Civil Juries and Complex Cases: Taking Stock after Twelve Years

Richard O. Lempert  
*University of Michigan Law School, rlempert@umich.edu*

Available at: [https://repository.law.umich.edu/book_chapters/455](https://repository.law.umich.edu/book_chapters/455)

Follow this and additional works at: [https://repository.law.umich.edu/book_chapters](https://repository.law.umich.edu/book_chapters)  
Part of the Civil Procedure Commons, and the Courts Commons

**Publication Information & Recommended Citation**


This Book Chapter 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 Book Chapters by an authorized administrator of University of Michigan Law School Scholarship Repository. For more information, please contact mlaw.repository@umich.edu.
Verdict
Assessing the Civil Jury System

Robert E. Litan
editor

The Brookings Institution
Washington, D.C.
CHAPTER SIX

Civil Juries and Complex Cases: Taking Stock after Twelve Years

Richard Lempert

TWELVE YEARS AGO, as the first Reagan administration was coming into office, it appeared that the civil jury, at least in complex cases, might be on the way out. The hostility of Chief Justice Warren Burger toward the civil jury was no secret and the circuit courts were split on the question of whether the Seventh Amendment guarantee of trial allowed an exception for complex cases. The issue was ripe for Supreme Court resolution. Moreover, a body of then-recent scholarship provided the Court with some historical justification for reading a complexity exception into the Seventh Amendment as well as with more modern policy arguments for eliminating the civil jury or dramatically altering its tasks in complex litigation. The Supreme Court did not, however, seize the moment, and the issue remains unresolved. Today most federal courts still feel obligated by the Seventh Amendment to try legal cases to juries no matter what their complexities as long as one party insists.

In failing to act, the Court was acting wisely, or so I argued in an article published in 1981. That article was based on the following premises: first, that the framers of the Seventh Amendment had important and enduring reasons for constitutionalizing the right to trial by jury in civil cases; second, that the claim that a complexity exception was implicit in the Seventh Amendment lacked adequate historical support; and third, that even if the Seventh Amendment contained no complexity exception, Fifth Amendment due process gave civil litigants the right to insist on bench trials if a judge could be expected to decide a litigant’s case.

Work on this paper was supported by the Cook and Cohn funds of the University of Michigan Law School. I would like to thank Shari Diamond, Phoebe Ellsworth, and Neil Vidmar for their comments on an earlier version of this chapter. I also thank Lisa Bernt for her research assistance and Gail Ristow for her careful reading of the final manuscript.
rationally and a jury, even with the aid of reformed procedures, could not.

I commended judicial inaction on the issue because at the time I wrote the empirical evidence was insufficient to determine whether some cases were so complex that only a bench trial was likely to yield a rational judgment. To prove this, I argued, meant that one would have to show: first, that it was possible to identify a set of cases so complex that juries did not deal rationally with them; second, that such failures of rationality were inherent in the institution of jury trial and not the result of mutable ways of treating jurors or developing cases for trial; and, third, that judges were likely to decide such cases more rationally than juries.\(^5\)

Twelve years have passed since I wrote my article. One might expect that we now have the data needed to determine whether the showings I argued for can be made. If so, perhaps the Supreme Court can give us a sound, empirically based resolution of the "complexity exception" issue that has for so long been on hold. This paper examines the research produced during the past twelve years to see if this is possible. It seeks to determine what we know now that we did not know then about the ability of juries to handle complex cases, about our capacity for improving that ability, and about the ability of judges to improve on jury performance in such cases. It asks whether there is an adequate empirical basis for concluding: that juries can or cannot cope with complex cases; that we can or cannot change the way jury trials are conducted so that rational jury decisionmaking will not be thwarted by complexity; and that judges can cope with complex issues that juries cannot master.

To avoid keeping the reader too long in suspense, let me say at the outset that the answer to each part of this question is no. Such research as has been conducted on these topics is either too flawed or too limited to provide answers with firm enough empirical foundations to justify Supreme Court reliance. Yet we have learned something. To anticipate the discussion that follows, the jury often appears to do surprisingly well in the face of complexity, particularly insofar as complexity is defined by length of trial and the introduction of massive arrays of evidence. We know little about the judge's capacity to cope with complexity, but what we do know gives us no reason to be confident that the judge will do better than the jury. We also lack the kinds of rigorous research needed to argue that reforms in case management or jury practice can solve perceived problems. We are at a point, however, where a number of reforms can be suggested with little risk that they will make things worse and considerable reason to believe they will
improve jury performance. Finally, theoretical developments in cognitive and jury psychology suggest a new perspective that we might wish to bring to bear in thinking about the complexity problem. I shall deal with these matters in turn.

