varying success. These criteria are usually expressed as simple relationships between pairs of variables uncomplicated by the consideration of other related variables. Much of jury research has been initiated by the natural desire to find simple, visible or otherwise identifiable variables that would help in juror selection. This menu of simple bivariate effects is subsumed under the scientific jury selection model and is utilized in the survey of the potential jury population. It is this data collection and analyses applying social science knowledge of sampling procedures, questionnaire construction, and powerful statistical techniques which lend the selection procedure its "scientific" reputation among research professionals as well as laypersons (Berk, 1977). Once analyzed, the results are used by attorneys and social scientists in the usual intuitive manner subject to all the idiosyncracies of differential expertise and application (Christie, 1976). But the scientific portion of the process should be replicable and reliable; and it is the survey and its analysis which concerns us here.

As we have seen, jury research is plagued by a mixture of research techniques and the questionable use of students as subjects. Efforts to control stimuli (the trials) and extraneous influence through controlled experiments have led to artificial situations of dubious generalizability.

Despite Saks' (1976) dismissal of scientific jury selection as ineffective and the failure of the scientific jury selection model to predict mock juror behavior in Penrod's (note 4) study, the use of scientific jury selection in the United States is increasing (Hunt, 1982). Because of the difficulties outlined earlier, there has never been a good test of the scientific jury selection procedure. The present study simulates and tests the 'scientific' component of the usual jury selection process by employing that portion of our data usually collected by jury selectors, assigning selection scores to the sample in the same manner as employed by the jury selectors, and then comparing these selection scores to the conviction scores as computed from the actual reported behavior of our juror respondent sample. We also test assumptions implicit in attempting to build a jury selection model.

Our questionnaire elicits more information than is ever available to jury selectors; not only are real jurors being canvassed, but their recently deliberated verdicts are noted and scored. If valid juror selection is possible, using all this information should enable us to construct a cogent model for jury selection. This model which uses the maximum information -- the Ideal Model -- is then cross-validated. If

the usual selection model itself fails to reliably predict juror behavior, a test of the Ideal Model will at least inform us whether some model can possibly do the job, and whether the assumptions, explicitly specified, do in fact hold.

There, then, are three general questions which are explored in this study:

1. Does the existing scientific jury selection procedure (as described by Penrod, note 4), predict juror verdicts?
2. Can we create an Ideal Model that reliably predicts juror verdicts?
3. Are the underlying scientific jury selection assumptions of strong and identifiable relationships between background variables, conviction proneness, and juror verdicts justified?

METHOD

Subjects

In order to understand the data collection, the reader must first understand the juror selection process in the Federal Court, Eastern District of New York -- the research setting.

As mandated by the Jury Selection and Service Act of 1968, every four years a jury wheel is randomly selected by computer from the list of registered voters in this jurisdiction: the counties of Kings, Queens, Richmond, Nassau, and Suffolk.

Every two weeks thereafter ,according to the scheduling of cases, a varying number of potential jurors are randomly selected from this jury wheel and summoned to appear for service. Some never answer the summons, others are excused for various reasons, the balance, usually less than 50% of those summoned (Simon, 1973), constitute the available juror pool, the venirepersons. The venirepersons are legally obligated to serve until dismissed, but unless actually serving on a trial the length of each venireperson's availability is customarily limited to two weeks.

The venireperson is paid $20 for each day spent in court. In order to reduce expenditures, the court does not require the venirepersons to come to court each day. The venirepersons call the court each evening and a recorded message informs them if their presence is required the

following day. In this way the jury clerk can have just the number of venirepersons necessary to fulfill the court needs each day. If venire-persons are selected as jurors early in their service, it is very likely that after the trial they might never be at court again and be dismissed from their obligations through the phone message; in fact, this is ordinarily the case; most jurors who serve on a jury in this jurisdiction are either dismissed from service immediately after a trial, or by phone, limiting their availabilily to jury researchers.

In order to secure permission for this research from the judges of this court, this study was required to:

1. guarantee the anonymity of the juror,
2. minimize interference with the usual court procedure,
3. refrain from any contact with the jurors until their service was complete.

In this court a jury may deliver its verdict any time of the day or night. When the jury's deliberations are over and the jurors agree on a verdict (or agree to disagree -- the hung jury), the foreperson notifies the judge who then assembles the principals in the courtroom to hear the verdict. After the verdict is delivered, the judge will thank the jury and dismiss them. What follows next depends on the time of day and the juror's length of service.

If the jurors are dismissed outside of the normal court hours (9:00 A.M. to 5:00 P.M.) those who have served a short time (the criterion

varies depending on the availability and size of the jurypool) are asked to report to the jury room the next day and are then re-entered into the jurypool. Those who have served long enough do not go back to the court at all. Neither of these groups were available to this study.

If the jurors are dismissed during the regular court day, they then report to the jury room. Those who have served for a short time have their names re-entered and are again available to service (and unavailable for this study). Those who have served a sufficient length of time are excused from further service. At this point the jury clerk handed the questionnaire (contained in pre-addressed manila return envelopes which require no postage) to the jurors, and asked them to read the contents when they got home. The covering letter was as follows:

> Dear Juror,
>
> The Center for Responsive Psychology, Brooklyn College, with the consent of the United States District Court for the Eastern District of New York is conducting a survey of jurors who are finishing service. Your voluntary assistance is greatly needed. Without your help, this project cannot be completed. This is not a test -- there are no right or wrong answers, and all answers will be kept strictily anonymous and confidential. PLEASE DO NOT PUT YOUR NAME ANYWHERE ON THESE FORMS. Some of our questions might seem a bit personal or irrelevant but the success of our project depends on your answering them frankly and completely so that statistical comparisons can be made. You do not have to answer every question, however to obtain as accurate a measure as possible, it is important that you try to complete the questionnaire. Please put the questionnaire back in the envelope it came in. The postage has already been paid.

> The results of this survey will be published and made available to judges, lawyers, law journals and the public in general. To reserve a copy of the report, just send a postal card to us at the above address.
>
> If you have any questions regarding the survey please call (212) 780-5960.
>
> We thank you in advance for your cooperation which we greatly appreciate.

If the case was a long one and the jury dismissed during the regular hours, the entire jury was given the envelope with the questionnaire; otherwise, as is usually the case, only those excused at that moment by the jury clerk were handed the envelope. Thanks to the cooperation of the personel in the Federal Court, Eastern District, a large sample of jurors were tested. The lengthy time frame required for the selection of subject jurors by the procedures outlined here assured this study of a representative sample of real jurors who had judged real criminal trials free from any unusual influences.

There were 800 questionnaires and envelopes printed. Approximately 100 of these were used as samples or given to judges or other court personnel who requested copies. There were 365 usable protocols returned by respondent jurors representing about 52% of the distributed questionnaires.

The final sample, then, consisted of 365 jurors who had just finished serving on criminal cases in the Federal Court, Eastern District of N.Y., and who returned the questionnaire given to them upon that completion. These 365 jurors were randomly split into 2 samples: an

exploratory sample of 182 (J1), and a cross-validation sample of 183 (J2). Table 2 presents the demographic breakdown of these two groups and that of the random sample selected from the same jurypool as described above on page 15f. We can see that the split halves were similar to each other in all the demographic categories but that there was an under-representation of the less educated in our sample when compared to the random jurypool sample. Such a discrepancy probably arose from differential selection of jurypoolers to be actual jurors during the voir dire, but could possibly have been the result of a differential questionnaire return rate.

TABLE 2

Comparison of "Early" Jurypool With Split Half Samples

Variable	Category	Early[1]	J1[2]	J2[3]
N		318	182	183
Sex	Male	49.7%	43.9%	43.3%
	Female	50.3%	56.1%	56.7%
Race	White	82.7%	87.6%	85.5%
	Black	11.9%	11.3%	11.7%
	Hispanic	3.5%	.6%	2.2%
	Asian	.3%	.6%	-
	Other	1.5%	-	.6%
Education	not HS Grad	16.7%	10.1%	6.7%
	HS Grad	39.4%	32.0%	30.9%
	Some College	24.6%	28.1%	29.2%
	College Grad	9.8%	16.3%	17.4%
	Grad/Prof School	9.5%	13.5%	15.7%
Income	Less than 5m	4.8%	.6%	1.1%
	5m to 9m	10.2%	4.7%	2.3%
	9m to 15m	27.6%	12.8%	15.3%
	15m to 25m	35.5%	43.0%	35.8%
	Greater than 25m	21.8%	38.2%	45.3%
Religion	Catholic	53.5%	41.0%	49.4%
	Jewish	20.1%	24.8%	21.1%
	Protestant	20.7%	22.4%	23.5%
	Other	1.3%	5.0%	4.2%

[1]The "early" jurypool refers to the research reported in Baker (note 1) and reviewed on p15f above.
[2]J1 refers to the exploratory sample.
[3]J2 refers to the validation sample.

The Questionnaire

The preferred approach to the study of real jurors is the study of real jurors who have judged real trials. This study accomplishes this through the use of self-report questionnaires at the jurors' moment of dismissal from service on a criminal trial. These questionnaires are composed of items taken from the actual instruments used in jury selection by the National Jury Project (Kairys, 1975) and refined in past research (Buckhout & Baker, 1977b; Ellison & Buckhout, 1981).

The questionnaire consists of 69 items. The first five items concern the jurors trial experience and serve to eliminate any responses that may result from errors in the distribution of questionnaires.

Items six through nineteen relate specifically to the last trial on which the juror sat.

After two items concerning personal experience with crime, there are 24 items concerning knowledge of, and attitudes to, crime related issues (called the "Q" items below).

Items 46 through 69 request demographic background information (called the "D" items), and are identical to the items used in the previous study on the equality of late and early jurypools reported on page 15f above.

Since the major aim of this study is to reproduce and test a typical jury selection procedure using Factor Analysis and Multiple Regression, these items are derived, primarily, from actual instruments used in jury

selection cases (e.g. Swinton). Almost all of the items, except those referring to the trial itself, have been used and tested at least once by myself or other researchers and/or jury selectors (Ellison & Buckhout, 1981). A few were suggested by judges and attorneys who had collaborated with me prior to this particular study.

The Conviction Score

The major obstacle to meaningful jury research is the trial itself. In mock jury research we can standardize the trials; no such option is open with actual trials, each is a unique event. How then can one obtain comparable measures for dependent variables when a jury survey may yield almost as many trials as subjects?

Jurors are usually asked to vote guilty or not guilty on a first ballot prior to any jury deliberation. This predeliberation vote of the respondent juror and the distribution of that vote within the jury of which he or she is a member was used to help define the dependent variable for the jury selection model.

The jurors who observe a trial and then deliberate together have shared the very same trial experience, viewed the same evidence, heard the same witnesses. Conceptually, the one juror who has shared this trial experience with the other jurors sitting with him or her and agrees with a clear majority of the others, can be thought of as one who is affected more by the evidence of the trial than by his or her proclivity to convict or acquit; if the juror stood alone against all other jurors his or her proclivity can be assumed to be more of a factor than the evidence. Taking the case of the conviction prone jurors, one would expect to find them voting guilty against the majority more often than those who are not conviction prone. This proclivity to convict or acquit can be measured and defined as a function of the predeliberation

ballot itself. Admittedly, this is an imperfect measure: trials heavily evidenced in the same direction as a juror's inclination do not permit that juror's disposition to manifest itself; but then we are dealing with groups and averages. The assigning of an individual score from aggregate data is a problem that all jury selections must deal with; this study is subject to that same limitation, and it is the resulting efficacy that is being evaluated.

What is needed is a measure that has the nice property of heavily weighting those who vote against the majority and lightly weighting those who vote with the majority, diminishing to zero when the vote is unanimous. In the present study the Conviction Score (CS) is defined as the positive value of the natural log of the proportion of the 12 jurors voting as the respondent when the respondent votes guilty; and a negative value of the log proportion when a respondent votes not guilty.

