

1989


Are Twelve Heads Better Than One?

Phoebe C. Ellsworth

University of Michigan Law School, pce@umich.edu

Available at: <https://repository.law.umich.edu/articles/121>

Follow this and additional works at: <https://repository.law.umich.edu/articles>

 Part of the [Courts Commons](#), [Criminal Procedure Commons](#), [Law and Psychology Commons](#), [Law and Society Commons](#), and the [Litigation Commons](#)

Recommended Citation

Ellsworth, Phoebe C. "Are Twelve Heads Better Than One?" *Law & Contemp. Probs.* 52 (1989): 205-24.

This Article is brought to you for free and open access by the Faculty Scholarship at University of Michigan Law School Scholarship Repository. It has been accepted for inclusion in Articles by an authorized administrator of University of Michigan Law School Scholarship Repository. For more information, please contact mlaw.repository@umich.edu.

ARE TWELVE HEADS BETTER THAN ONE?

PHOEBE C. ELLSWORTH*

I

INTRODUCTION

Few advocates of the jury system would argue that the average juror is as competent a tribunal as the average judge. Whatever competence the jury has is a function of two of its attributes: its number and its interaction. The fact that a jury must be composed of at least six people,¹ with different backgrounds, experiences, and perspectives, provides some protection against decisions based on an idiosyncratic view of the facts. Not only must the jury include at least six people, but they must be chosen in a manner that conforms to the ideal of the jury as representative of community opinion.² The jury's competence, unlike that of the judge, rests partly on its ability to reflect the perspectives, experiences, and values of the ordinary people in the community—not just the most common or typical community perspective, but the whole range of viewpoints. Representativeness is important not only for ensuring “the essential nature of the jury as a tribunal embodying a broad democratic ideal,”³ but because it affects the jury's competence directly. Failure to assure that any given group has a fair chance of participation “deprives the jury of a perspective on human events that may have unsuspected importance in any case that may be presented.”⁴

Copyright © 1989 by Law and Contemporary Problems

* Professor, Law School and Department of Psychology, University of Michigan.

This research was supported in part by a grant from the National Science Foundation panel on Law and Social Science.

1. *Ballew v. Georgia*, 435 U.S. 223, 228 (1978) (five-member jury does not satisfy the jury trial guarantee of the sixth amendment).

2. *Strauder v. West Virginia*, 100 U.S. 303 (1880) (equal protection clause guarantees the defendant that the state will not exclude members of his race from the jury venire on account of race); *Thiel v. Southern Pacific Co.*, 328 U.S. 217 (1946) (in an exercise of its supervisory power, the Court granted a new trial where jury had been chosen in a manner which systematically excluded persons who were wage workers); *Taylor v. Louisiana*, 419 U.S. 522 (1975) (fair cross-section requirement of the sixth amendment is violated by the systematic exclusion of women); *Duren v. Missouri*, 439 U.S. 357 (1979) (state statute providing for automatic exemption from jury service for any woman requesting not to serve which produces jury venires averaging less than 15% female violates the sixth amendment's fair cross-section requirement); *Batson v. Kentucky*, 476 U.S. 79 (1986) (prosecutor's use of peremptory challenges to exclude blacks from a jury trying a black defendant establishes an equal protection claim of purposeful discrimination). These cases are a small selection of major landmarks in the development of the definition of representativeness as a fundamental characteristic of fair juries.

3. Ellsworth & Getman, *Social Science in Legal Decision-making*, in *LAW AND THE SOCIAL SCIENCES* 596 (L. Lipson & S. Wheeler eds. 1986).

4. *Peters v. Kiff*, 407 U.S. 493, 503-04 (1972) (plurality opinion of Marshall, J., joined by Douglas and Stewart, JJ.).

A jury decision, however, is more than an average of the verdict preferences of six or twelve citizens who represent a variety of experiences. Ideally, the knowledge, perspectives, and memories of the individual members are compared and combined, and individual errors and biases are discovered and discarded, so that the final verdict is forged from a shared understanding of the case. This understanding is more complete and more accurate than any of the separate versions that contributed to it, or indeed than their average. This transcendent understanding is the putative benefit of the deliberation process.

There is a good deal of evidence that jurors commonly organize the mass of testimony they hear into a story⁵ with characters, motives, and plot. Incoming information is assimilated to the basic story framework, and information that does not fit is often forgotten or explained away. It is well known among psychologists that much of what is perceived is a function of the perceiver; it is a particular *construal* of the events perceived, rather than a true reflection.⁶ Several different perceivers will come up with several somewhat dissimilar accounts of a sequence of events. Once having arrived at a construal, or a story or explanation of the same sequence of events, most people find it very difficult to imagine a different way of interpreting the same events, and this leads them to underestimate severely their own creative contribution to their "memory."⁷ Even though most people recognize in principle that a good deal of perception is really interpretation, they are unable to make adequate inferential adjustments for this process in guiding their own behavior, often behaving exactly as they would if their interpretation were the only possible one.⁸

If it does nothing else, group deliberation (except in extraordinarily one-sided cases) forces people to realize that there are different ways of interpreting the same facts. While this rarely provokes a prompt revision of their own views, it necessarily reminds the jury members that their perceptions are partly conjectural—an obvious truth, but one that is otherwise unlikely to occur to them. A judge does not have this vivid reminder that alternative construals are possible.

A judge, however, has experience on the bench and training in the law. Critics of the jury often focus on the incompetence of people chosen as jurors, compared to that of the judge. At best the venire consists of a representative sample of the community, with a few members having genuine expertise, a

5. See W. BENNETT & M. FELDMAN, *RECONSTRUCTING REALITY IN THE COURTROOM* (1981); Pennington & Hastie, *Evidence Evaluation in Complex Decision Making*, 51 *J. PERSONALITY & SOC. PSYCHOLOGY* 242 (1986); Pennington & Hastie, *Explanation-Based Decision Making: Effects of Memory Structure on Judgment*, 14 *J. EXPERIMENTAL PSYCHOLOGY: LEARNING, MEMORY & COGNITION* 521 (1988).

6. See D. SCHNEIDER, A. HASTORF & P. ELLSWORTH, *PERSON PERCEPTION* (1979); Ross, *The Problem of Construal in Social Inference and Social Psychology*, in *A DISTINCTIVE APPROACH TO PSYCHOLOGICAL RESEARCH: THE INFLUENCE OF STANLEY SCHACHTER* (N. Grunberg, R. Nisbett & J. Singer eds. 1987).

7. D. Griffin, D. Dunning & L. Ross, *The Role of Construal Processes in Overconfident Predictions about the Self and Others* (1988) (unpublished manuscript, Stanford University).