Juries and Complex Cases

There are many dimensions to complexity, but one feature that stands out in the discussion of complex cases is protraction. The horrible examples in the literature on complex cases, that is those cases cited as clearly unsuitable for jury trial by those who would abrogate the right to jury trial, are cases that take a long time to try to a verdict. Trial length is important to the argument against jury trial because lengthy trials raise serious problems of juror memory; are associated with massive amounts of information for the jury to comprehend; mean that large numbers of jurors, including a disproportionate number of those most likely to be especially capable, are excused from jury service; and can impose hardships on jurors who do serve, hardships that in theory might interfere with juror performance by causing resentment.

Case Studies

Perhaps because protraction is seen as a central feature of complex litigation, psychologists in their studies of mock juries have seldom focused explicitly on the issue of jury factfinding in complex cases. Instead, most of what we have learned during the past decade about the jury’s capacity to cope with complexity is anecdotal; it is based on close attention to jury behavior in particular cases chosen because of their research convenience or celebrity. Table 6-1 identifies and summarizes important characteristics of cases reported in the literature that meet two principal criteria. First, the case had to be one that could be regarded as complex by virtue of either its length or its subject matter. Second, the case description had to focus largely on the jury’s performance, and the author’s assessment of this performance had to be based at least in part on interviews with some or all of the jurors who participated in the case. While the focus of this paper is on civil trials, cases involving criminal trials are included in table 6-1. Although the legal question regarding the right to jury trial in complex cases may, for constitutional reasons, have a different answer in the criminal context than it has in the civil context,
rationally and a jury, even with the aid of reformed procedures, could not.

I commended judicial inaction on the issue because at the time I wrote the empirical evidence was insufficient to determine whether some cases were so complex that only a bench trial was likely to yield a rational judgment. To prove this, I argued, meant that one would have to show: first, that it was possible to identify a set of cases so complex that juries did not deal rationally with them; second, that such failures of rationality were inherent in the institution of jury trial and not the result of mutable ways of treating jurors or developing cases for trial; and, third, that judges were likely to decide such cases more rationally than juries. 5

Twelve years have passed since I wrote my article. One might expect that we now have the data needed to determine whether the showings I argued for can be made. If so, perhaps the Supreme Court can give us a sound, empirically based resolution of the "complexity exception" issue that has for so long been on hold. This paper examines the research produced during the past twelve years to see if this is possible. It seeks to determine what we know now that we did not know then about the ability of juries to handle complex cases, about our capacity for improving that ability, and about the ability of judges to improve on jury performance in such cases. It asks whether there is an adequate empirical basis for concluding: that juries can or cannot cope with complex cases; that we can or cannot change the way jury trials are conducted so that rational jury decisionmaking will not be thwarted by complexity; and that judges can cope with complex issues that juries cannot master.

To avoid keeping the reader too long in suspense, let me say at the outset that the answer to each part of this question is no. Such research as has been conducted on these topics is either too flawed or too limited to provide answers with firm enough empirical foundations to justify Supreme Court reliance. Yet we have learned something. To anticipate the discussion that follows, the jury often appears to do surprisingly well in the face of complexity, particularly insofar as complexity is defined by length of trial and the introduction of massive arrays of evidence. We know little about the judge's capacity to cope with complexity, but what we do know gives us no reason to be confident that the judge will do better than the jury. We also lack the kinds of rigorous research needed to argue that reforms in case management or jury practice can solve perceived problems. We are at a point, however, where a number of reforms can be suggested with little risk that they will make things worse and considerable reason to believe they will
improve jury performance. Finally, theoretical developments in cognitive and jury psychology suggest a new perspective that we might wish to bring to bear in thinking about the complexity problem. I shall deal with these matters in turn.

Juries and Complex Cases

There are many dimensions to complexity, but one feature that stands out in the discussion of complex cases is protraction. The horrible examples in the literature on complex cases, that is those cases cited as clearly unsuitable for jury trial by those who would abrogate the right to jury trial, are cases that take a long time to try to a verdict. Trial length is important to the argument against jury trial because lengthy trials raise serious problems of juror memory; are associated with massive amounts of information for the jury to comprehend; mean that large numbers of jurors, including a disproportionate number of those most likely to be especially capable, are excused from jury service; and can impose hardships on jurors who do serve, hardships that in theory might interfere with juror performance by causing resentment.