Since the natural log of a proportion is always negative and we want the score to reflect conviction proneness,

$$\begin{aligned} CS &= (-1)\ln(P) && \text{if respondent votes guilty, and} \\ &= \ln(P) && \text{if respondent votes not guilty.} \end{aligned}$$

Where CS is the conviction score, ln is the natural log, P is the proportion of the 12 jurors voting as respondent; e.g. CS = .69 = (-1)ln(6/12) if respondent is 1 of 6 voting guilty.

The CS varies from 2.48 to -2.48. A score of 2.48 reflects the highest level of conviction proneness and occurs when 11 jurors vote not guilty or neutral and the respondent juror votes guilty. A score of -2.48 reflects the lowest level of conviction proneness and occurs when 11 jurors vote guilty or neutral and the respondent votes not guilty. The values between 2.48 and -2.48 reflect intermediate levels of conviction proneness.

Most predeliberation ballots find some individual jurors who do not have a firm opinion as to the guilt of the defendant. Voting neutral when the evidence is overwhelming is fully as effective as casting a contrary vote; a neutral vote indicates a disposition and in fact can be considered as indicative of a juror's conviction proneness as one which makes a definite stand. Therefore, the neutral vote is weighted

$$
\begin{aligned}
CS &= (-1)\ln(1-Pm) && \text{if >5 jurors vote not guilty,} \\
&= \ln(1-Pm) && \text{if >5 jurors vote guilty,} \\
&= 0 && \text{otherwise.}
\end{aligned}
$$

Where: CS is the conviction score, ln is the natural log, Pm is the proportion of the majority jurors, and the score is therefore opposite in sign to the majority vote of at least six: e.g. CS = -.88 = ln(1-7/12) when 7 jurors vote guilty. If no such preponderance manifests itself on the first ballot, then the neutral vote is scored zero.

Jurors were assigned conviction scores as a function of their predeliberation votes as detailed above. This conviction score reflects

a jurors conviction proneness and serves as a dependent variable representing the juror's actual behavior. More specifically, Table 3 shows the scores that were assigned to a juror voting as indicated in the columns when he or she was a member of a jury whose distribution of predeliberation votes is as recorded in the rows.

TABLE 3

Conviction Scores for the Respondent Juror
as a Function of
Some Typical Voting Distributions[1]

When the Respondent Votes[2]					
guilty	(CS)	neutral	(CS)	not guilty	(CS)
0	-	0	-	12	(0.00)
2	(1.80)	3	(.88)	7	(-.54)
2	(1.80)	2	(1.10)	8	(-.41)
5	(.88)	3	(.00)	4	(-1.10)
7	(.54)	3	(-.88)	2	(-1.80)
7	(.54)	1	(-.88)	4	(-1.10)
11	(.09)	1	(-2.48)	0	(.00)
12	(0.00)	0	-	0	-

[1]each row represents a single trial.
[2]CS assigned to respondent voting as indicated.

The Jury Selection Model

Figure 1 depicts a model of the scientific jury selection procedure as conceived and examined in this study. Table 4 contains the list of variables collected by questionnaire, and the derived variables embodied in the model diagram. Comparing the notation in the model with that of the table we see that:

TABLE 4

List of Variables used in analyses

Dependent Variables

Proneness Index: PI=f(Q)
Conviction Score: CS=f(JR,JY)

Independent Variables

Pertaining to Respondent

Variable	Q[1]	Variable label
Resvctm	(20)	Was respondent the victim of a crime
Famvctm	(21)	Anyone in family the victim of a crime
Orgsize	(46)	Number of employees in workplace
Supervs	(47)	Is respondent in a supervisory postion
Numsuper	(48)	Number of persons supervised
Newsp	(49)	What newspaper read regularly
Lives	(50)	County of residence
Rentown	(51)	Type of residence
Resage	(52)	Respondent's age
Marital	(53)	Marital status
Educatn	(54)	Schooling completed
Sex	(55)	Gender
Resocc	(56)	Respondent's occupation
Religion	(58)	Respondent's religion
Gochurch	(59)	How often attends religious services
Ancestor	(60)	Country of origin of family
Belong	(61)	Belong to a social/civic orginization
Faminc	(62)	Total family income
Party	(63)	Political party supported
Resethn	(68)	Respondent's race/ethnicity
Lang	(69)	Is English respondent's native language

Pertaining to Judged Trial

Defsex	(10)	Defendant's sex
Defage	(11)	Defendant's age
Defethn	(12)	Defendant's race/ethnicity
Defocc	(13)	Defendant's occupation
Charge	(6)	Defendant's criminal charge
Deftest	(7)	Did defendant testify
Lawtest	(8)	Did law enforcement agent(s) testify

[1]Q refers to the Question number in Appendix A.

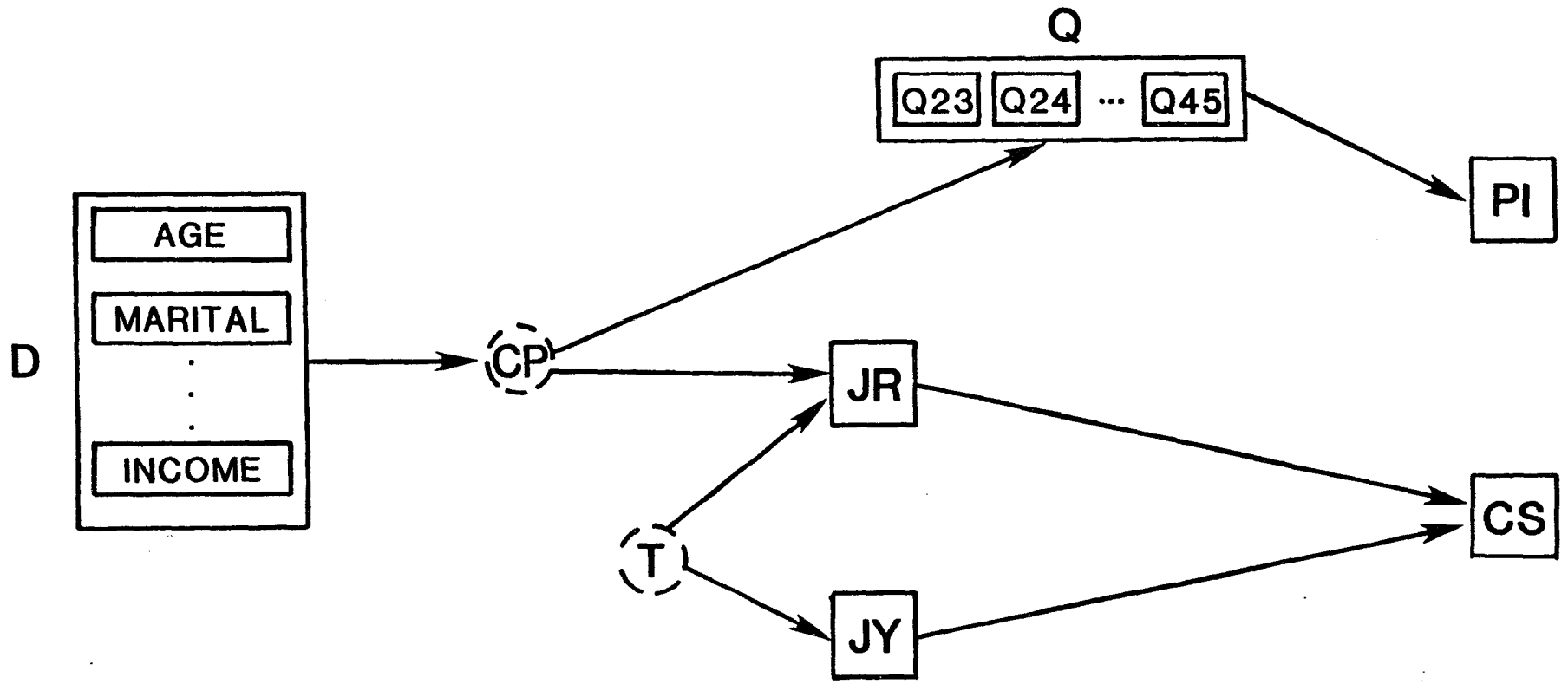

Figure 1. THE JURY SELECTION MODEL

- D represents the large number of demographic and background variables gathered by the questionning of the over 350 jurors.
- Q represents the 24 attitude and legal knowledge questions asked of the juror respondents -- questions 22 through 45 of the questionnaire.
- JY is the jury predeliberation vote at the trial.
- JR is the juror respondent's vote.
- PI is the proneness index, F(Q), the first factor of an unrotated factor analysis of Q. Derived as explained below.
- CS=f(JR, JY), the conviction score as computed above -- the juror vote weighted by the ambiguity of the trial.

Furthermore, there are two (2) theoretical variables not contained in Table 4 but which are important features of Figure 1

- CP is conviction proneness, defined as the tendency of an individual to consider a defendant guilty. It is a theoretical construct according to which those with higher conviction proneness would be more likely to consider a defendant guilty than those with lower conviction proneness while viewing the same trial.
- T represents the juror's perception of the experience of the particular viewed trial: the type of criminal charge, the competence of the opposing attorneys, the credibility of the various witnesses, and all the other measureable and unmeasureable variables that go into weaving the unique fabric of each trial.

The Field Model

In this study, the existing scientific jury selection procedure, as reported by various researchers (Berk, 1977; Ellison & Buckhout, 1981; Shulman et al., 1973) and described in detail by Penrod (note 4), is called the Field Model. In terms of the model in Figure 1, it is an attempt to estimate the latent conviction proneness (P) from background variables (D) and measures of attitudes (Q).

First, a survey is conducted among the potential jurypool -- usually registered voters in the relevant jurisdiction. All the questions asked in this study, except those related specifically to the trial viewed by the juror, have been used in such a survey. The survey is then analyzed to statistically estimate the conviction proneness of different identifiable segments of the sample. A stand-in variable for conviction proneness, the proneness index (PI), is developed through the factor analysis of attitude questions (Q). The proneness index (PI) is regarded as a proxy for the latent conviction proneness (CP) of the jurors and regressed upon the various background variables (D) to find the most important predictors of conviction proneness (CP) which the field model assumes is a good predictor of actual juror behavior.

Having generated a regression equation, it is then usually applied in court in a most informal manner by the jury selectors who use the results to advise the attorneys and confirm (or disconfirm) the attorneys' educated guesses.

In this study the following Field Model procedures were implemented:

- An unrotated principal factor analysis (Rummel, 1970) of the 23 attitude items (Q) was performed. Previous factor analyses of similar sets of items had isolated a conviction proneness/authoritarian main factor accounting for most of the common variance. The resulting factor score (the proneness index -- PI) was assigned to each member in our sample and used as a conviction proneness (CP) stand-in variable.
- The proneness index (PI) was then regressed on the background variables (D) in the questionnaire using forward stepwise regression and other exploratory procedures. Categorical variables of more than two categories (e.g. criminal charge, race, occupation) were coded into sets of dichotomous dummy variables (Cohen & Cohen, 1975).

The derived regression equation can now be used in the Field Model to predict the proneness index (PI), the proxy for conviction proneness (CP). Jury selectors use this prediction equation in a most general sense (Berk et al., 1977; Ellison & Buckhout, 1981). Previous analyses of similar sets of variables in this jurisdiction have estimated a significantly higher conviction proneness for older catholic jurors who go to church frequently. This is the profile of a juror the defense attorney does not want (in the absence of better information). A defense

attorney could use all or any part or any combination of parts of such a profile: rather younger than older, rather non-catholic, rather less frequent church attendance, rather younger and non-catholic, etc., etc. Applying the prediction equation to each juror in this study yields a specific estimated proneness index ($\hat{pi}$). We also have (as jury selectors do not) conviction scores (CS) as computed from their actual behavior.

- Therefore, in this study we were able to specifically test the efficacy of the Field Model through a simple correlation of the predicted proneness index and conviction scores: $r(\hat{pi}.CS)$. A significant correlation (at the .05 level) would be considered a confirmation of the model.