8. *Id.*

large number who are simply average citizens, and a few others who are distinctly below average. In practice, many of the better-educated jurors are excused from service, and others who show knowledge or ability relevant to the particular case at trial may be challenged during the voir dire. Attorneys select jurors for incompetence.⁹ Thus, some have argued that the average jury is not only less competent than the average judge, but is also less competent than a random sample of twelve citizens from the community.

Pointing out the discrepancy in the qualifications of the average juror and the average judge is one common tactic used to criticize the jury. A second is to point to particular cases, usually highly publicized cases, where the jury verdict seems inappropriate. When a John Hinckley is acquitted or a \$10 billion damage award is set, many citizens are outraged and look for someone or something to blame. When the judge decides, typically they blame the individual judge. When the jury decides, they blame the jury system (or occasionally the governing law). The practice of drawing conclusions about an institution on the basis of a tiny handful of highly publicized cases is, of course, scientifically and logically unsound, though no doubt inevitable.

Historically, the debate over the competence of juries has been less enlightening than it might have been. In particular, there are two conspicuous omissions. First, there is a great reluctance to define competent decisionmaking. Social scientists who turn to the legal literature in search of criteria by which to evaluate the jury are likely to find it a frustrating experience. It is extremely difficult to design research that will contribute useful information to the debate on competence when the concept of competence is not defined. Second, most of the social science research and much of the legal debate has focused primarily on the jury's verdict, an extremely crude measure of competence, and one that tells us very little about what juries actually do.

Among social scientists, one way to look at jury functioning is to break down the jury's task into components, and look at the way the jury deals with each one. Pennington and Hastie¹⁰ have provided a useful list:

- 1) The jury members must "encode" the information they get at trial. A competent jury must pay attention to the testimony and remember it.
- 2) The jury must define the legal categories. A competent jury should define these categories as they are presented in the judges' instructions.
- 3) The jury must select the admissible evidence and ignore evidence that is inadmissible.
- 4) The jury must construct the sequence of events.
- 5) The jury must evaluate the credibility of the witnesses.

9. See J. VAN DYKE, *JURY SELECTION PROCEDURES* (1977); V. HANS & N. VIDMAR, *JUDGING THE JURY* 63-78 (1986).

10. Pennington & Hastie, *Juror Decision Making Models: The Generalization Gap*, 89 *PSYCHOLOGICAL BULL.* 246, 249-55 (1981).

- 6) The jury must evaluate the evidence in relation to the legal categories provided in the instructions. That is, certain elements of the story the jury constructs are particularly important in determining the appropriate verdict. The jury must identify these elements and understand how differences in the interpretation of the facts translate into differences in the appropriate verdict choice.
- 7) The jury must test its interpretation of the facts and the implied verdict choice against the standard of proof: preponderance of evidence, clear and convincing evidence, or beyond a reasonable doubt.
- 8) The jury must decide on the verdict.

This list is meant to provide a set of reminders of what a jury must do to accomplish its task. There is no assumption that the jury performs these tasks in the order listed, nor that it should. No decisionmaking group indefatigably sticks to the task throughout the course of its discussions, and juries are no exception. In discussing my research on jury deliberations, I present data and some impressions of how the jury performs these tasks; I also discuss some other aspects of jury deliberation.

The research itself involved the close analysis of the first hour of deliberation of eighteen mock juries.¹¹ Because of the small sample size, statistical analysis of the data generally would be misguided. The study is most usefully considered as an intensive case study of the process of jury deliberation, although the fact that there are eighteen cases rather than one makes it considerably more useful than the usual case study, because it allows for some assessment of the variability of juries exposed to the same stimulus. A major drawback is that none of the juries reached a verdict in the hour allotted to them.¹² Thus, the study is most useful as an exploration of how juries structure their task, how well they deal with the facts and the law, and what things they discuss. It is very likely that at some point juries move into an "end game" that may differ substantially from the phases preceding it.

II

METHOD

A. Subjects and Overview

Two hundred and sixteen adults eligible for jury service in Santa Clara or San Mateo County, California, participated in the deliberation study and

11. The deliberation data were collected as part of a study of the relationship between death penalty attitudes and perceptions of guilt. Results relevant to this issue are reported in Cowan, Thompson & Ellsworth, *The Effects of Death Qualification on Jurors' Predisposition to Convict and On the Quality of Deliberation*, 8 *LAW & HUM. BEHAV.* 53 (1984).

12. Five of the juries had reached 10-2 splits by the end of the hour, as assessed by post-deliberation questionnaires on verdict preferences. The trial used in the research was the same one reported in R. HASTIE, S. PENROD & N. PENNINGTON, *INSIDE THE JURY* (1983). They found that twelve-person juries operating under a unanimity rule took an average of two hours and 18 minutes to reach a verdict.

provided usable data.¹³ Thirty-three of them were recruited from the venire lists of the Santa Clara County Superior Court after completion of their terms as jurors. Of the remaining 183 subjects, 156 had responded to a classified advertisement in local newspapers asking for volunteers for a study of "how jurors make decisions." The final twenty-five people were referred by friends who had heard of the study or who were subjects themselves. Each subject was paid ten dollars for participation.

The sample was fairly representative of the suburban upper-middle-class community surrounding Stanford University, except that males and minorities were underrepresented. The sample was 93 percent white and 65.3 percent female. The average age of the subjects was forty-three, and 63 percent of the sample was employed outside the home. Married persons constituted 45 percent of the sample; 26 percent were single, 19 percent divorced, 4 percent separated, and 6 percent widowed. The median educational level was slightly less than a college degree. Democrats made up 46.2 percent of the sample, Republicans 32.4 percent, and Independents 11.6 percent, with the remainder divided among unregistered voters and small parties. In the category of religious affiliation, one-third of the sample listed themselves as Protestant, 15 percent as Catholic, and 9 percent as Jewish; 26 percent listed no affiliation, and 17 percent listed other religions. Finally, 46 percent of the sample had previously performed jury duty, while 37 percent had actually served on juries.

Subjects watched a videotape of a simulated homicide trial that represented all major aspects of an actual criminal trial. After hearing the evidence, arguments, and instructions, the jurors gave an initial verdict. Jurors were then assigned to twelve-person juries and allowed to deliberate for one hour.