Case Studies

Perhaps because protraction is seen as a central feature of complex litigation, psychologists in their studies of mock juries have seldom focused explicitly on the issue of jury factfinding in complex cases. Instead, most of what we have learned during the past decade about the jury’s capacity to cope with complexity is anecdotal; it is based on close attention to jury behavior in particular cases chosen because of their research convenience or celebrity. Table 6-1 identifies and summarizes important characteristics of cases reported in the literature that meet two principal criteria. First, the case had to be one that could be regarded as complex by virtue of either its length or its subject matter. Second, the case description had to focus largely on the jury’s performance, and the author’s assessment of this performance had to be based at least in part on interviews with some or all of the jurors who participated in the case. While the focus of this paper is on civil trials, cases involving criminal trials are included in table 6-1. Although the legal question regarding the right to jury trial in complex cases may, for constitutional reasons, have a different answer in the criminal context than it has in the civil context,
the problems that complexity poses for juror decisionmaking do not necessarily differ with case type.6

Serious problems arise from reliance on anecdotal case histories. Cases chosen for study may not be typical of the range of complex cases either in the nature and extent of their complexity or in the ways that judges, jurors, and lawyers reacted to that complexity. Moreover, the case study methodology is not designed to be reproducible. Different authors viewing different cases choose to emphasize different features, and where they do focus on the same issues, standards for judgment may vary. These problems are compounded in the current instance, for many of the studies collected in table 6-1 were reported by journalists and not by social scientists. Not only did these journalists wish to tell a good story, but they were seldom concerned with what their observations could tell us generally about jury behavior in the face of complexity. Nevertheless, I believe that much can be learned from looking at how jurors performed across a range of complex cases. Any conclusions reached on the basis of what we can learn from these cases studies must be tentative, but even tentative knowledge is better than a knowledge vacuum.

Table 6-1 summarizes important features of each of the cases described. Several of the cases are among the most highly publicized and celebrated cases that have been tried in this country in recent years. Other cases are less well known; and a number may be considered run-of-the-mill complex cases. None of the cases appears as formidable for a factfinder as cases like SCM Corp. v. Xerox Corp. or Zenith Radio Corp. v. Matsushita Elec. Indus. Corp., which are the kinds of cases most commonly cited by those arguing for a complexity exception to the Seventh Amendment, but were a complexity exception established, at least some of the listed civil cases might, upon motion of a party, be removed from the civil jury docket.

The case descriptions come from a diverse group of observers. The four cases attributed to the ABA's 1989 report come from a study designed by a team of social scientists and lawyers. The social scientists who were in charge of the data gathering and are, I presume, responsible for the descriptions presented, are Elizabeth Loftus, Jane Goodman, and Edith Greene, Ph.D. psychologists who have written extensively on issues relating to jury trial. Arthur Austin, who described the trial and retrial in the case of Cleveland v. Cleveland Electric Illuminating Company (C.E.I.), an antitrust case that pitted a city against a privately owned utility, was at the time of his study a Professor of Jurisprudence at Case Western Reserve University. Molly Selvin and Larry Picus, who
<table>
<thead>
<tr>
<th>Characteristic</th>
<th>ABA sexual harassment</th>
<th>ABA antitrust</th>
<th>ABA arson</th>
</tr>
</thead>
<tbody>
<tr>
<td>Court</td>
<td>Federal</td>
<td>Federal</td>
<td>Federal</td>
</tr>
<tr>
<td>Subject</td>
<td>Sexual harassment</td>
<td>Antitrust—price maintenance conspiracy</td>
<td>Criminal conspiracy to commit arson</td>
</tr>
<tr>
<td>Length of trial*</td>
<td>4 weeks</td>
<td>9 days</td>
<td>2 weeks</td>
</tr>
<tr>
<td>Sources of complexity other than length</td>
<td>3 years of behavior to sort out, 175 exhibits</td>
<td>Unraveling meaning of an intricate chain of business transactions; difficult instructions on conspiracy; use of business jargon</td>
<td>Conspiracy law, complex fact question; five-day recess in midtrial</td>
</tr>
<tr>
<td>Jury aids</td>
<td>Note-taking; each side limited to 25 hours of testimony</td>
<td>Special verdict form, Note-taking allowed; preliminary instructions</td>
<td>One juror took notes</td>
</tr>
<tr>
<td>Jury size</td>
<td>8</td>
<td>6</td>
<td>12</td>
</tr>
<tr>
<td>Inherent difficulty of evidence</td>
<td>Low</td>
<td>Moderate</td>
<td>Low</td>
</tr>
<tr>
<td>Jurors who completed college</td>
<td>Not available</td>
<td>0 (estimate)</td>
<td>3–5 (estimate)</td>
</tr>
<tr>
<td>Defensibility of verdict</td>
<td>High</td>
<td>High</td>
<td>High</td>
</tr>
<tr>
<td>Comment</td>
<td>None</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>
Table 6-1 *(cont’d)*