Assumptions of the Field Model. In Figure 1 we have a representation of the jury selection model. Different demographic groups have differing degrees of conviction proneness (CP). In other words, under similar trial conditions different groups are more likely to consider defendants guilty and to bring in guilty verdicts. The Field Model estimates the relationship between the background variables (D) and the proneness index (PI) in a random sample of the potential juror population, and then applies the resulting regression equation to estimate the conviction proneness (CP) of the members of the jurypool selected for the trial (a different sample) under the assumption of PI and CP equivalence. There is considerable looseness between this underlying theo-

retical model and the actual measures used to specify these theoretical constructs. Background variables (D) can be measured fairly accurately, but researchers are often not sure which variables are the important ones, and they are not sure which categories are most meaningful; this study uses those which have proven most useful in past jury selection cases. The problems of representing conviction proneness (CP), a theoretical construct of attitude, is not so easily resolved. As we can see from Figure 1, the proneness index (PI) is the measure of the posited conviction proneness (CP) constructed from the 23 attitude questions; while the conviction score(CS) is derived from the actual behavior of the juror sample and is conceived as conviction proneness (CP) modified by the juror's perception of the viewed trial (T).

The upper part of Figure 1 represents the data of the Field Model; the lower part represents the additional information collected in this study. Specific legal attitude and knowledge questions(Q) previously used in past jury selections have been chosen to unearth conviction proneness (CP). These are imperfect measures, of course, which contain the additional sources of error one picks up from all questionnaires. Since the juror verdict one would expect is modified by actual trial conditions, favorableness of evidence, competence of attorneys, and the composition of the jury itself, the juror vote as an indicator of the theoretical likelihood is necessarily an imperfect measure modified by

the context of the viewed trial. The demographic and other background variables (D) can be said to be measured without error despite the uncertainty concerning the relevant categories. The factor score (PI) which acts as a stand-in-variable in the Field Model is a doubly impaired indicator burdened with the uncertainties of exploratory factor analysis (Harris, 1975) as well as the normal error to be expected in all measures. The conviction score (CS) is an attempt to estimate the theoretical conviction proneness (CP) of an individual while considering the context in which the juror acts. It is not to be regarded as a number which represents an absolute quantity so much as a number which can be used to compare individuals on their conviction proneness (CP) in an attempt to estimate potential from past behavior. It too is brimming with error.

The Field Model equates these two error laden derived measures: the conviction score (CS) and the proneness index (PI); it estimates the relationship between the background variables (D) and the proneness index (PI), and further makes the strong assumption that the proneness index (PI) is itself strongly related to the conviction score (CS); it is further assumed that this relationship holds over different samples.

In order for such an enterprise to succeed it is apparent that the posited relationships not only have to exist, but must be strong enough to overcome the errors at every stage of the analysis: random errors

do not correlate, so that even a strong underlying relationship might produce a low correlation between error laden observed measures. Therefore, as can be readily seen, the Field Model process of estimating conviction proneness (CP) by constructing a regression equation which predicts the factor score (PI) from the background variables (D), and then assuming that the predicted values themselves can predict juror verdicts (JR) does not have a strong likelihood of success even if the posited relationships do exist.

The Ideal Model

It is possible, of course, that the Field Model is unworkable; that is, the analysis does not reveal any connections strong enough to be detected because of the weight of errors inherent in the process. However, in any case, a model using all the information available to this study is created and cross-validated.

The model is called ideal because this study can regress actual recorded past behavior on a number of measured variables. If a set of predictors can not be reliably determined with the wealth of information at this study's disposal, it is difficult to imagine how it could ever be done in the field. Failure to create a cross-validated model can be considered as a rejection of a most basic scientific jury selection assumption -- that we can determine how demographic groups differ in their conviction proneness -- and would therefore be considered as a rejection of the scientific jury selection model as presently conceived.

RESULTS

Data Analysis Procedure

Supplying answers to the research questions necessitated an analysis in two distinct stages which are explained in detail in the sections below. The first stage involved exploratory procedures needed to derive the "best" predictor equation possible using the information available in the exploratory sample (J1). The second stage involved using the derived predictor equations to assign a score to each juror in the validation sample (J2) and comparing that predicted score to the actual score recorded for each juror to determine if jury selection based on surveys actually "works." When the relationships implied by the research questions were clearly in one direction, one-tailed statistics were used to test the associations.

Referring once again to Figure 1, we now summarize the four correlations performed to answer the research questions.

Research Question 1

- Does the existing jury selection procedure predict juror verdicts?

In terms of Figure 1: Is there a significant correlation between CS and the predicted PI when using the information ordinarily available from juror surveys?

Two statistical tests are necessary to answer this question: To determine if the proneness index (PI) is predictable in J2 using the information supplied by J1, we first apply the regression equation derived in the exploratory sample (J1) to the validation sample (J2):

$$\hat{pi2} = B_1 + B_2D_1 + \ldots + BnDn,$$

where $D_1 \ldots Dn$ are the D variables selected by exploratory regression procedures on J1, B_1, $B_2 \ldots Bn$ are the associated unstandardized beta weights, and $\hat{pi2}$ is the predicted proneness index (PI) of the jurors in J2. Once the predicted proneness index for sample J2 is calculated we can then test the cross sample predictability of the derived regression equation through

$$r(\hat{pi2}, PI2), \qquad (1)$$

where pi2 is the predicted proneness index for sample J2 and PI2 is the measured proneness index for sample J2. A significant positive correlation would inform us that the field model can indeed predict the proneness index across samples, an essential first step in predicting juror verdicts.

To determine if the predicted proneness index score (pi2) is significantly related to the conviction score in the validation sample, we test

$$r(\hat{pi}2, CS2), \qquad (2)$$

where CS2 is the actual conviction score of the jurors in J2, and $\hat{pi}2$ is as defined for test (1). A significant positive correlation would inform us that the Field Model can predict CS and would validate the Field Model as defined here.

Research Question 2

• Can we create an Ideal Model that reliably predicts juror verdicts? In terms of Figure F1: can we reliably predict CS across samples? Two steps are required to answer this question. First we apply the regression equation predicting CS as derived in sample J1 to the jurors in J2:

$$\hat{cs}2=B_1 + B_2D_1 + \ldots + BnDn + B_{11}T_{11} + \ldots + BnnTnn,$$

where $D_1 \ldots Dn$ are the D variables selected by exploratory regression procedures on J1, B_1 $B_2 \ldots Bn$ are the associated unstandardized beta weights, $T_{11} \ldots Tnn$ are the trial related variables, and $\hat{cs}2$ is the

predicted conviction score (CS) of the jurors in J2. Next we test the cross sample predictability of the derived regression equation through

$$r(CS2, \hat{cs2}), \qquad (3)$$

where CS2 is the actual conviction scores of the jurors in J2, and cs2 is the predicted conviction score (CS) of the jurors in J2. A significant positive correlation would tell us that an Ideal Model is possible which can predict the juror verdict proxy, the conviction score.

Research Question 3

- Are the underlying scientific jury selection assumptions of strong and identifiable relationships between background variables, conviction proneness, and juror verdicts justified?

In terms of Figure F1: is there a correlation between D and CS, and/or between PI and CS?

Correlation (3) tests the relationship between the demographic variables (D) and the conviction score (CS) across samples; and a test of

$$r(CS, PI), \qquad (4)$$

where CS is the conviction score, and PI is the proneness index, would tell us if there is a significant positive relationship between the attitudinal variables (Q) and the conviction score (CS) which can possibly be determined.

The balance of this chapter details the procedures applied in the selection of the variables used in the statistical tests and the results of the analyses.

Deriving the Proneness Index Equation

The first step in predicting conviction proneness is necessarily the generation of a conviction proneness proxy. Twenty-three (23) attitude questions were factor analyzed using a Principal Factor Analysis[1]. The first factor analysis was performed only upon those respondent jurors in the exploratory sample (J1) with complete data on all 23 items (N=147). In order to conserve N and increase the power of our analyses a criteria of 80% completion was established: if a respondent had more than four missing items he or she was deleted from the data set, otherwise the mean score of the particular missing item for the sex X education cell in which the respondent belonged was assigned when that item was missing. With the 80% completion criteria the J1 respondent pool numbered 178. A visual comparison of the two factor analyses

---

[1] All of the data manipulations and statistical tests reported below were performed using the Statistical Analysis System (SAS Institute Inc.,1982).

(N=147 and N=178) showed similar factor structures. The results of the Principal factor analysis on J1 are shown in Table 5

TABLE 5

Components of the Conviction Proneness Proxy (PI)
in the Exploratory Data Set

Q[1]	Variable	Mean[2]	Std	F1 Loading	Factor Score Coeff
23	Probglty	.93	.92	.55	.13
24	Buyout	1.28	.88	.28	.06
25	Social	1.62	.86	.06	.02
26	Laweff	1.59	.96	-.22	-.04
27	Jgeinf	.73	.83	-.07	-.01
28	Goalong	.36	.74	.23	.05
29	Accomp	.87	.75	.41	.10
30	Faith	1.43	1.25	.49	.11
31	Single	.38	.66	.42	.09
32	Jgeknows	.34	.66	.38	.08
33	Death	1.72	1.22	.34	.07
34	Attstate	.91	1.12	.57	.14
35	Certain	2.50	.91	.39	.08
36	Witness	1.14	1.21	.55	.13
37	Looks	.56	.81	.29	.05
38	Obednce	1.75	1.05	.63	.20
39	Jurtalk	1.13	.93	.06	.01
40	Reputatn	1.44	1.00	.44	.11
41	Useforce	1.29	.99	.46	.10
42	Innocent	1.35	1.17	.24	.04
43	Noreport	.89	.91	.30	.05
44	Poltruth	2.56	.66	-.27	-.06
45	Fifth	1.35	.94	.59	.16

Factor	Eigenvalue[3]	Proportion of Common Variance
1	3.57	.55
2	1.22	.19

[1]Q refers to the Question number on questionnaire (see Appendix A).
[2]Likert values for Q23 - Q45 were reversed and modified to: 0 = disagree strongly, 1 = disagree, 2 = somewhat agree, 3 = agree strongly.
[3]There were 2 factors with eigenvalues greater than 1.

The first factor (non-rotated) accounted for 55% of the common variance, and the multiple correlation (R^2) of the 23 variables with the first factor equaled .84. The standardized scoring coefficients used to assign the conviction proneness proxy to each J1 juror are also listed in Table 5 The computations were done separately for J1 and the cross-validation sample (J2) since, for the purposes of this study, they are to be considered two different samples of the same potential juror population.

The other exploratory procedures were performed only on J1, and Table 4 lists all of the variables tested as possible predictors.

Eighteen(18) variables plus a random variable were entered in a forward stepwise procedure. There were 112 subjects with complete data on these 18 variables. Four variables were entered and kept in the generated regression equation. The four variables were then tested in a regression equation including some cases previously deleted because of missing information on potential variables (thereby increasing the N to 140) with the results shown in Table 6

Further improvements of the prediction equation were tested hierarchically: that is, each new variable or set of variables was added to the predictor set of Language, Education, NYTimes, and Orgsize; and the additional variance accounted for by the extended model tested by

$$F = \frac{(R_2^2 - R_1^2)/k}{(1-R_2^2)/N-k_2-k_1-1} \quad \text{for DF} = \frac{k_2}{N-k_2-k_1-1}$$

TABLE 6

The Best Model Predicting the Proneness Index (PI) as Derived from the Exploratory Sample (N=140)

Variables	Unstandardized Beta Weight	t
Intercept	1.56	
Education[1]	- .23	-3.87 ***
Language[2]	- .89	-2.74 **
Orgsize[3]	- .13	-2.03 *
NYTimes[4]	.44	2.42 *

R^2 = .19; F(4,135) = 7.71 ****

* p < .05
** p < .01
*** p < .001
**** p < .0001

[1]Education refers to Q 54 (see Appendix A); 0=HS incomplete, 1=HS complete, 2=some college, 3=college grad, 4=grad/prof school attendence.
[2]Language refers to Q 69; 1=english native language, 2=other native language.
[3]Orgsize refers to Q 46; 0=under 10 employees in workplace, 1=10 to 30 employees, 2=30 to 100 employees, 3=over 100 employees, 4=over 1000 employees.
[4]NYTimes refers to Q 49; 1=reads NY Times,0=reads other.