B. The Trial Videotape

After viewing simulated trial materials prepared by several other social scientists interested in jury behavior, and after considering creating materials of our own, we chose the videotape prepared by Reid Hastie for use in his research on jury unanimity.¹⁴ This tape is representative of the procedures, setting, style, and issues that commonly occur in actual homicide trials. The case was complex enough to afford several plausible interpretations and verdict preferences. It resembled most real murder trials in that there was no question that the defendant had killed the victim; rather, the evidence centered on the precise sequence of events preceding the killing and on the defendant's state of mind at the time. Finally, the tape was far more vivid and

13. The study included 20 juries, but two had to be dropped from the analysis, one because of equipment failure in the sound recording, and the other because one of its members was an amateur actor who had recently starred in a production of *12 Angry Men* (Orion-Nova/UA 1957) and who dominated deliberation using arguments and reasoning drawn from that play.

14. R. HASTIE, S. PENROD & N. PENNINGTON, *supra* note 12.

realistic than any others we have encountered. It was highly unlikely that we could have constructed a better tape with our resources.

Hastie's videotape is a reenactment of an actual homicide case based on a complete transcript of the original trial. Each actor portraying a defense or prosecution witness was provided with "a summary of the case highlighting his or her testimony."¹⁵ The judge and the lawyers, portrayed by an actual judge and two experienced criminal attorneys, were given "unabridged copies of the actual judge's instructions, selections of relevant testimony, and the actual attorney's opening and closing arguments as they were originally presented."¹⁶ The attorneys were asked to develop their cases as they would for a real trial. The witnesses were asked to review their materials to get their version of the events firmly in mind. Then, for the actual taping, all actors put aside their materials and engaged in "spontaneous improvisations closely following the original case."¹⁷

We modified the tape in two ways for the present research. First, we shortened it slightly by deleting one defense witness whose testimony added little. Second, we replaced the segment of the original tape containing the judge's instructions, which had been based on Massachusetts law, with a new sequence in which the applicable California law was given. The new instructions were derived from *California Jury Instructions, Criminal* (CALJIC, 1970)¹⁸ and were assembled in consultation with trial attorneys and law professors. Professor William Keogh of Stanford Law School portrayed the presiding judge in the new sequence. The version of the tape we used lasted two and one-half hours, including one-half hour of judge's instructions. Pretesting indicated that the tape was regarded as convincing and realistic.¹⁹

In the trial videotape, the defendant, Frank Johnson, is charged with first-degree murder for the stabbing of Alan Caldwell outside a neighborhood bar. The prosecution brings evidence that the defendant and victim had argued in the bar earlier that day, and that Caldwell had threatened the defendant with a straight razor. Johnson had left after the argument, but had returned with a friend that evening. Johnson was carrying a fishing knife in his pocket. Caldwell later came into the bar, and he and the defendant went outside and began to argue loudly. Two witnesses testify that they saw Johnson stab down into Caldwell's body. The victim's razor was subsequently found folded in his left rear pocket.

15. *Id.* at 47.

16. *Id.*

17. *Id.*

18. COMMITTEE ON STANDARD JURY INSTRUCTIONS, CRIMINAL, CALIFORNIA JURY INSTRUCTIONS, CRIMINAL (3d ed. 1970).

19. Aside from the successful use of this tape by Hastie, Penrod, and Pennington, whose study provides substantial evidence of its realism, the pretest evidence comes from seven attorneys, a judge, two law professors, and about 30 nonprofessionals to whom we showed the tape before we ran the actual study. Although we took no formal ratings, none of the pretest subjects found any flaws in the film, except to point out that the defendant looked a little old for his stated age and that a real trial would have more delays. See R. HASTIE, S. PENROD & N. PENNINGTON, *supra* note 12.

For the defense, Johnson testifies that he had returned to the bar that evening on the invitation of his friend and had entered only after ascertaining that Caldwell was not there. Caldwell had come in later and had asked Johnson to step outside, presumably for the purpose of patching up their quarrel. Once outside, Caldwell had hit him and had come at him with a razor. Johnson had pulled out a fishing knife which he often carried in his pocket and Caldwell had run onto the knife. Johnson's friend corroborates much of this testimony. In cross-examination, the defense attorney casts doubt on the ability of the prosecution's eyewitnesses to see the scuffle, and shows that medical evidence cannot establish whether the defendant stabbed down into the victim or the victim ran onto the knife.

Four verdicts are possible in this case, depending upon the jury's findings of the facts. The defendant may be guilty of first-degree murder, of second-degree murder, of voluntary manslaughter, or he may be not guilty for reason of self-defense or accidental homicide.

C. Procedure

The study was conducted on weekend afternoons at Stanford University. Each subject group consisted of twelve to thirty-six subjects. Upon their arrival, all subjects were shown to an auditorium, asked to fill out an informed-consent form and a preliminary questionnaire, and given a brief overview of the study. The questionnaire focused on background (demographic) characteristics, general attitudes toward the death penalty and toward criminal defendants, and general attitudes with respect to crime control and due process.²⁰ The experimenter then introduced the trial videotape:

Now we would like to show you a videotape reenactment of an actual criminal trial. This trial took place in Boston, and nothing has been changed, except, of course, for the names. You will be asked to reach a verdict based on the facts of the case and the law the judge explains to you, just as you would if you were an actual juror. We would like you to pay close attention to the testimony and the judge's instructions and to try to take the case as seriously as you would if you were actually serving on a jury. While you are listening to the case, you should not communicate with anyone else, and you should not take any notes, because we want your experience to be as much like that of a real juror as possible.

During the two and one-half hours while the subjects were watching the videotape, the experimenters assigned the participants to twelve-person juries. All jury assignments were random, subject to the following constraints: (1) juries could not contain members who were acquainted; (2) juries could not contain more than one student member; and (3) numbers of male jury members and of persons recruited from the jury list were roughly equalized across juries.²¹

20. See Fitzgerald & Ellsworth, *Due Process vs. Crime Control: Death Qualification and Jury Attitudes*, 8 LAW & HUM. BEHAV. 31 (1984); Packer, *Two Models of the Criminal Process*, 113 U. PA. L. REV. 1 (1964).

21. One-half of the juries included from two to four members who would be ineligible to sit on real juries in capital cases due to their attitudes against the death penalty; the other half did not.

As soon as the videotape was over, the experimenter asked the subjects to indicate their verdict preferences on an initial verdict questionnaire by checking one of four choices: first-degree murder, second-degree murder, manslaughter, or not guilty. After collecting the initial verdict questionnaires, the experimenter read off the jury assignments and directed the members of each jury to a separate room for deliberation. The rooms were seminar rooms, each containing a table with ample space for five people along each side and one at each end. A videotape camera was placed behind the person sitting at one end of the table, and two ceiling microphones were used to record the audio track. The camera and microphones recorded the deliberations for later analysis. The equipment also allowed the experimenters to view the deliberations on a monitor outside the deliberation room, in order to detect problems that might jeopardize the validity of the study.