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>ABA trade secrets</th>
<th>Pennzoil v. Texaco</th>
<th>Cleveland v. C.E.I. #1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Court</td>
<td>Federal</td>
<td>State</td>
<td>Federal</td>
</tr>
<tr>
<td>Subject</td>
<td>Trade secrets and restraint of trade</td>
<td>Tortious interference with contract</td>
<td>Antitrust, attempt to monopolize</td>
</tr>
<tr>
<td>Length of trial(^a)</td>
<td>6 weeks</td>
<td>4½ months</td>
<td>8 weeks</td>
</tr>
<tr>
<td>Sources of complexity other than length</td>
<td>Highly technical testimony, subtle patent issues, law of tortious interference</td>
<td>Difficult financial analyses and business concepts</td>
<td>Difficult concepts</td>
</tr>
<tr>
<td>Jury aids</td>
<td>Preliminary instructions; limits on testimony; note-taking allowed</td>
<td>Special questions posed; juror questions allowed for part of trial</td>
<td>Special interrogatories</td>
</tr>
<tr>
<td>Jury size</td>
<td>6</td>
<td>12</td>
<td>6</td>
</tr>
<tr>
<td>Inherent difficulty of evidence(^c)</td>
<td>High</td>
<td>Liability low</td>
<td>Moderate</td>
</tr>
<tr>
<td>Jurors who completed college(^d)</td>
<td>2–3 (estimate)</td>
<td>1–3 (estimate)</td>
<td>0</td>
</tr>
<tr>
<td>Defensibility of verdict</td>
<td>Moderate(^e)</td>
<td>Liability: moderate; damages: low</td>
<td>Low</td>
</tr>
<tr>
<td>Comment</td>
<td>Yes(^h)</td>
<td>Yes(^i)</td>
<td>Yes</td>
</tr>
</tbody>
</table>

\(^a\) \(^{1}\) \(^{2}\) \(^{3}\) \(^{4}\) \(^{5}\)
<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Cleveland v. C.E.I. #2</th>
<th>U.S. v. DeLorean (criminal)</th>
<th>U.S. v. GAF Corp.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Court</td>
<td>Federal</td>
<td>Federal</td>
<td>Federal</td>
</tr>
<tr>
<td>Subject</td>
<td>Antitrust, attempt to monopolize</td>
<td>Criminal conspiracy</td>
<td>Stock manipulation (criminal)</td>
</tr>
<tr>
<td>Length of trial</td>
<td>11 weeks</td>
<td>4 months</td>
<td>6 weeks</td>
</tr>
<tr>
<td>Sources of complexity other than length</td>
<td>Difficult concepts</td>
<td>Potentially distracting videotape</td>
<td>Technical financial concepts</td>
</tr>
<tr>
<td>Jury aids</td>
<td>Copy of instructions allowed in deliberations; special interrogatories; preliminary instructions summarized issues and defined terms; written final instructions</td>
<td>Written instructions for each juror; note-taking</td>
<td></td>
</tr>
<tr>
<td>Jury size</td>
<td>6</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>Inherent difficulty of evidence</td>
<td>Moderate</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>Jurors who completed college</td>
<td>0</td>
<td>7</td>
<td>0–1</td>
</tr>
<tr>
<td>Defensibility of verdict</td>
<td>High</td>
<td>High^k</td>
<td>Moderate (hung)</td>
</tr>
<tr>
<td>Comment</td>
<td>Yes</td>
<td>Yes^a</td>
<td>Yes^a</td>
</tr>
</tbody>
</table>