Where R_1^2 = variance accounted for by the old model, R_2^2 = variance accounted for by the extended model, N = number of respondents, k_1

= number of predictors in the old model, and k_2 = number of predictors added for the extended model.

The null hypothesis tested is that, for the population, there is no increment in PI variance accounted for when the new variables are added to the set of predictors. A 't' test of the increase in R^2 associated with the addition of a single variable yields the identical figure as a 't' test of that variable in ordinary least squares regression (OLS). To test for curvilinear relationships the square of each predictor variable was added to the set of four predictors one at a time and showed no significant increase in variance accounted for as reflected by the increase in R^2. Two way interactions were tested one at a time and in all possible combination adding no significant variance; three way interactions were tested and likewise non-significant.

Nominal variables with more than two categories were dichotomized in different ways and did not add to the explanatory power of the 'best' equation. In fact, all of the demographic variables which could possibly have an effect on the prediction equation were tested and examined at different dichotomous and ordinal levels separately and in various combinations without any significant increase in explained variance. Additionally, the strong relationship between respondents' age and education was investigated and the respondents' age substituted for education with a resultant significant drop in exlained variance ($R_2^2-R_1^2=.07$, $p<.05$); despite their strong relationship subsitution was

not effective. The religiosity variable (Q 59) was significantly related to the proneness index, $r = -.19$, $p<.03$; but adding the variable either in its original ordinal coding or binary coded did not increase the predictive power of the 'best' equation. Breaking out the education variable at different points for binary coding instead of the ordinal coding used for the prediction equation did not increase the R^2 significantly.

In short, it was not possible to either improve or refine the 'best' equation originally spewed forth by the first stepwise regression procedure. The 'best' equation as reported in Table 6 is indeed the best prediction equation of conviction proneness derived from these data.

Validating the PI Prediction Equation

The analysis up to this point has paralleled that of the jury selectors; the use of stepwise regression and other exploratory procedures have generated a prediction equation.

We now go a step further. Validating a prediction equation with a different sample selected from the same population as the sample used for exploratory purposes is a procedure recommended uniformly by those who have written about stepwise regression and/or exploratory analysis (e.g. Cohen & Cohen, 1975, p. 102 ff; Draper & Smith, 1981, p. 419 f). And yet, as far as can be determined from the literature, jury selectors do not split their samples and cross-validate their

results. In validating the "best" equation we are not interested in redoing an exploratory analysis with the validating sample and then matching equations, but in determining if our "best" equation can significantly predict the proneness index in the validating sample. This, essentially, is how the jury selection survey is used in the real world: a profile is developed from a sample of the potential juror population (the survey), and used to select jurors from another sample from the same population (the available jurypool). In this study the prediction equation:

$$\hat{pi}=1.56 -.89LANG -.23ED -.13ORGSIZE +.44NYTIMES$$

was used to predict the proneness index of the validation sample (J2). These predicted proneness index scores (pi2) correlated significantly with the measured proneness index of the validation sample (PI2):

(1) $r(\hat{pi2},PI2)=.23$, one tailed $p<.01$.

Jury selection survey analysis can indeed predict conviction proneness as usually defined.

The Proneness Index and the Conviction Score

It has been demonstrated that the Proneness Index prediction equation can predict across samples. However, validating the predicted proneness index for selection purposes involves more than deriving a stable prediction equation. The essential justification for the jury selection survey is the assumption that the predicted proneness index is strongly related to the potential behavior of the jurors under actual trial conditions. This assumption is not testable under ordinary selection procedures for many of the reasons outlined previously, but the data collected in this study can test this basic presupposition.

In the exploratory sample there is a significant but moderate correlation between the predicted proneness index and conviction score, $r = .16$, one-tailed $p<.03$. However, in the validation sample the correlation is not significant,

$$(2) \qquad r(\hat{pi}2, CS2)=.09, \text{ one tailed } p>.10.$$

Therefore we conclude that the proneness index prediction equation derived from one sample can not be used to predict the conviction scores of another sample beyond chance expectations.

Berk (1977) among others has suggested that one should expect sizeable errors and should not use the results of selection surveys

unless differences between groups on possible selection variables are huge as well as significant. In that spirit the two samples were grouped into pi quintiles; the highest quintile in each sample was compared to the lowest using student's 't': the null hypothesis tested is that there is no difference in conviction score between the upper and lower quintiles as defined by the predicted proneness index. The results of the 't' tests parallel the correlations (see Table 7).

TABLE 7

Conviction Score Means of Upper and Lower Predicted Pronenes Index Quintiles in the Two Samples

Sample	N	Mean	Standard Deviation
Exploratory			
Upper Quintile	28	.27[1]	.55
Lower Quintile	27	-.06	.80
Validation			
Upper Quintile	25	-.02[2]	.77
Lower Quintile	28	-.11	.78

[1]Difference between means, $t(53)=1.83$, one tailed $p<.04$.
[2]Difference between means, $t(51)=0.46$, one tailed $p>.10$.

In the exploratory sample the high quintile had a significantly higher mean conviction score, $t(53) = 1.83$, one-tailed $p<.04$; and the cross-validation sample showed no significant difference, $t(51) = .46$, one-tailed $p>.10$. Therefore the conclusion is reinforced that juror convic-

tion behavior as measured by the conviction score is not predictable by the proneness index equation from one sample to the other.

Deriving the Conviction Score Prediction Equation

Conviction scores were assigned to each juror in sample J1 and sample J2 as outlined previously: the initial deliberation vote was weighted by the jury vote distribution to generate the conviction scores, CS1 and CS2 respectively. Very much the same exploratory procedure which generated the PI prediction equation was followed in deriving the CS prediction equation. A forward stepwise regression was performed on the same 18 variables plus a random variable. Only one variable met the criterion of entry into the equation -- Famvctm: has any member of your family been the victim of a crime? The basic equation has just one predictor and that variable significantly predicts the opposite of any normal expectation: those respondent jurors who reported that some member of the family has been a victim of a crime have significantly lower conviction scores than those who didn't, $t(111)= -2.82$, $p<.01$! Famvctm is a binary variable. Binary coded independent variables in regression equations actually test the differences between dependent variable means of the two groups coded on the independent variable with the identical result one would get with a 't' test of these two groups represented by the binary variable. Before proceeding the counter-intuitive finding was investigated further. At this point it is

important to realize that the previous analysis regarding the proneness index examined the attitudes of the respondents without reference to any trial experience in accordance to the usual jury selection procedure which tests potential jurors. The present analysis has no such restriction; all the information available can be used to derive an optimum prediction model.

A separation of the trials into victimless(victim=0) and crimes which had a personal victim(victim=1) clarified the heretofore strange finding. Table 8 presents the results of adding Victim and the Victim X Famvctm interaction to the prediction equation. Notice that there is no significant main effect for Victim, but that the interaction is significant, $t(163) = 2.75$, $p<.01$.

Looking at Figure 2 one sees that those respondents who report a family member victim of crime have noticeably lower conviction scores when they are deliberating on a victimless crime. It is as if those respondents were saying, "why are we wasting our time on this when there is real crime there in the streets."

TABLE 8

The Model Predicting Conviction Score (CS) in the Exploratory Sample with Victim Oriented Independent Variables (N=164)

Variables	Unstandardized Beta Weight	t
Intercept	.24	
Famvctm[1]	- .51	- 4.36 ****
Victim[2]	- .15	- 1.21
Famvctm x Victim	.57	2.75 **

R^2 = .11; F(3,160) = 6.56 ****

** p < .01
**** p < .001
****** p < .0001

[1]Famvctm refers to Q 21 (see Appendix A); 1=yes, 0=no.
[2]Victim refers to Q 6; 1=victim crime, 0=victimless crime.

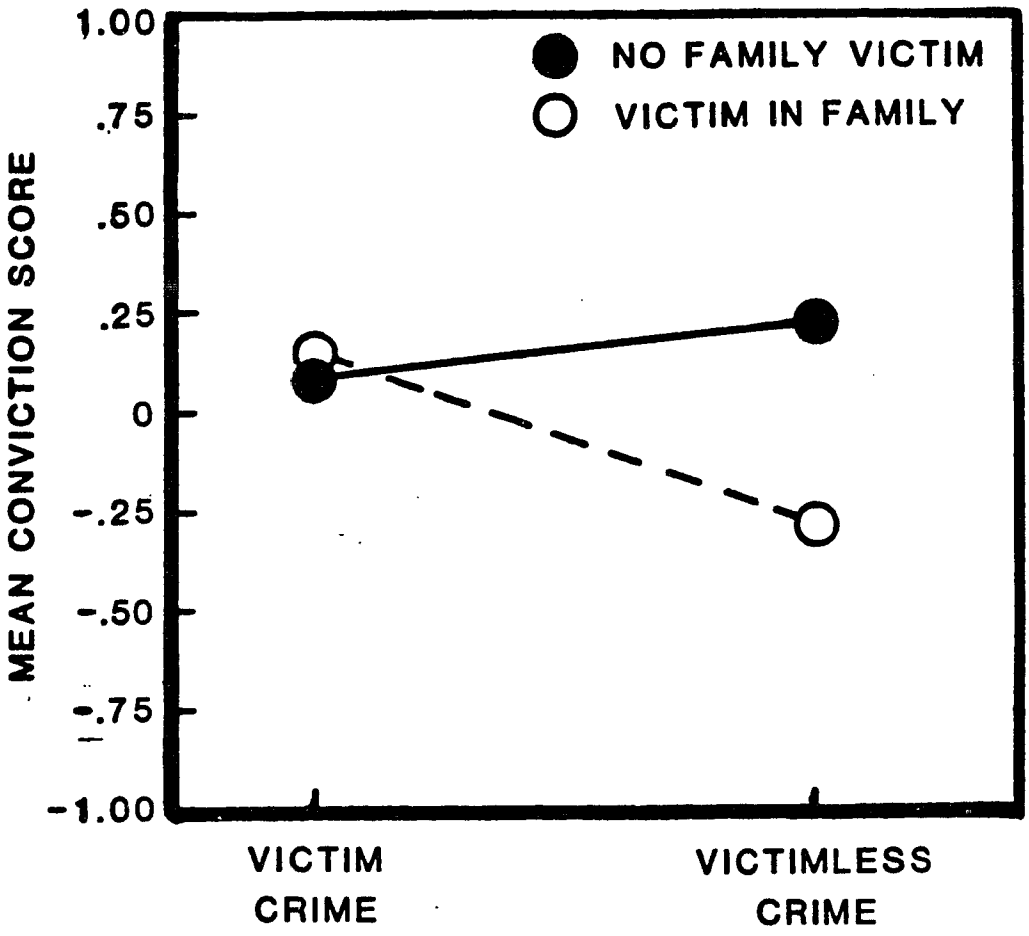

Figure 2. The interaction in the pentultimate equation predicting conviction score (CS) in the exploritory sample.

The model of CS elicited so far informs us that the search for simple predictors characteristic of the jurors might be misdirected. Here, obviously, the effect of the person's experience is changed drastically by the content of the jury trial itself. A better procedure might be the investigation of the trial characteristics which affected the conviction score as well as the search for personal predictors which magnify or interact with such predictors.

The trials were coded as drug or non-drug; violent or not violent, white defendant or not white defendant, male or female defendant, and an ordinal scale of the defendant's occupational status. The correlation of a binary variable with a continuous variable is called the Point Biserial Correlation and is distributed as 't' with $n_1 + n_2 - 2$ degrees of freedom. The 't' score obtained is identical to that one would obtain from the usual students 't' test of means between the two groups here coded as 0 and 1. Table 9 displays the correlations between the trial variables and the conviction score; there seems to be no trial variable other than 'Victim' which significantly affects the conviction score.

The basic equation was then used as the root of possible predictor equations with the increases in R^2 once again serving as the criteria for model improvement: if the test of the increase in R^2 with the use of additional variables was significant, then the variables were added to the predictor equation. A plot of the conviction score means and the education variable (Figure 3) revealed a curvilinear relationship.