Once the subjects were settled in the jury room, the experimenter told them that their next task was to discuss the case and try to reach a verdict. They were assured that their immediate postvideotape verdict was confidential and that they need not feel committed to it. They were also told that most juries began by taking a straw vote, and that in any case they should choose a foreman before beginning their deliberation. The experimenter continued as follows:

As you discuss the case, it is important to put yourselves into the role of jurors. Imagine that you are a real jury and that your verdict will actually determine the fate of the defendant you saw on the tape. We want you to make your decision only on the basis of what you saw on the tape. Although the characters in the trial you saw were actors, we want you to treat them as if they were real. In short, we want you to make the decision you would make if you were a real jury and if you had seen in court exactly what you saw on the tape.

The experimenter closed by informing the subjects that they had one hour in which to deliberate, and that they should try to reach a decision in that time, although quite possibly one hour would not be long enough to reach a consensus. The purpose of this instruction was simply to assure that the subjects worked on their deliberation seriously and tried to reconcile their differences of opinion. We did not ask them to take a vote at the end of the hour, and we did not expect them to reach a verdict.

Subjects were then left to discuss the case. Although they appeared to be slightly self-conscious in the presence of the recording equipment for the first minute or two, the jurors became highly involved in the discussion and seemed to forget about the camera as soon as the deliberations revealed disagreements among the members, which, for each jury, occurred almost immediately. After an hour the experimenter returned, stopped the deliberation, and handed out the postexperiment questionnaires.

The videotaped jury deliberations were transcribed, and the transcripts were divided into units. In devising the coding scheme, I identified thirty

Differences between these two types of juries are not discussed here. See Cowan, Thompson & Ellsworth, *supra* note 11.

major issues in the case. A unit, by definition, could contain no more than one of these issues. Short utterances occasionally contained none; long utterances were divided into units corresponding to the number of issues. Each transcript was coded by one or more of three trained coders.²² Each unit was coded for the general nature of the statement (issue, fact, law, vote, procedural comment, and so on), correctness, prodefense or proprosecution position, and the particular fact, issue, or point of law that was mentioned.²³ Coders met weekly with me to resolve questions and settle differences.

III

CHOOSING A FOREMAN

All juries began by choosing a foreman, not surprisingly, since the experimenter had instructed them to do so. The foreman was always chosen very quickly, with a minimum of discussion. For ten of the eighteen juries, the process of foreman selection can be summed up by the phrase "choose a man who says he has experience." Although 65 percent of the jurors were female, sixteen of the eighteen foremen were male.²⁴ On the jury composed of eleven women and one man, the man was chosen.

When the jurors had arrived in the room and settled in their seats, someone would point out that their first job was to choose a foreman, and then typically someone would ask, "Has anybody had any experience with this sort of thing?" A man would claim experience, and the other jurors would agree that he should take the job. Occasionally two men would claim experience and a brief "after you, Alphonse" discussion would ensue until one of them said, "all right, I'll do it." These two scenarios account for foreman selection in ten of the eighteen juries. Since we knew which of our subjects had actually served on real juries, we were able to find out whether the people chosen as foreman were actually more likely to have had prior jury experience than the other jurors. They were not more experienced: 39 percent of the foremen had served on juries, as compared with 36 percent of the other jurors, an insignificant difference. Thus, a foreman is someone who claims experience, not necessarily someone who has it.²⁵

22. The author is deeply grateful to John Krueger, Margaret Campbell, and particularly Michael J. Sandmire, who served as coders and helped to develop the coding system, and to Pamela Burke, who also worked on the coding scheme and devised ways to calculate reliabilities.

23. Coders were given lists of 110 case facts, 18 major issues, and 60 legal instructions (all included "other" choices). At various points during the coding, two coders were asked to code the same jury, in order to calculate inter-coder reliability. Considering all of the coding categories, coders made 62 decisions for each unit. Overall agreement ranged from 47% to 69%. For the major categories, rates of agreement were as follows: Issues (62-98%); Facts (72-80%); and Legal Norms (73-87%). All differences of opinion were resolved by group discussion.

24. This gender bias in choice of a foreperson has changed little over the last 40 years. See Strodtbeck, James & Hawkins, *Social Status in Jury Deliberations*, 22 AM. SOC. REV. 713 (1957). It occurs not only in mock-jury research, but in real trials. See Kerr, Harmon & Graves, *Independence of Multiple Verdicts by Jurors and Juries*, 12 J. APPLIED SOC. PSYCHOLOGY 12, 24-25 (1982); Note, *Gender Dynamics and Jury Deliberations*, 96 YALE L.J. 593 (1987).

25. It is possible, of course, that the men who volunteered interpreted the term "experience" as broadly covering experience leading committee meetings, work groups, club meetings, or other

On the remaining eight juries, five foremen (four male, one female) were chosen because they were sitting in one of the two end seats, and three (two male, one female) were individuals who had opened the discussion by volunteering for the position. Altogether, nine of the foremen were sitting at the head of the table, and four others were sitting in the chair right next to the head. Table position is by no means a subtle proxemic cue that exerts an unconscious influence on the jurors: In the majority of cases the jurors explicitly gave table position as their reason for their choice—"you should do it, you're sitting in the right place."²⁶

These data suggest that jurors give little consideration to their selection of foremen. They are generally given no information on what qualifications to look for, so they have little to guide them but their background knowledge and stereotypes of the jury, gained from the media and other sources. In the movies, the foreman sits at the head of the table. In addition, at the time that the foreman is chosen, most jurors may still regard their task as a relatively simple one, because the extent and depth of disagreement on the jury has not yet been revealed. In light of people's inability to imagine alternative interpretations of a set of facts they have already interpreted,²⁷ jurors probably overestimate the popularity of their own position and the degree of consensus on the jury before they begin deliberating.²⁸ Thus, they may not think it makes much difference who is chosen foreman, because they see the case as a straightforward one and do not anticipate serious disputes. Finally, since no disagreements have yet been revealed, it is likely that strong norms of courtesy prevail at the time that the foreman is selected. Once someone has been suggested, the others may think it is impolite to question his or her ability.

IV

STRUCTURING THE TASK

Once the foreman was selected, the juries took one of two approaches to the task. One-half of the juries began by taking a vote, roughly evenly divided between show-of-hands, secret ballot, and a go-round procedure where each juror states a position and says a little about his or her reasons for taking that position. The other half of the juries began by discussing the facts and issues

gatherings. We do not have sufficient background data on our jurors to test whether those chosen as foremen actually had more of this sort of broadly construed "experience" than the other jurors. In any case, no prospective foremen volunteered specific information about these other kinds of experience, and no jury members requested it.