TABLE 9

Correlation (N) of Conviction Score (CS) with Trial Related Variables in the Exploratory Sample

	Drugs[1]	Defsex[2]	Defage[3]	Defwhite[4]
Conviction Score	.08(168)	-.08(175)	-.04(166)	-.02(175)

	Defocc[5]	Deftest[6]	Lawtest[7]
Conviction Score	0.0(131)	.12(174)	.09(174)

Note: all ps > .10.
[1]Drugs refers to Q 6 (see Appendix A); 1=drug related crime, 0=others.
[2]Defsex refers to Q 10; 1=male, 0=female.
[3]Defage refers to Q 11.
[4]Defwhite refers to Q 12; 1=white defendant, 0=others.
[5]Defocc refers to Q 13 recoded as an ordinal variable ranging from 1=laborer to 6=professional.
[6]Deftest refers to Q 7; 1=yes, 0=no.
[7]Lawtest refers to Q 8; 1=yes, 0=no;

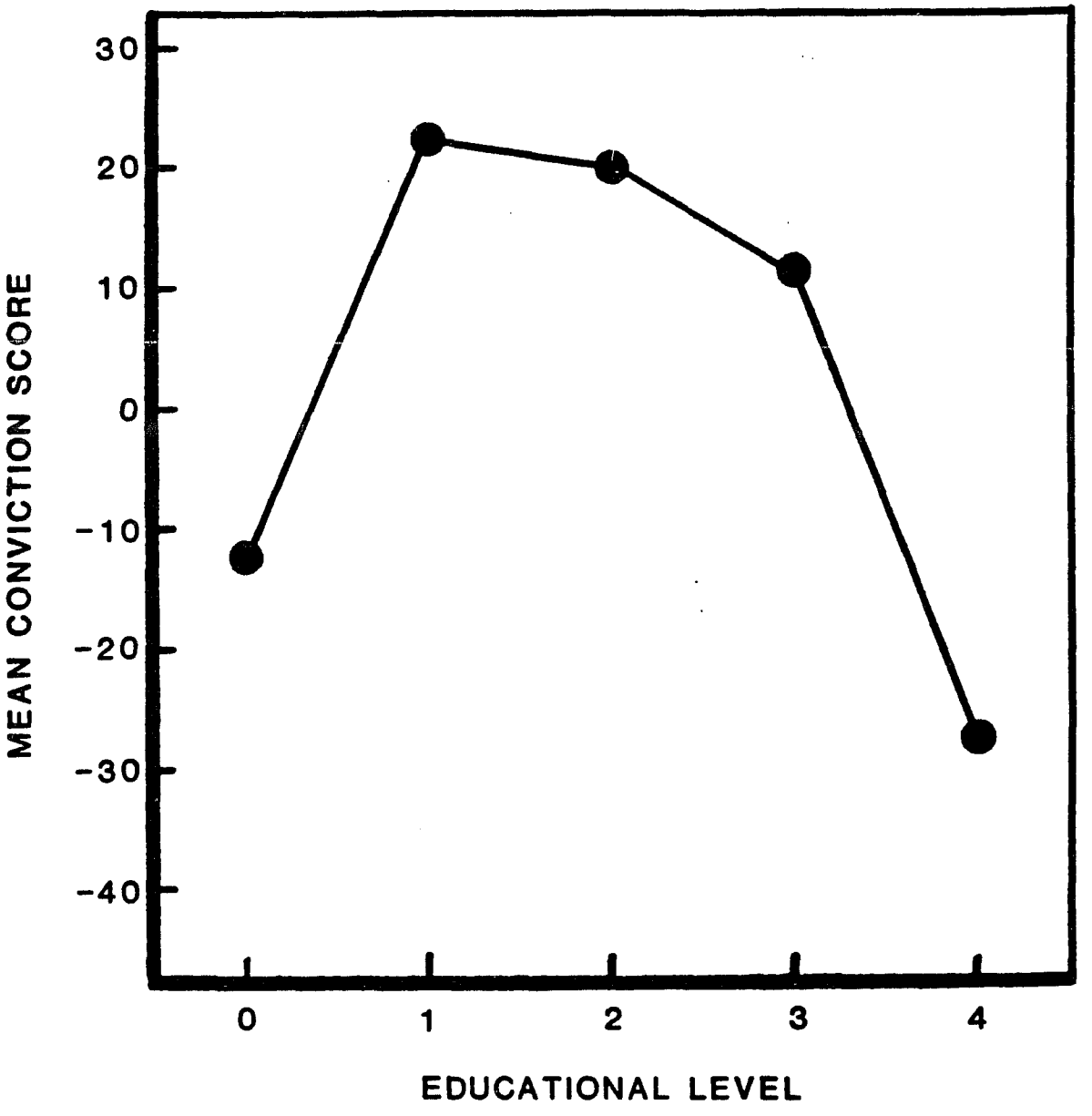

Figure 3. The relationship of respondent's educational level and conviction score (CS) in the exploratory sample. Where 0 = less than H.S. education, 1 = H.S. Graduate, 2 = some College, 3 = College Graduate and 4 = Professional or Graduate School attendence.

The relationship between education and the conviction score was investigated three ways:

- education as an ordinal variable and education squared were added to the basic model with a non-significant increase in variance of .0237, $F(2,154) = 2.64$, $p>.1$;
- the education variable was coded into a three set dummy variable: high, low, and others, This extended model showed a non-significant additional variance of .305, $F(2,154) = 2.64$, $p>.1$;
- the education variable was split into graduate school attendance and no graduate school attendance, a binary variable, and there was a significant improvement in the prediction equation, of .0218, $F(1,155) = 3.95$, $p<.05$; this was adjudged the best model using an education variable, and that coding scheme was used to extend the basic equation now shown in Table 10

Other possible predictor variables listed in Table 4 and not yet tested were added to the model, but no further improvement in the model was possible with these data. A visual examination of the residuals plotted against predicted values showed no abnormalities.

TABLE 10

The Best Model Predicting Conviction Score as Derived From the Exploratory Sample (N=160)

Variables	Unstandardized Beta Weight	t
Intercept	.30	
Famvctm[1]	- .52	- 4.44 ****
Victim[2]	- .20	- 1.62
Famvctm x Victim	.56	2.73 **
Edhigh[3]	- .29	- 1.99 *

R^2 = .15; F(4,155) = 6.65 ****

* p < .05
** p < .01
**** p < .0001
[1]Famvctm refers to Q 21 (see Appendix A); 1=yes, 0=no.
[2]Victim refers to Q 6; 1=victim crime, 0=victimless crime.
[3]Edhigh refers to Q 75;1=Grad School attendence,0=others.

Validating the Conviction Score Prediction Equation

The derived CS prediction equation:

$$\hat{cs} = .30 - .20VICTIM - .52FAMVCTM + .56INTVCTM - .29EDHIGH.$$

was used to assign predicted conviction scores ($\hat{cs}2$) to juror respondents in the validation sample (J2). The correlation of $\hat{cs}2$ with the conviction score was not significant,

(3) $r(CS2, \hat{cs2}) = .11$, one tailed $p > .05$.

The validation sample was then separated in high and low cs2 scorers and tested for differences in actual conviction scores. Once again the 't' test paralleled the correlation (see Table 11); there was no signficant difference between the means, $t(100) = 1.37$, one-tailed $p > .05$.

TABLE 11

Actual Conviction Score Means of High and Low Predicted Juror Conviction Scores in the Exploratory Sample

Predicted Scores	N	Mean	Standard Deviation
High	75	.03[1]	.76
Low	35	-.18	.71

[1]Difference between means, $t(108) = 1.37$, one tailed $p > .05$.

Therefore one must conclude that the proxy for juror conviction behavior, the conviction score, is not predictable from one sample to the other.

The Conviction Proneness - Conviction Score Assumption

The fundamental assumption of scientific jury selection is the strong relationship between juror conviction proneness and juror voting behavior. The test of this relationship in the two samples, J1 and J2, gave contradictory results. The correlation of the constructed proneness index (PI) and the derived conviction score (CS) in J1 was .02, one tailed $p>.10$, not significantly different from zero. However, the correlation of these variables in J2 was .13, one way $p<.05$; showing that jurors with higher proneness index scores were more likely to have higher conviction scores in this sample. When the two samples were combined the correlation was trivial and not significant:

(4) $r(CS,PI)=.06$, one tailed $p>.10$.

In our juror sample there was no predictable relationship between the measures of juror conviction proneness and juror conviction behavior.

Discussion

This study has tested a small but important component of the scientific jury selection procedure. It is the survey and its attendant analysis that lends the semblance of 'science' to the entire process. In brief, this study demonstrated that while a stable conviction proneness proxy can be derived from demographic and general data not related to specific cases, it cannot dependably predict the conviction behavior of real jurors across samples; furthermore, it was shown that even with the additional information available to this study (and never available to the jury selectors) a stable prediction equation of conviction behavior based on such data could not be derived for real jurors across samples. In terms of the research questions posed by this study,

1. The proneness index (the measure of conviction proneness) prediction equation derived from the exploratory sample was used to assign predicted proneness index scores to the jurors in the validation sample. The predicted proneness index scores did not correlate significantly with the conviction scores (the measure of conviction behavior) in the validation sample, demonstrating that jury selection surveys unrelated to specific cases do <u>not</u> predict juror verdicts across samples.
2. The best possible conviction score regression equation derived from the exploratory sample was used to assign predicted conviction scores to the jurors in the validation sample. These

predicted conviction scores did not correlate significantly with the actual conviction scores in the validation sample showing that an ideal model predicting juror verdicts across samples could not be created for non-specific criminal cases even with the information at this study's disposal.

3. The measure of conviction proneness, the proneness index, and the measure of conviction behavior, the conviction score, did not correlate significantly. This finding along with the failure of the proneness index regression equation to reliably predict juror verdicts across samples, informs us that the assumption of strong and identifiable relationships between background variables, conviction proneness, and juror verdicts are not justified.

It is instructive to note that when the jury selection procedure is followed on a single sample, the methods seem to "work"; that is, significant predictor equations are generated. Furthermore, it should be noted that the predictors of conviction proness are fairly stable across samples. However, in jury selection the crucial prediction is that of juror behavior: Can a juror profile generated from a sample of the juror population be used to select favorable jurors in the available jurypool? It is here that jury selection based on surveys, like the one used in this study, fail.

All exploratory procedures in general, and stepwise regression in particular, capitalize on data set idiosyncrasies in determining the best equation. It can be shown that the experimentwise error rate when looking at 18 possible predictor variables is closer to .60 when nominally set a .05 (Harris, 1975). Thus we can see that a single predictor derived from a stepwise regression -- such as the Family Victim variable derived above as a predictor of the conviction score -- is one and one-half times as likely to not be indicative of the true relationship in the population as it is to be a true indicator. To complicate matters most computer stepwise regression programs suggest the use of an alpha level of .10 thereby increasing the experimentwise error rate to .85! Without cross validation these procedures can often lead to incorrect inferences. Cross validation in another sample resets the experimentwise error rate to .05. Multiple comparison techniques such as Bonferonni or Scheffe adjustments of the critical value (Harris, 1975), are useful in theoretical situations but can sometimes lead to the opposite, Type II error.

In an applied situation when important decisions hinge upon the shape of the 'best' equation, it is incumbent upon the analyst to cross-validate on a different sample from the same population before application (splitting in half is one way). In the real world, there is no other check!

The Juror

The very name, scientific jury selection, implies expertise in juror selection over and above the usual courtroom routine. Why else would it be necessary? why else would defendants and attorneys go through the time and trouble involved in using it? This study can be seen as a test of this claim. For a procedure this expensive in time and money to justify its existence it must be able to differentiate the conviction prone from the defendant prone jurors to a much greater degree than the ordinary practicing attorney with regard to actual conviction behavior. As we have seen, this replication of the scientific jury selection has not been able to so differentiate; a basic assumption of jury selection using survey results has not been corroborated: there was no relationship between the general attitudes on which the estimate of conviction proneness is based, and the actual conviction behavior of the respondent jurors.

But how is that possible?

The easily swayed and manipulable juror has become a standard of American folklore. It is almost common sensical to accept the claims of scientific jury selectors since we all 'know' intuitively that jurors can be manipulated and the aura of 'science' is such that we <u>expect</u> the very claims made. But the dearth of scientific as opposed to anecdotal evidence makes these claims problematic.