26. See Strodtbeck & Hook, *The Social Dimensions of a Twelve-Man Jury Table*, 24 *SOCIOMETRY* 397 (1961).

27. See Ross, *supra* note 6; Griffin, Dunning & Ross, *supra* note 7.

28. See Ross, Greene & House, *The "False Consensus Effect": An Egocentric Bias in Social Perception and Attribution Processes*, 13 *J. EXPERIMENTAL SOC. PSYCHOLOGY* 279 (1977). The tendency to overestimate the popularity of one's own position is pervasive in social perception. One might expect it to be exaggerated among members of a jury, since they all know that they have seen the exact same trial. Thus, many potential plausible reasons for expecting differences of opinion are ruled out.

in the case. The judge's instructions contained a caution to the jurors not to become unduly committed to their position but to remain open-minded. A few jurors interpreted these instructions to mean that they should not begin deliberations with a vote.

Hastie²⁹ and his colleagues have proposed that when a jury postpones a formal vote, it is freer to raise issues and discuss them open-mindedly. When a jury begins by voting, people feel committed to the position they have publicly expressed, and spend their time defending their position rather than trying to understand the facts and the law. Our data generally support the findings of Hastie's research. Juries that postponed voting talked more than juries that began with a vote, spent more time talking about the important issues in the case, and brought out more facts.³⁰ One might hypothesize that juries that voted early would spend more time discussing the relevant law, because they would need to define the legal verdict categories before they could vote. This, however, was not the case. Early versus late voting did not predict amount of time spent discussing the law.

Whether or not a jury began by voting, it was quickly apparent to the members of the jury that they disagreed about the appropriate verdict. Although the jurors had watched a simulated trial, their differences of opinion were real. As soon as these disagreements emerged, the character of the deliberation changed. During the selection of the foreman there was an atmosphere of conviviality in the jury room, along with some degree of self-consciousness. A few jurors joked about the videotape camera. Once the discussion or an early vote revealed differences of opinion, the jurors became involved in the task. There were no more references to the camera and very few jokes of any sort. They kept their attention focused on the case. On the average 47 percent of their utterances (defined in terms of units),³¹ concerned the facts of the case (range: 28 percent-57 percent); 32 percent addressed the important contested issues (for example, defendant's state of mind, provocation, angle of the knife thrust, ability of various witnesses to see the crime) (range: 22 percent-38 percent); 21 percent dealt with the law and the judge's instructions (range: 10 percent-40 percent); and 7 percent were votes or discussions about calling for a vote (range: 3 percent-17 percent). These proportions are quite comparable to those found by Hastie, Penrod, and Pennington, whose juries saw the same case but deliberated to a final verdict. The criticism that juries approach their task in a frivolous manner receives no support from this study or from any other serious empirical research on the jury.³²

29. See R. HASTIE, S. PENROD & N. PENNINGTON, *supra* note 12; see also Hawkins, *Interaction Rates of Jurors Aligned in Factions*, 27 AM. SOC. REV. 689 (1962).

30. The nine juries who voted early voted within the first 36 units of deliberation (see *supra* text accompanying note 22). Of the nine juries that postponed voting, none voted earlier than Unit #254, and four did not vote at all during the hour of deliberation.

31. See *supra* text accompanying note 23.

32. See generally Hans & Vidmar, *supra* note 9.

Whether or not the jury began with a vote, the general progression of the deliberation moved from an emphasis on facts toward an emphasis on law. In juries that did not begin by voting, the initial discussion resembled a random walk through the facts and issues. A topic would be raised, discussed briefly, and replaced by a totally different topic, with little attempt to organize the discussion and no attempt to resolve the issues. These juries conformed very closely to Kalven and Zeisel's observation that "the talk moves in small bursts of coherence, shifting from topic to topic with remarkable flexibility. It touches an issue, leaves it, and returns again."³³ During the hour of deliberation, the important facts and issues would come up again and again, while trivial issues would be dropped, and new issues added. Typically, as an issue was examined and re-examined, there would be movement toward consensus.

For example, one of the most important pieces of evidence in the trial was the coroner's statement that he found the victim's razor folded up in his back left pocket. Had the victim been coming at the defendant with the razor, a self-defense scenario would have been very plausible. The defendant and his friend claimed to have seen the razor drawn; two other witnesses testified that they did not see the razor. Most juries raised this issue early and dropped it without fully considering the implications. In subsequent discussions, someone would raise the possibility that the victim somehow, in a reflex-like action, could have folded up the razor and pocketed it after he was stabbed, or that someone else (the policeman, the ambulance doctor, or a passer-by) might have picked it up and put it in the dead man's pocket. The jury would eventually conclude that these possibilities were far-fetched. This conclusion would lead them to agree that the victim never pulled the razor during the fatal confrontation. As a consequence, some juries would reject the possibility of self-defense and a few would turn their attention to the relevant question of the defendant's possible belief that the razor was drawn. In general, over the course of deliberation, jurors appear to focus more on the important facts and issues, come to a clearer understanding of them, and approach consensus on the facts.

In juries that began with a vote, the discussion tended to be slightly more organized. The average distribution of verdicts prior to deliberation was one for first-degree murder, two for second-degree murder, six for manslaughter, and two for not guilty. Although none of the juries showed exactly this pattern, most of them had a majority of votes in the two middle categories with outliers for not guilty or for both not guilty and first-degree murder. A common tactic was for the middle jurors to begin by asking the outliers to explain their deviant position, typically starting with the proponents of first-degree murder. Whether or not the jury began with a vote, however, issues were raised and dropped fairly unsystematically, then raised again; slowly, progress was made. Little by little, most juries resolved the issues of fact and

33. See H. KALVEN & H. ZEISEL, *THE AMERICAN JURY* 486 (1966).

spent an increasing proportion of their time on the central issue: the defendant's state of mind.

A. The Jury's Ability to Deal with the Facts

Kalven and Zeisel conclude that "the jury does by and large understand the facts and get the case straight."³⁴ The juries in our study spent more time discussing the facts of the case (47 percent of the units included references to facts brought out in testimony) than anything else. These were rarely purely factual statements. Most of the time facts were raised in connection with a contested issue ("If the officer could clearly identify Frank and saw that he was doing something so that he should say, 'Frank, don't do it,' then that shows that he had very good vision."), a reference to common sense or knowledge ("There's only one way that bone could have been struck and it had to be like that. You couldn't strike down like that and hit right here."), a hypothetical scenario ("Suppose that—that he gets up with the knife, and Caldwell has the razor in this hand, and Caldwell, who is what, 200 pounds, six feet, 200 pounds, lunges toward him."), or a reference to the law ("The razor was in that man's back left pocket . . . so it couldn't have been murder by reason of the man defending himself.").