A good part of the distrust of the jury is the tacit, but common, disregard many 'experts' and intellectuals have for democratic institutions which the media has absorbed and dispersed: how can a citizenry, half of whom have less than average intelligence, decide important questions properly? how can Bill and Mary six-pack[2], who have elevated Mork and Mindy to TV pre-eminance, be trusted to make sensible decisions? (Herbert Spencer called a jury 12 men with average ignorance.) But, despite the intuitive appeal of the 'ignorant juror,' it seems, as Saks (1976a) has insisted, that jurors really do pay attention to the evidence, leaving very little to be accounted for by other less important variables. What is most ironic about jury criticism, in factual as well as fictional accounts, is that it flows from imagination and speculation. Those who have actually served on a jury that came to a verdict (and I am one) usually are impressed with the seriousness, the skill, and the competence with which the jurors approach their task (e.g. Kennebeck, 1975; Schor in Ellison & Buckhout, 1982; Shulman, 1973).

---

[2] A name bestowed on the TV watching couple by some market researchers.

Caveats

In the Federal Courts jury deliberations are protected from intrusion or observation. The difficulties encountered by a researcher wanting to study the jury are many and perplexing. This study has attempted to examine jurors and juror selection while adapting to the realities of the court system using the very same techniques developed by jury selectors themselves. While we have not tested the other jury selection components, the invalidation of surveys as predictors of conviction behavior must be seen as calling the entire enterprise into question since every part of the process depends to some extent on the profiles generated through the analysis of these surveys. But the strength of this finding should be moderated somewhat by the circumstances of this particular investigation.

The content of the survey was not identical to the content of many surveys which ask additional questions specifically related to the defendant. We have not tested the type of jury selection which receives the greatest attention -- the prosecution of some famous or notorious person where the public might have strong opinions which could influence verdicts. For example, if one were surveying for a politician defendant one might have agree-disagree statements such as:

- When public officials get caught breaking the law, they are usually able to get themselves out of trouble.

- It is reasonable for a person in a high position to reward his friends.
- Anyone who is a leader of a local political party is probably corrupt.
- A certain amount of waste in government is inevitable.

Attitudes expressed towards specific defendants or a specific class of defendants could have led to different results from those reported above. This might be especially true for those cases in which media celebrants are involved, or in which value laden issues are being espoused by defendants.

Moreover, the attitudes expressed by the respondents are recorded after the trials and after the delivered verdict.

Attitudes collected immediately after jury experience might well be different from those collected at other times. However, one would expect an even closer relationship between expressed attitudes and conviction scores under such conditions. The failure to show a significant relationship between these attitudes and the immediately foregoing behavior in the jury room is therefore even more telling.

Another limitation of this study is the derived behavior measure, the conviction score. The protected status of the juror effectively prevents any direct measure of behavior other than the verdicts themselves. It is as if the jury deliberation room were the black box itself: in go the jurors, out comes the verdict. The verdict is a group deci-

sion, but the input is from individual jurors and it is the individual which the surveys try to assess. This study defines the juror's conviction behavior prior to the final verdict by comparing him or her to the other members of the jury (as the respondent reports it) and deriving a conviction score. Jurors are examined, not juries. The contraints imposed on researchers in the court system by considerations of fairness and protection of the defendant virtually preclude the comparison of entire juries under actual trial conditions.

Finally, the sample examined in this study is different from the sample used in most jury studies. Researchers make claims about jurors based on findings which actually come from studies of the public -- a different population. It is true that the jurypool is, theoretically at least, randomly selected from the public, but jurors are not randomly selected.

We have not tested the general population here; we have not even tested the general potential juror population, or even a sample of the jurypool; we have tested a sample of jurors who have already been selected and approved by defense attorneys and prosecutors alike. The sample examined for this study is a doubly selected group of jurors; selected in the most ordinary unremarkable courtroom routine: self-selected because each is one of the less than 50% who answer the court summons; attorney-selected because each juror in our sample has gone through voir dire.

Christie (1976) has argued that the primary function of jury selection is the elimination of potential jurors who cannot be fair. This study is of selected jurors, those who have survived voir dire; a different sample, surely, than the jurypool despite the similarity in demographic profile.

There is a real problem in the use of mock juries to research jury behavior. It is not only that behavior in the 'real' world is different from behavior in mock juries; behavior in the courtroom is as artificial as any contrived experimental situation, but the artificiality is imposed by a different set of rules, different from those we usually play by. These rules are implicitly and explicitly defined at the voir dire; this is a deadly serious game we indulge in and those that can't play properly are soon dismissed which might account for the lack of differences among those who remain.

Anyone involved in a criminal trial would certainly welcome jurors of like opinions or at least some who do not regard one with prejudice. Strangely enough in the sample of jurors we have examined here, it made no difference to the outcome of the trial. This is not to say that voir dire is unnecessary. Indeed one can conclude the reverse: voir dire is so effective in eliminating hopelessly prejudiced jurors that juries actually decide cases on the evidence; most individual jurors can somehow accept the role of impartial judges when required to do so.

It is true that we are only looking at mundane cases here and that different questions would be asked that would be more specific to the trials viewed; but we must also realize that the attorneys have pre-selected our sample knowing the contents of these cases and that they would do the same for the more media celebrated cases. This study might be viewed as a test of survey usefulness over and above the selection prowess of attorneys! Using survey information to help select jurors in mundane cases has failed this test and the burden of proof regarding the efficacy of the scientific jury selection procedure still rests on the shoulders of those who would use it.

In Conclusion

As noted earlier, Etzioni (1974b) attacks the very existence of scientific jury selection. Many of the published answers to Etzioni are not responsive to the questions he raises. To say that jury selection is not infallible is not to dispute Etzioni's thesis. He makes very clear the fallibility of the scienfific jury selection technology. He merely claims that it increases the likelihood of receiving a favorable verdict and therefore interferes with the normal workings of the CJS; a rebalancing of the scales to the eventual detriment of the very groups the selection originators such as Jay Shulman professed to be helping. Etzioni states that notorious individuals and the wealthy will be helped in their jury trials, while the vast majority of trial participants will be handicapped.

I am pleased to report that the change in likelihood that Etzioni predicted does not, in fact, exist. Another poor response to be sure; but I feel the danger is not in more accurate jury selections, but in changing the legal system based on the results of questionnable studies. In some respects this study can be considered a reaffirmation of our jury system. The failure to show a strong relationship between juror attitudes and behavior in the jury room supports the usefulness of voir dire in jury selection and argues against the always present proposals to change jury trial procedure in the name of efficiency or fairness.

Thibaut & Walker(1975) have argued that if the adversary system is faulted because of the unfair advantages it confers on the rich and powerful, it is the society which is to blame not the judicial system. Being rich and powerful is an advantage in every one of our endeavors, in every aspect of our society. This is not a just society, it is (hopefully) a society becoming just. Instead of changing our legal procedures, what must be done is somehow to ensure equal resources to the litigants involved.

Appendix A

JUROR SURVEY

prepared by

The Department of Psychology
Brooklyn College, C.U.N.Y.

PLEASE PUT A CHECK IN THE BOX FOR YOUR ANSWER

DO NOT WRITE IN THIS SPACE

1. During this session, were you ever called in to join a panel for jury selection? ☐ YES ☐ NO — ☐ 6

2. How many times were you brought in this session and questioned about serving as a juror in a criminal case? _____ — ☐ 7

3. Were you seated as a juror/alternate in a criminal case during this session? ☐ YES ☐ NO — ☐ 8

4. If you served as a juror or alternate, how many cases did you hear and deliberate on? _____ — ☐ 9

5. If you were not seated as a juror, can you indicate who was responsible for excusing you?

☐ The judge
☐ The defense
☐ The prosecution
☐ Don't know who
☐ Doesn't apply, I was a juror

— ☐ 10

IF YOU SERVED, PLEASE FOCUS ON YOUR LAST CRIMINAL JURY CASE WHEN ANSWERING THE FOLLOWING QUESTIONS. IF YOU DID NOT SERVE ON ANY CRIMINAL JURY FOR THIS ENTIRE SESSION, PLEASE SKIP THE FOLLOWING QUESTIONS AND BEGIN ANSWERING THE QUESTIONS ON PAGE 3 (QUESTION 20).

6. What was the defendant charged with in your trial? (Describe)

_____ — ☐ 11

7. Did the defendant testify? ☐ YES ☐ NO — ☐ 12

8. Did any law enforcement agents testify in your trial? ☐ YES ☐ NO — ☐ 13

9. How long was your jury trial? _____ — ☐☐ 14 15

PLEASE FOCUS ON THE PRINCIPAL DEFENDANT IF THERE WAS MORE THAN ONE.

DO NOT WRITE IN THIS SPACE

10. Defendant's sex: ☐ MALE ☐ FEMALE — 16

11. Defendant's age: _____ — 17 18

12. Defendant was: ☐ White ☐ Black ☐ Asian American ☐ Spanish American ☐ Other _____ (specify) — 19

13. Defendant's occupation: _____ — 20

NOW, RECALL THE JURY DELIBERATIONS AFTER THE TRIAL

14. What was the result of the first vote taken by the jury:

Number voting guilty _____ — 21 22

Number voting not guilty _____ — 23 24

Number neutral _____ — 25 26

15. You voted: ☐ GUILTY ☐ NOT GUILTY ☐ NEUTRAL — 27

16. What was the final vote taken by the jury:

Number voting guilty _____ — 28 29

Number voting not guilty _____ — 30 31

17. You voted: ☐ GUILTY ☐ NOT GUILTY — 32

18. How long did your jury deliberate before reaching a verdict? _____ — 33 34 35

19. As you now examine your feelings about your vote, how would you rate your feelings on the following scale. (Please circle a number) — 36 37 38

0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
I have no confidence that I made the right decision										I'm absolutely confident that I made the right decision

PLEASE CHECK THE BOX FOR YOUR ANSWER.

DO NOT WRITE IN THIS SPACE

20 Have you ever been a victim of a crime? ☐ YES ☐ NO ☐ 39

21. Has anybody in your family been a victim of a crime? ☐ YES ☐ NO ☐ 40
Describe relationship _____

22. In Criminal cases the judge tells the jury that the defendant has to be proven guilty beyond a reasonable doubt. Regardless of the law, many people feel that the defendant has to prove himself innocent if he is to be acquitted. If you sit as a juror you would ☐ 41

☐ Want the defendant to prove his innocence
☐ Want to hear both sides try to prove their case
☐ Follow the judge's instructions

FOR THE FOLLOWING, PLEASE CHECK HOW MUCH YOU AGREE OR DISAGREE WITH THE STATEMENT.

	Agree Very Much	Agree Some What	Disagree Some What	Disagree Very Much	
23. If the authorities go to the trouble to bring someone to trial, he is probably guilty.	☐	☐	☐	☐	☐ 42
24. When businessmen get caught breaking the law, they are usually able to buy themselves out of trouble.	☐	☐	☐	☐	☐ 43
25. Poor social conditions are the cause of crime.	☐	☐	☐	☐	☐ 44
26. The personality and conduct of the lawyers can often affect a juror's opinion.	☐	☐	☐	☐	☐ 45
27. The juror is often influenced by the judge in making up his or her mind.	☐	☐	☐	☐	☐ 46
28. If one juror holds a different point of view from the foreperson and the rest of the jurors, that juror should go along with the majority.	☐	☐	☐	☐	☐ 47
29. Testimony of an accomplice to a crime should generally be disbelieved.	☐	☐	☐	☐	☐ 48
30. Every person should have complete faith in God.	☐	☐	☐	☐	☐ 49

	Agree Very Much	Agree Some What	Disagree Some What	Disagree Very Much	DO NOT WRITE IN THIS SPACE	
31. Single women who have had intimate relationships with more than one man tend to be untrustworthy.	☐	☐	☐	☐	☐	50
32. The judge in a trial is the best person to decide what really happened.	☐	☐	☐	☐	☐	51
33. The death penalty is more effective than a life sentence without parole in keeping people from committing murder.	☐	☐	☐	☐	☐	52
34. Statements made by either attorney during a trial are important parts of the evidence.	☐	☐	☐	☐	☐	53
35. In order to reach a verdict or finding a juror must be absolutely certain that the evidence supports the finding.	☐	☐	☐	☐	☐	54
36. Defendants in a criminal case should be required to take the witness stand.	☐	☐	☐	☐	☐	55
37 A good juror can usually spot the motives of a party to a trial by observing his/her appearance.	☐	☐	☐	☐	☐	56
38. Obedience to authority is the most important virtue children should learn.	☐	☐	☐	☐	☐	57
39. When at home a juror will often times discuss a case.	☐	☐	☐	☐	☐	58
40. Testimony about the reputation of a witness is important evidence of the facts in a case for a juror.	☐	☐	☐	☐	☐	59
41. Police should not hesitate to use force to maintain order.	☐	☐	☐	☐	☐	60
42. Finding a defendant not guilty means that he/she is innocent of the crime.	☐	☐	☐	☐	☐	61
43. A person who fails to report to the police on the activity of someone else is probably equally guilty.	☐	☐	☐	☐	☐	62
44. I generally assume that when a law enforcement officer testified, he is less likely to tell the truth than other witnesses.	☐	☐	☐	☐	☐	63
45. A witness who takes the 5th amendment (refuses to testify) is probably hiding his guilt of a crime.	☐	☐	☐	☐	☐	64

BACKGROUND DATA

DO NOT WRITE IN THIS SPACE

PLEASE ANSWER THE FOLLOWING QUESTIONS AS HONESTLY AS YOU CAN. REMEMBER, THESE QUESTIONNAIRES ARE TOTALLY ANONYMOUS. YOU MAY CHOOSE NOT TO ANSWER ANY QUESTION: HOWEVER, TO GIVE US AN ACCURATE MEASURE, ALL QUESTIONS SHOULD BE ANSWERED.

46. The organization I work for employs the following number of people.

- ☐ Self employed
- ☐ Under 10
- ☐ 10-30
- ☐ 30-100
- ☐ More than 100
- ☐ More than 1000

☐ 65

47. Are you in a supervisory position? ☐ YES ☐ NO ☐ 66

48. How many people do you supervise? _____ ☐☐ 67 68

49. What newspaper do you read regularly? _____ ☐ 69

50. In what section of New York do you live? _____ ☐ 70

51. Do you:
- ☐ Own a house
- ☐ Rent
- ☐ Own a cooperative or condominium

☐ 71

52. How old are you? _____ ☐☐ 72 73

53. What is your marital status:
- ☐ Married
- ☐ Single
- ☐ Separated
- ☐ Divorced
- ☐ Widowed

☐ 74

54. How much schooling have you completed?
- ☐ None, or High School incomplete
- ☐ High School complete
- ☐ College incomplete or two year degree
- ☐ Four year college complete
- ☐ Two year technical college
- ☐ Graduate or professional school complete or incomplete

☐ 75

55. What is your sex? ☐ MALE ☐ FEMALE ☐ 76

56. What is/was your occupation (MALES ONLY) _____ _____

What is/was your occupation (FEMALES ONLY) _____ _____

☐ 77

DO NOT WRITE IN THIS SPACE

57. Now could you tell me if your spouse is:

MALES DESCRIBE WIFE

- ☐ Employed
- ☐ Unemployed
- ☐ Retired
- ☐ Student
- ☐ Housewife
- ☐ Does not apply

FEMALES DESCRIBE HUSBAND

- ☐ Employed
- ☐ Unemployed
- ☐ Retired
- ☐ Student
- ☐ Does not apply

☐ 78

58. What is your religion? _____ ☐ 79

59. How often do you attend religious services?

- ☐ More than once a week
- ☐ Weekly
- ☐ Several times a month
- ☐ Once a month or less
- ☐ Never

☐ 80

60. Everybody is an American but what country does does your family originally come from? _____ ☐ 6

61. Do you belong to any social or civic organization? ☐ YES ☐ NO ☐ 7

62. Could you please tell me your approximate total annual family income. (What group does it fall into?)

- ☐ Under $5,000
- ☐ $5,000 to $9,000
- ☐ $9,000 to $15,000
- ☐ $15,000 to $25,000
- ☐ Over $25,000

☐ 8

63. Could you please tell me what political party you usually supported in the past?

- ☐ Republican
- ☐ Democrat
- ☐ Liberal
- ☐ Independent
- ☐ Conservative
- ☐ Other _____ (specify)

☐ 9

64. Have you ever been called to serve as a juror before? ☐ YES ☐ NO ☐ 10

65. Have you ever served as a juror before? ☐ YES ☐ NO ☐ 11

66. If served as a juror, was it a federal or state case? ☐ Federal ☐ State ☐ Both ☐ 12

67. Was it a criminal or civil case ☐ Criminal ☐ Civil ☐ Both ☐ 13

DO NOT WRITE IN THIS SPACE

58. Are you:

- ☐ White
- ☐ Black
- ☐ Asian American
- ☐ Spanish American
- ☐ Other _____ (specify)

☐ 14

59. Is English your native language?

☐ YES ☐ NO

☐ 15

AGAIN, THANK YOU VERY MUCH FOR YOUR COOPERATION

Reference Notes

Baker, E. A report to the Federal Judges. Unpublished paper, City University of New York, 1978.

Bray, R. M. The mock trial: Problems and prospects for jury research. In R. Dillehay, (Chair), Innovations in jury research: Concepts, methods, and policy implications. Symposium presented at the annual meeting of the American Psychological Association, Washington, 1976.

Mitchell, H. E., & Byrne, D. Minimizing the influence of irrelevant factors in the courtroom: The defendant's character, judge's instructions, and authoritarianism. 1974, unpublished paper, Purdue University.

Penrod, S. D. Study of attorney and 'scientific' jury selection models. Unpublished doctoral dissertation, Harvard University, 1979.

Sigworth, H. & Henze, F. Juror's comprehension of jury instructions in south Arizona. Unpublished paper, University of Southern Arizona, 1975.

References

Adler, F. Socioeconomic factors influencing jury verdicts. New York University Review of Law and Social Change, 1973, 3, 1.

Adorno, T. W., Frenkel-Brunswic, E., Levinson, D. J., & Sanford, R. N. The Authoritarian Personality New York: Harper & Row, 1950.

Alker, H. R., Hosticka, C., & Mitchell, M. Jury selection as a biased process. Law and Society Review, 1976, 11 (1), 9-41.

Andeanaes, J. The moral or educational influence of criminal law. Journal of Social Issues, 1971, 27 (2), 17-32.

Apodeca, Cooper and Madden v. Oregon. 92 US 1928, 1972.

Balch, R. W., Griffiths, C. T., Hall, E. L., & Winfree, L. T. The socialization of jurors: The voir dire as a rite of passage. Journal of Criminal Justice, 1976, 30 (2), 152-160.

Beiser, S.N. Are juries representative? Judicature, 1973 57 (5).

Berg, K. S., & Vidmar, N. Authoritarianism and recall of evidence about criminal behavior. Journal of Research in Personality, 1975, 9, 147-157.

Berk, R. A. Social science and jury selection: A case study of a civil suit. In G. Bermant, C. Nemeth, & N. Vidmar (Eds.), Psychology and the Law. Lexington, Mass.: Heath, 1976.

Berk, R. A., Hennessy, M., & Swan, J. The vagaries and vulgarities of "scientific" jury selection: A methodological evaluation. Evaluation Quarterly, 1977, 1, 143-158.

Berman, J., & Sales, B. D. A critical evaluation of the y systematic approach to jury selection. Paper presented to the American Psychological Association, Washington, D. C., 1976.

Blumberg, A.S. Criminal Justice. Chicago: Quadrangle Books, Criminal Justice. Chicago: Quadrangle Books, 1967.

Blunk, R. A., & Sales, B. D. Persuasion during the voir dire. In D. Sales (Ed.), Psychology in the legal process New York: Spectrum Publications, 1977.

Blunk, R. A., & Sales, B. D. Persuasion during the voir dire. In D. Sales (Ed.), Psychology in the legal process New York: Spectrum Publications, 1977.

Boehm V. Mr. Prejudice, Miss Sympathy and the authoritarian personality: An application of psychological measuring to the problems of jury bias. Wisconsin Law Review, 1968, 3 734-750.

Bray, R. M., & Noble, A. M. Authoritarianism and decisions of mock juries: Evidence of jury bias and group polarization. Journal of Personality and Social Psychology, 1978, 41, 256-260.

Brodsky, S.L. (Ed.). Psychologists in the criminal justice system. Urbana, Ill.: U. of Illinois Press, 1973.

Brodsy, S. L., & Robey, A. On becoming an expert witness: Issues of orientation and effectiveness. Professional Psychology, 1972, 3, 173-176.

Broeder, D. W. The University of Chicago jury project. Nebraska Law Review, 1958, 38, 744-761.

Bruner, J. S., & Potter, M. C. Interference in visual recognition. Science, 1964, 144, 424-425.

Buckhout, R., Weg, S., Reilly, V., & Frohboese, R. Jury verdicts: Comparison of six versus twelve person juries and unanimous versus majority decision rule in a murder trial. Center for Responsive Psychology Report No. CR-12, 1976. New York: Brooklyn College, C.U.N.Y.

Buckhout, R. U. S. v. Swinton: A case history of jury selection. Social Action and the Law, 1978, 4 (4), 27-29.

Buckhout, R. Jury without peers. Center for Responsive Psychology Monograph No. CR-2, 1973. New York: Brooklyn College, C.U.N.Y.

Buckhout, R., & Baker, E. Jury attitudes and the death penalty. Social Action and the Law, 1977(a), 3 (6), 80-81.

Buckhout, R., & Baker, E. Surveying the attitudes of seated jurors. Social Action and the Law, 1977(b), 4, (6), 98-101.

Campbell, D. T., & Stanley, J. C. Experimental and Quasi-Experimental Designs for Research. Chicago: Rand McNally, 1963.

Cani, J.S. Statistical evedence in jury discrimination cases. Journal of Criminal Law and Criminology., 65 (2), 234-238.

Carp, R. The behavior of grand juries: Acquiescence or justice? Social Science Quarterly, 1975, 55 (4), 853-870.

Christie, R. Probability v. precedence: The social psychology of jury selection. In G. Bermant, C. Nemeth, and N. Vidmar (Eds.), Psychology and the Law. Lexington, Mass. Heath, 1976.

Cohen, J., & Cohen, P. Applied multiple regression/correlation analysis for the behavioral sciences. New York: John Wiley, 1975.

Davis, J.H., Kerr, N.L., Atkin, R.S., Holt, R., & Mark, D. The decision process of six and twelve person mock juries assigned unanimous and two-third majority rules. Journal of Personality and Social Psychology, 1975, 32 (1), 1-14.

Diamond, Shari, S. A jury experiment reanalyzed. University of Michigan Journal of Law Reform. 1974, 7,520-532.

Dorman, M. King of the courtroom: Percy Foreman for the defense. New York: Delacorte press, 1969.

Draper, N., & Smith, H. Applied regression analysis. New York: John Wiley & Sons, Inc, 1981.

Ellison, K. W., & Buckhout, R. Psychology and criminal justice. New York: Harper and Row, 1981.

Elwork, A., & Sales, B. D. Psycholegal research on the jury and trial processes. In C. Petty, W. Curran, & L. McGarry (Eds.), Modern Legal Medicine and Forensic Science. Philadelphia, Pa.: F. A. Davis, 1977.

Emerson, C.D. Personality tests for prospective jurors. Kentucky Law Journal, 1968, 56, 832-854.

Erlanger, H.S. Jury research in America, its past and future. Law and Society Review, 1970, 4 (3), 345-370.

Etzioni, A. Creating an Imbalance. Trial, 1974a, 10 (6), 28;30.

Etzioni, A. Science: Threatening the jury trial. Washington Post, May 26, 1974(b).

Faust, R., & Carlson, J. The impact of age and other stratification variables on attitudes toward justice. Center for Responsive Psychology Monograph No. CR-32, 1977. New York: Brooklyn College, C.U.N.Y.