Most of the juries managed to sort out the factual issues fairly well during the process of deliberation. Conflicting testimony (for example, about the angle of the knife thrust) was recognized as such, so that juries ended up correctly attributing different versions of the story to different witnesses. Questions regarding the distance and angle of vision of the various witnesses were generally resolved correctly, and errors of fact generally were corrected. None of the juries maintained an erroneous perception of an important fact after the hour of deliberation. Implausible suggestions generally were discussed and rejected, as in the case of someone else putting the razor in the victim's pocket.

Individual jurors tended to focus on testimony that favored their initial verdict preference: Testimony about the previous confrontation between the two men was generally raised by jurors who favored a murder verdict, whereas testimony that the victim punched the defendant immediately before the killing was generally raised by jurors who favored manslaughter or self-defense. This tendency is not a weakness, but rather a benefit of the deliberation process—the opportunity it affords for comparing several different interpretations of the events along with the supporting factual evidence. It may be seen as a healthy instantiation of the "counterbalancing of various biases"³⁵ cited by the Supreme Court as one of the advantages of the jury system.

Not all of the discussion of facts was focused on the important issues—digressions about the defendant's health or his family occurred, for example.

34. *Id.* at 149.

35. *Ballew v. Georgia*, 435 U.S. 223, 234 (1978).

Likewise, major issues were sometimes omitted, particularly when there were no members of the jury, or only one, who favored a verdict for which consideration of that issue was important. It should also be remembered that deliberation ended after an hour. On the whole, however, the data from this study support Kalven and Zeisel's conclusion that the jury does quite well at its task of finding the facts.³⁶

The case was, of course, a simple homicide case, of a sort that was likely to be familiar to the jurors from television and the movies, and the testimony only lasted about two hours. Thus, we cannot tell how well our juries would have done had they been faced with a much longer case involving unfamiliar technical issues.

For most of the juries in this study, discussion of the facts and issues dominated the first part of the hour. Among the juries that voted early, there was usually some discussion of the judge's instructions in order to arrive at the verdict categories, but the discussion was generally quite superficial. During the course of the factual discussions, the central issues of disagreement emerged, and jurors attempted to persuade each other. Agreement on the facts, however, did not lead to substantial agreement on the central issue of the case: the defendant's state of mind. Jurors tried to persuade each other that their construals of the facts were commonsensical. The discussions often became heated, few opinions were changed, and at some point (often, but not necessarily in connection with a vote), the jurors would turn to the legal definitions of the verdict choices for guidance.

B. The Jury's Ability to Deal with the Law

Juries worked hard to understand the law. They spent an average of 21 percent of their time discussing the judge's instructions, primarily during the latter half of the hour. Following the hour of deliberation, jurors were given an eighteen-question true-false test on elements of the judge's instructions. On average the jurors answered 11.7 of the questions correctly, a result not significantly different from random guessing. In the larger study of which these deliberation data are a part, there were also seventy-two subjects who saw the videotaped trial, indicated their verdict preference, and filled out the postexperiment questionnaires but did not participate in jury deliberations. The questionnaires revealed no differences between these subjects and those subjects who had deliberated as juries in understanding of the judge's instructions. On a postdeliberation multiple-choice test of factual issues, however, jurors performed quite well, answering correctly an average of 8.8 out of fourteen questions (since there were four response alternatives, 3.5 correct answers would be expected by chance). Jurors also performed better than those subjects who did not deliberate. These results suggest that the deliberation process works well in correcting errors of fact but not in correcting errors of law.

36. See H. KALVEN & H. ZEISEL, *supra* note 33.

An examination of the statements jurors made about the law during the course of their deliberations provides further detail in this gloomy picture. We coded all statements jurors made about the law as correct, incorrect, or unclear. Remarks were coded as correct even if they were incomplete. For example, the statement “first-degree murder involves premeditation” would be scored as correct. Statements were scored as incorrect if they were unambiguously wrong; for example, “second-degree murder involves premeditation.” Statements that were coded as unclear were usually statements about verdict-evidence relationships; for example, “If Johnson knew that Caldwell would be there, it’s premeditation.” While this statement is technically false—in that returning to a bar knowing that one’s enemy is there does not necessarily imply intent to kill—it was scored as unclear, because the juror could have meant that Johnson’s knowledge was a relevant consideration in determining premeditation. Thus, we were lenient in coding references to the law as correct, and we did not code statements as incorrect unless there was no plausibly correct construal.

Given this rather generous coding, we found 631 (51 percent) correct references to the law and 609 that were not correct (28 percent unclear; 21 percent definitely incorrect). Thus, only half of the references to the law were accurate, even when credit was given for partial accuracy. One-fifth of the references were clearly, seriously wrong. Whereas factual errors tended to be corrected during deliberation, errors of law were not corrected. Omitting from the analysis the unclear statements, we compared the proportion of times that an error was corrected by the end of the hour with the proportion of times that a correct statement of the law was replaced by an error. Considering only these instances where the jury changed its position, 52 percent of them involved replacing an erroneous response with a correct one, and 48 percent involved replacing a correct response with an erroneous one. The results are quite distressing, since they mean that the jury does not recognize the right answer when it hears it. Juries who have heard the right definition are as likely to reject it as juries who have heard the wrong one. The jury as a whole does not profit from the abilities of its best members when it comes to questions of law.

For example:

Juror A: Second-degree stated that it doesn’t have to be—he doesn’t have to premeditate that far in advance.” (scored as correct);

Juror B: If it’s not premeditated, it can’t be murder.” (scored as incorrect; the jury accepts this);

Juror X: Now if I got up and I walked over there and you hit me, as I was coming over, then you hit me, then I pulled a knife out and stabbed you, that’s manslaughter.” (unclear);

Juror Y: No. That’s self-defense.” (unclear);

Juror X: Yeah, that’s right. Self-defense would be manslaughter.” (incorrect);

Juror Z: It’s *involuntary* manslaughter.” (incorrect).

During the course of deliberation, jurors generally fought to defend their correct opinions of the facts but not their correct versions of the legal standards. Typically the most forcefully expressed position prevailed, whether or not it was correct. Most of the jurors' discussions of substantive law (that is, the definitions of the verdict categories) conveyed an impression of considerable uncertainty ("Was it . . . I think it was something about passion?"), and jurors who seemed confident about the law were often believed, whether or not their statements corresponded to the judge's instructions. Of the 1752 units across all juries that referred to the law, only seventy-five (4 percent) were error corrections. Only 12 percent of the 609 incorrect and unclear statements were corrected. Examining each jury's last definition of the four verdict choices during the course of the hour, we found that no jury was correct on all four of them. It appears that most jurors failed to absorb a great many of the judge's instructions and that the process of deliberation did not correct this problem. Hastie, Penrod, and Pennington³⁷ used different judicial instructions for the same case (and their juries deliberated until they reached a verdict); their results were similar to ours.