Finkelstein, M.O. A statistical analysis of guilty plea proctices in the federal courts. Harvard Law Review, 1975, 89 (2).

Foss, R.D. A critique of jury simulation research. Paper presented at the American Psychological Association Meeting, Chicago, September, 1975.

Frank, J. Courts on trial: Myth and reality in American justice. Princeton, N. J.: Princeton University Press, 1949.

Friedman, H. Trial by jury: Criteria for conviction. Jury size and type I and type II errors. The American Statistician, 1972, 26 (2), 21-23.

Gerbasi, K.C., Zuckerman, M., & Reis, H. T. Justice needs a new blindfold: A review of mock jury research. Psychological Bulletin, 1977, 84 (2), 323-345.

Ginger, A.F. The practicing lawyer's search for justice: Judicial reform. The Forensic Quarterly. 1971, 45 (3).

Gleason, J.M., & Harris, V.A. Race, socioeconomic status, and perceived similarity as determinants of judgments by simulated jurors. Social Behavior and Personality,1975, 3 (2), 175-180.

Gobert, J. Can psychologists tip the scales of justice? Psychology Today, Feb, 1984, 18, pp. 38-39.

Golman, J., Maitland, K.Q., & Norton, P.L. Psychological aspects of jury performance. Journal of Psychiatry and Law, 1975.

Hans, V.P., & Doob, A.N. Section 12 of the Canada evidence act and the deliberations of simulated juries. Criminal Law Quarterly, 1976, 18 (2), 235-253.

Harris, R. J. A primer of Multivariate Statistics. New York: Academic Press, 1975.

Hunt, M. Putting juries on the couch. The New York Times, Nov. 28, 1982, pp. 72, 78.

Johnson v. Louisiana. 92 US 1935, 1972.

Jurow, G. New data on the effect of a death qualified jury on the guilt determination process. Harvard Law Review, 1971, 84, 567-611.

Jury selection and service act of 1968. U. S. C., 28, 1861-1869.

Kadish, M. R., & Kadish, S. H. The institutionalization of conflict: Jury acquittals. Journal of Social Issues, 1971, 27 (2), 199-218.

Kairys, D. Juror selection: The law, a mathematical method of analysis and a case study. The American Criminal Law Review, 1972, 10 (4), 771-777.

Kairys, D., Schulman, J., & Harring, S. The Jury System: New Methods for Reducing Prejudice. Philadelphia: National Jury Project and the National Lawyers Guild, 1975.

Kalven Jr., H. Please, Morris, don't make trouble: Two lessons in courtroom confrontation. Journal of Social Issues, 1971, 27 (2), 219-230.

Kalven Jr., H., & Zeisel, H. The American Jury. University of Chicago Press, 1966.

Kantrowitz, A. The science court experiment. Trial, March, 1977, pp. 48-49; 54.

Kaplan, M., & Kemmerick, G. Juror judgment as information integration: Combining evidential and nonevidential information. Journal of Personality and Social Psychology. 1974, 30, 493-499.

Kaplan, M. F., & Miller, L. E. Reducing the effects of juror bias. Journal of Personality and Social Psychology, 1974, 36, 1443-1455.

Kaplan, M., & Scherscing, C. Juror deliberation: An information integration analysis. In B. D. Sales (Ed.), Perspectives in Law and Psychology. Volume II: The Jury, Judicial, and Trial Process. Plenum, 1978.

Konecni, V. J., & Ebbesen, E. B. External validity of research in legal psychology. Law and Human Behavior, 1979, 3, 3

Kubia, J.F. Instrumental, chemical, and psychological aids in the interrogation of witnesses. Journal of Social Issues, 1957, 2, 40-49.

Landy, D., & Aronson, E. The influence of the character of the criminal and his victim on the decisions of simulated jurors. Journal of Experimental and Social Psychology, 1969, 5, 141-152.

Lemert, E. M. Human deviance, social problems, and social control. Englewood Cliffs, N. J.: 1967.

Lempert, R. D. Uncovering "nondiscernible" differences: Empirical research and the jury size cases. Michigan Law Review, 1975 73 (4), 643-708.

Lerner, M.J. The justice motive in social behavior: Introduction. Journal of Social Issues, 1975, 31 (3).

Levine, F. J., & Tapp, J. L. The psychology of criminal identification: The gap from Wade to Kirby. Research Contributions of the American Bar Association, 1973, 4, Whole

Levine, M. Scientific method and the adversary model: Some preliminary thoughts. American Psychologist, 1974, 661-677.

Mccart, S.W. Trial by jury: A complete guide to the jury system. New York: Chilton Books, 1966(revised).

Marshall, J. Law and psychology in conflict. New York: Bobbs-Merill, 1966.

McConahay, J. B., Mullin, C. J., & Frederick, J. The uses of social science on trials with political and racial overtones: The trial of Joan Little. Law and Contemporary Problems, 1977, 41 (1), 10-17.

Meehl, P.E. Law and the fireside induction: Some reflections of a clinical psychologist. Journal of Social Issues, 1971, 27 (4), 65-100.

Miller, M., & Hewitt, J. Conviction of a defendant as a function of juror-victim racial similarity. Journal of Social Psychology, 1978, 105, 159-160.

Mischel, W. Personality and assessment. New York: John Wiley, 1968.

Mitchell, H. E., & Byrne, D. The defendant's dilemma: Effects of jurors' attitudes and authoritarianism on judicial decisions. Journal of Personality and Social Psychology, 1973, 25 (1), 123-129.

Mode, E.B. Probability and criminalistics. American Statistical Association Journal, 1963, 58, 628-640.

Moore Jr., H.A. Redressing the balance. Trial, 1974, (6), 29; 31; 35.

Moore, L.E. The jury: Tool of kings, palladium of liberty. Cincinnati, Ohio: W.H. Anderson, 1973.

Morris, R.E. Witness performance under stress: A sociological approach. Journal of Social Issues, 1957, 2, 17-22.

Mossman, K. Jury selection: An expert's view. Psychology Today, 1973, May, pp. 78-79.

Mosteller, F., & Tukey, J. W. Data analysis and regression: a second course in statistics. Reading, Mass.: Addison-Wesley Publishing Company, 1977.

Nader, L. Forums for justice: A cross-cultural perspective. Journal of Social Issues, 1975, 31 (3).

Nagel, S., & Neef, C. Deductive modeling to determine an optimum jury size and fraction required to convict. Washington University Law Quarterly, 1975, (4), 973-976.

National Jury Project West. Jury Work. New York: Clark Broadman, 1983 (2nd ed.).

Nemeth, C., & Sosis, R. H. A simulated jury study: Characteristics of the defendant and the jurors. The Journal of Social Psychology, 1973, 90, 221-229.

Newman J. R. The World of Mathematics New York: Simon & Shuster, 1956. Volume iv.

Pennington, N., & Hastie, R. Psychological Bulletin. 1981, 89 (2), 246-287.

Reed, J. Jury deliberations, voting and verdict trends. Southwest Social Science Quarterly, 1965, 45, 361-370.

Rourke, F.L. Psychological research on problems of testimony. Journal of Social Issues. 1957, 2, 40-49.

Rummel, R. J. Applied factor analysis. Evanston: Northwestern University Press, 1970.

Sage, W. Psychology and the Angela Davis jury. Human Behavior, 1973, 2 (1), 56-61.

Saks, M. J. Jury Verdicts. Lexington, Mass. Lexington Books, 1977.

Saks, M.J. Ignorance of science is no excuse. Trial, 1974, 10 (6), 18-20.

Saks, M.J. The limits of scientific jury selection: Ethical and empirical Jurimetrics Journal. 1976(a), 17 (1), 3-22.

Saks, M. J. Social scientists can't rig juries. Psychology Today, 1976(b), 9 (8), 48-57.

Sampson, E.E. On justice as equality. Journal of Social Issues, 1975, 31 (3), 45-64.

Shapiro, M. Stability and change in judicial decision-making: Incrementalism or stare decisis? In L. M. Friedman & S. Macaulay, Law and the behavioral sciences New York: Bobbs-Merrill, 1969.

Shulman, J., Shaver, P., Colman, R., Emrich, B., & Christie, R. Recipe for a jury. Psychology Today, May 1973, pp. 37-84.

Siegel, S. Nonparametric statistics: For the behavioral sciences. New York: McGraw-Hill Book Company, 1956.

Sigall, H., & Landy, D. Effects of the defendant's character and suffering on juridic judgment: A replication and clarification. Journal of Social Psycho___ 1972, 88, 149-150.

Sigall, H., & Ostrove, N. Beautiful but dangerous: Effects of offender attractiveness and nature of the crime on juridic judgment. Journal of Personality and Social Psychology, 1975, 31, 410-414.

Simon, C. K. The juror in New York City: A survey of attitudes and experiences. A report by the subcommittee on the jury system of the department of committees for court administration of the appellate divisions, first and second departments, June, 1973.

Simon, R.J. The jury and the defense of insanity. Boston: Little, Brown, 1967.

Simson, G.J. Jury nullification in the American system: A skeptical view. Texas Law Review, 1975, 54 (3), 488-525.

Smith v. Texas. 311 US 128, 1940.

Sofer, C. Organizations in Theory and Practice. New York: Basic Books, 1972.

Sonquist, J., Baker, E., & Morgan, J. Searching for Structure. Ann Arbor, Michigan: Institute for Social Research, 1973.

Suggs, D. & Sales, B. The art and science of conducting the voir dire. Professional Psychology, 1978, 9, 367-388.

Swain v. Alabama. 380 US 202 208, 1965.

Taylor v. Louisiana. 419 US 522, 1975.

Tedeschi, J. T., & O'Donovan, D. Social power and the psychologists. Professional Psychologist, 1971, 2 (1), 59-64.

Thibaut, J., & Walker, S. Procedural justice: A psychological analysis. analysis. Hillsdale, N.J.: Lawrence Erlbaum, 1975.

Toch, H. (ed). Legal and criminal psychology. New York: Holt, Rinehart, & Winston, 1961.

Ugwuegbu, D. C. E. Black jurors' personality trait attribution to a rape case defendant. Social Behavior and Personality, 1966, 4, 193-201.

Valenti, A. C., & Downing, R. L. Differential effects of jury size verdicts following deliberation as a function of the apparent guilt of the defendant. Journal of Personality and Social Psychology 1975, 22 (4), 655-663.

Van Dyke, J. M. Jury Selection Procedures: Our Uncertain Commitment to Representative Panels. Cambridge, Mass.: Ballinger, 1977.

Vidmar, N. Effect of decision alternatives on the verdicts and social perceptions of simulated jurors. Journal of Personality and Social Psychology 1975, 22 (2), 211-218.

Walster, E., & Walster, G.W. Equity and social justice. Journal of Social Issues, 1975 31 (3) .

Weiss, C. H. Evaluation research in the political context. In M. Guttentag & D. Struening (Eds.), The Handbook of Evaluation Research. Beverly Hills: Sage, 1975.

Williams v. Florida. 399 US 78, 1970.

Wilson, D. W., & Donnerstein, E. Guilty or not guilty? A look at the simulated jury paradigm. Journal of Applied Social Psychology, 1977, 7, 175-190.

Wolf, R. L., & Arnstein, G. Trial by jury: An evaluation method. Phi Delta Kappen, 1975, 57, 185-190.

Woodward, J. A., & Overall, J. E. The significance of treatment effects in ordered category data. Journal of Psychiatric Research, 1977, 13, 169-177.

Zeisel, H. Twelve is just. Trial, 1974, 10 (6), 12.

Zeisel, H., & Diamond, S. S. Convincing empirical evidence on the six member jury. University of Chicago Law Review, 1974, 41, 281.

Zeisel, H., & Diamond, S. S. The effect of peremptory challenges on jury and verdict: An experiment in a federal district court. Stanford Law Review, 1978, 30 (3), 491-531.

Zeisel, H. & Diamond, S. S. The jury selection in the Mitchell-Stans conspiracy trial. American Bar Foundation Journal 1976, 1, 151-174.