Further evidence that the jurors learned less than they should have from the judge's instructions comes from an examination of the frequency with which various aspects of the law were discussed during deliberation. The instructions most often discussed involved points of law that the jurors were very likely to have heard about before they heard the case; thus, there is a strong possibility that much of their discussion of the law was based not on the instructions they had heard from the judge but on prior knowledge.

For example, of the sixty possible elements of the judge's instructions, the one discussed most frequently was the definition of first-degree murder requiring premeditation and deliberation. Jurors were usually correct in their definitions (sixty-five correct statements, five incorrect, thirty-seven unclear). They very rarely went any further, however, in trying to define first-degree murder. While there were 180 references (including questions) to the requirement of premeditation and deliberation, there were only thirteen mentions of the definition of premeditation and deliberation, and only thirty-eight attempts to distinguish first-degree from second-degree murder (eight correct, four incorrect, twenty unclear, four questions, and two error corrections). These results suggest that much of the jurors' discussion of the law on first-degree murder may have been based on the well-known phrase "premeditation and deliberation," and did not benefit from the new information provided by the judge's instructions. In addition, one might argue that the disproportionate amount of time spent discussing premeditation was inappropriate in this case, since fewer jurors favored first-degree murder than any other verdict choice.

Likewise, the familiar phrase "heat of passion" was the most commonly discussed element of manslaughter and accounted for 125 units of which a

37. See R. HASTIE, S. PENROD & N. PENNINGTON, *supra* note 12.

third were incorrect or unclear statements. More surprisingly, “involuntary manslaughter” was raised in ninety-three units, and the references were correct in twelve of these, clearly incorrect in fifty-one, and unclear in eight others. Fifteen of the fifty-one errors were corrected. It is not surprising that most of the references were incorrect, since the judge had stated that the verdict category “involuntary manslaughter” was not relevant to this case. This particular error was highly consequential on two juries in which jurors who had voted for “not guilty” changed their votes to manslaughter because they were persuaded that self-defense was the same as involuntary manslaughter. The fact that “involuntary manslaughter” was discussed almost as much as “heat of passion” in relation to the manslaughter verdict provides further evidence that juries rely at least as much on legal knowledge gained outside the courtroom as they are on the judge’s instructions.

In regard to self-defense, nearly half of the jurors’ statements fell into the category “self-defense other,” meaning that they were references to the law of self-defense that did not correspond to any of the points raised in the judge’s instructions.

Another way of looking at juries’ use of the legal instructions to guide their decisionmaking task is to look at the amount of time they spent explicitly discussing the distinctions between the verdict categories. Of the 1285 references to the verdict choices, only 10 percent (128) addressed the distinctions between verdict choices. Of these, 26 percent were correct statements, 11 percent were definitely incorrect, 42 percent were unclear, and 21 percent were questions.

Table 1 summarizes the juries’ discussions of the verdict categories. Clearly erroneous statements were relatively rare, except in relation to the manslaughter verdict, where a fifth of all references to the law (including questions) were incorrect statements. If we include the unclear statements—statements that could conceivably be correct—the picture is considerably bleaker. (For example, the statement “If you kill your girlfriend it’s manslaughter” was coded as unclear, whereas the statement “It’s not manslaughter unless you kill your girlfriend” would be coded as an error.) For first-degree murder and self-defense, correct statements outnumbered incorrect and unclear statements by about four to three, but for second-degree murder and manslaughter, incorrect and unclear statements outnumbered correct statements.

Although most of the law discussed by the jurors involved the substance of the verdict categories, jurors devoted 7 percent of their discussion of the law to the reasonable doubt standard, and 10 percent to the judge’s instructions about the jurors’ duties. The juries were extremely accurate on reasonable doubt, both on the instruction that the jury should be convinced beyond a reasonable doubt, and on the instruction that if they were not convinced beyond a reasonable doubt of a given guilty verdict, they must find the defendant guilty of a lesser included offense or not guilty. Not one person on any jury, however, raised the question of the definition of reasonable doubt.

TABLE I
 JURORS' DISCUSSION OF THE VERDICT CATEGORIES

| | <i>Total Units</i> | <i>Correct Statements</i> | <i>Incorrect Statements</i> | <i>Questions</i> | <i>Unclear Statements</i> | <i>Error Corrections</i> |
|--------------|------------------------|-------------------------------|---------------------------------|------------------|-------------------------------|------------------------------|
| Murder 1 | 279 | 98 | 19 | 86 | 69 | 7 |
| Murder 2 | 293 | 70 | 41 | 90 | 82 | 10 |
| Manslaughter | 348 | 104 | 75 | 86 | 64 | 19 |
| Not Guilty | 365 | 142 | 51 | 84 | 72 | 16 |

Percentage of Correct and Incorrect Statements

| | <i>Correct Statements and Error Corrections</i> | <i>Incorrect and Unclear Statements</i> |
|--------------|---|---|
| Murder 1 | 37.6% | 31.5% (6.8 incorrect, 24.7 unclear) |
| Murder 2 | 27.3% | 42.0% (14 incorrect, 28 unclear) |
| Manslaughter | 35.3% | 40.0% (21.5 incorrect, 18.5 unclear) |
| Not Guilty | 43.2% | 33.6% (14 incorrect, 19.6 unclear) |

Like "premeditated murder," the phrase "beyond a reasonable doubt" is one that is likely to be familiar to jurors from prior experience, so we cannot conclude that they learned this standard from the judge's instructions. Attempts to apply the reasonable doubt standard to the facts of the case were evenly divided between correct applications and applications that were incorrect or unclear (for example, "The doctor wasn't clear about the angle of the knife going in so I gave him the reasonable doubt from first-degree to second-degree," scored as unclear). The reasonable doubt standard was almost always raised by jurors who were trying to persuade a harsher faction to move toward their position. After extensive argument over the facts and issues had failed to persuade people, some members of the more lenient faction would turn to the reasonable doubt standard as a persuasive device.

Procedural instructions were also used as arguing tactics. Of the 172 remarks made about jurors' duties, 114 were devoted to three of the eleven instructions given by the judge: that jurors should only be influenced by the evidence and law presented in court (forty-nine remarks); that jurors should not speculate about sustained objections (twenty-two); and that jurors should not consider the penalty or consequences of the verdict (forty-three). These comments were also used primarily as a weapon to close off lines of argument that a juror disagreed with, and generally took the form "We can't speculate about that," or "We're not allowed to consider that." Jurors applied these rules correctly thirty-nine times, and were clearly incorrect forty-five times (for example, "We can't speculate about what [the defendant] was thinking or what he wasn't thinking"). Only fifteen of these forty-five errors were corrected. A great deal of concern has been expressed about jurors' inability to disregard extra-evidentiary factors, appropriately so.³⁸ These data

38. See Kassin & L. Wrightsman, *The American Jury on Trial: Psychological Perspectives* 99 (1988); Dane & Wrightsman, *Effects of Defendants' and Victims' Characteristics on Jurors' Verdicts*, in *THE PSYCHOLOGY OF THE COURTROOM* 83 (N. Kerr & R. Bray eds. 1982); Carretta & Moreland, *The Direct and Indirect Effects of Inadmissible Evidence*, 13 *J. APPLIED SOC. PSYCHOLOGY* 291 (1983); Hans & Doob,

indicate, however, that jurors may also use the judge's cautionary instructions to stifle discussion of unpalatable, but clearly relevant, evidence.

V

CONCLUSION

In summary, the process of deliberation seems to work quite well in bringing out the facts and arriving at a consensus about their sequence. Errors are corrected, and irrelevant facts and implausible scenarios are generally weeded out, at least in deliberations over this relatively simple homicide. The juries also do a good job of gradually narrowing down discussion to the important issues. On the whole, however, the discussion of the facts does not produce changes in votes, since jurors' verdict preferences in the case were rarely a function of a clear mistake on the facts.

Changes in votes are likely to occur after discussion of the law.³⁹ Unfortunately, the jurors' understanding of the law was substantially inferior to their understanding of the facts and issues. Much of the jurors' discussion of the law revolved around phrases they were likely to have known before they heard the judge's instructions. The instructions may have been effective in reminding the jurors of terms they had heard before, but the instructions were not very effective in educating them in new areas, or even in focusing their attention on the meaning of the familiar terms.

This failure to apply the law correctly was by no means a failure to take the law seriously. Discussions of the law took up one-fifth of the deliberation time and were carried out with great intensity, frequently with an apparent sense of frustration. The jurors understood that a key aspect of their task was to interpret the evidence in terms of the appropriate legal categories. They struggled to do so, but often failed.

Members of two of the eighteen juries sent out to the experimenters for help with the instructions. In our study we did not give them any help, because we were comparing the relative abilities of two kinds of juries and did not want to introduce a confounding variable. In a real case, many judges would have provided assistance. An informal survey of fifteen local judges indicated that there was wide variability in judges' responses to requests for help. One judge said that he would actually try to explain the difficult instructions in plain English; a few provided the jury with a written copy or tape recording of the instructions when the jury retired to the jury room. At the opposite extreme, some judges said that they did not believe in interfering with the natural process of deliberation by explaining, or even reiterating, the

Section 12 of the Canada Evidence Act and the Deliberations of Simulated Juries, 18 CRIM. L.Q. 235 (1975); Sue, Smith & Caldwell, *Effects of Inadmissible Evidence on the Decisions of Simulated Jurors: A Moral Dilemma*, 3 J. APPLIED SOC. PSYCHOLOGY 345 (1973); Wolf & Montgomery, *Effects of Inadmissible Evidence and Level of Judicial Admonishment to Disregard on the Judgments of Mock Jurors*, 7 J. APPLIED SOC. PSYCHOLOGY 205 (1977); Note, *Social Science Findings and the Jury's Ability to Disregard Evidence Under the Federal Rules of Evidence*, LAW AND CONTEMP. PROBS., Autumn 1989, at 341.

39. See also R. HASTIE, PRESENTATION TO THE FELLOWS OF THE CENTER FOR ADVANCED STUDY IN THE BEHAVIORAL SCIENCES (1987).

instructions. The majority of judges said that they responded to requests for clarification either by rereading the questionable instruction or by rereading the entire set of instructions from beginning to end. The judges also agreed that very few juries requested further information about the law. The judges also said, however, that they did not inform the jury before deliberation that it was allowed to ask for clarification of the instructions.

There is no *a priori* reason to believe that the jurors' misunderstanding of the law is a function of their mental capacities. It seems more plausible that the system is set up to promote misunderstanding. Factors blockading the serious jury trying to perform its task include the following: the convoluted, technical, language;⁴⁰ the dry and abstract presentation of the law following the vivid, concrete, and often lengthy presentation of evidence; the requirement that jurors interpret the evidence before they know what their decision choices are;⁴¹ the fact that juries usually do not get copies of the instructions to take with them into the jury room;⁴² the lack of training in the law for jurors as part of their jury duty; the general failure to discover and correct jurors' preconceptions about the law; the failure to inform jurors that they are allowed to ask for help with the instructions; and the fact that those who do ask for help are often disappointed by a simple repetition of the incomprehensible paragraph.

Research on jurors' comprehension of judge's instructions is increasing, but there is still very little. We do not even know whether juries that ask for help do better than juries that try to muddle through on their own. Research on specific techniques for improving juror comprehension indicates that improvement is possible. At any rate, it seems profoundly unfair to criticize juries for failing to perform well a task that, by all the usual educational criteria, has been stacked against them.

40. See A. ELWORK, B. SALES & J. ALFINI, MAKING JURY INSTRUCTIONS UNDERSTANDABLE (1982); Charrow & Charrow, *Making Legal Language Understandable: Psycholinguistic Study of Jury Instructions*, 79 COLUM. L. REV. 1306 (1979); Severance & Loftus, *Improving the Ability of Jurors to Comprehend and Apply Criminal Jury Instructions*, 17 LAW & SOC'Y REV. 153 (1982).

41. Kassin & Wrightsman, *On the Requirements of Proof: The Timing of Judicial Instruction and Mock Jury Verdicts*, 37 J. PERSONALITY & SOC. PSYCHOLOGY 1877 (1979); Elwork, Sales & Alfini, *Juridic Decisions: In Ignorance of the Law or in Light of It?*, 1 LAW & HUM. BEHAV. 163 (1977); Prettyman, *Jury Instructions—First or Last?*, 46 A.B.A. J. 1066 (1960); V. Smith, *The Psychological and Legal Implications of Pre-trial Instruction in the Law* (1987) (dissertation submitted to the Department of Psychology, Stanford University).

42. R. NIELAND, *PATTERN JURY INSTRUCTIONS: A CRITICAL LOOK AT A MODERN MOVEMENT TO IMPROVE THE JURY SYSTEM* (1979).