^k High (hung)
Table 6-1 (cont’d)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Court</td>
<td>State</td>
<td>Federal</td>
<td>State</td>
<td></td>
</tr>
<tr>
<td>Subject</td>
<td>Violating California Corporations Code and fraud in selling bonds (criminal)</td>
<td>Toxic tort</td>
<td>Breach of contract, conversion</td>
<td></td>
</tr>
<tr>
<td>Length of trial</td>
<td>10+ weeks</td>
<td>4 months</td>
<td>7 days</td>
<td>8+ weeks</td>
</tr>
<tr>
<td>Sources of complexity other than length</td>
<td>Need to understand market in junk bonds and duties of buyers and sellers; evidence entirely circumstantial; multiple counts and plaintiffs</td>
<td>Technical (hydrogeological) conflicting expert testimony</td>
<td>Complex medical issues</td>
<td></td>
</tr>
<tr>
<td>Jury aids</td>
<td>Special interrogatories (poor), trial in phases</td>
<td>Jurors given trial notebooks containing background information; preliminary instructions on some issues; note-taking allowed; four-question verdict form</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jury size</td>
<td>12 Moderate</td>
<td>6 High</td>
<td>6</td>
<td>12 Low</td>
</tr>
<tr>
<td>Inherent difficulty of evidence</td>
<td>Moderate</td>
<td>High</td>
<td>Liability: moderately high; damages: low (?)</td>
<td>Low</td>
</tr>
<tr>
<td>Jurors who completed college</td>
<td>1</td>
<td>1</td>
<td>Not available</td>
<td>0–3</td>
</tr>
<tr>
<td>Defensibility of verdict</td>
<td>High</td>
<td>Low⁰</td>
<td>Liability: high; damages: low</td>
<td>Liability: high; damages: low</td>
</tr>
<tr>
<td>Comment</td>
<td>Yes⁰</td>
<td>Yes⁰</td>
<td>Yes¹</td>
<td>Yes¹</td>
</tr>
</tbody>
</table>

³ Jurys who completed college⁴ Inherent difficulty of evidence⁵ Jurors who completed college⁶ Defensibility of verdict⁷ Comment
Sources: For ABA sexual harassment, antitrust, arson, and trade secrets cases: ABA (1989).

a. In some cases the trial time was given in days or hours. These figures were converted into days or weeks by assuming five hours to a trial day and five days to a trial week.
b. The use of special questions or verdict forms had the potential to complicate the jury's task as well as to ease it.
c. Where a large amount of difficult, unfamiliar scientific information bore on the central issues in the case, technical difficulty is coded as high. Where technical or specialized information seemed somewhat easier to understand, or where full understanding seemed less crucial to correct decisionmaking because the evidence was not so central, because it was redundant with other evidence that was easier to understand, or because a vaguer understanding would suffice, difficulty is coded as moderate or low, depending on how these factors appeared to play out.
d. Where occupational but not educational information was given, I attempted to estimate the number of college-educated jurors on the basis of their occupations.
e. Jurors with good understanding guided the others, some of whom were confused.
f. Jurors requested but could not get a transcript of crucial testimony; they complained of having to reconstruct testimony from memory.
g. Without compromise forced by the holdout, the verdict would have been highly defensible.
h. The most able juror assumed the leadership role. The jury reached a compromise verdict because of one holdout. The lack of indexing made it difficult to find documents in the jury room.
i. The trial had two judges. The first made several mistaken evidentiary rulings that strongly hampered the defense case; the second posed a jury charge that undermined the defendant's position. Defense counsel failed to put on an expert witness to explain a key concept that the jury misunderstood and failed to offer the jury any evidence or argument on the damage issue. The plaintiff's attorney contributed $10,000 to the trial judge's reelection campaign (the judge's largest contribution). He also contributed $10,000 to the judge's administrative judicial superior.
j. The judge seemed biased against the city. The jurors were not properly concerned with the issue of the relevant market. The jurors' attitudes toward the size and identity of the parties affected the verdict. The jurors did not understand the concept of proximate cause and failed to limit certain evidence to impeachment. After an early straw vote, the jury did not function well. The jurors did not understand their instructions, which were given orally in a one-and-a-half hour lecture by the trial judge, but they did seem to understand the conduct testimony.
k. The jury hung five to one against the judge's apparent preferred verdict.
l. Unlike the first case, the judge provided the jury with a written copy of his instructions to use during its deliberations.
m. The judge did an excellent job in keeping things clear and being unbiased. The jury discussed matters thoroughly and systematically; juror mistakes were corrected by other jurors.

q. The jury interpreted its task in terms of guilt or innocence. The jurors attempted to deal with the evidence systematically and to look at the evidence from both sides. The verdict form had a "not determined" response category, which in effect decided the issue for the defense without the jury knowing it; neither the plaintiff's counsel nor the judge explained that a "not determined" answer was equivalent to a "no." Only one juror—the only college graduate—understood the implications of a "not determined" response. He did not explain this to the others because he liked the result.
r. The authors argue that most damage awards were excessive, but that is because defense counsel never explained that if the plaintiffs had asbestosis the disease would not progress equally in all plaintiffs. The authors also argue that one plaintiff's award was cut back because of his nationality, but the award was still substantial and fair compared with what others received. This plaintiff received the same amount in punitive damages.
s. The case lasted as long as it did because it was tried only four hours a day, four days a week.

There was an intrinsically resolved to settle. The defendant misunderstood the judge's instructions on damages and misinterpreted the facts. The judge allowed a questionable cause of action to be added, which opened the door to the punitive damages. Crucial evidence on damages was excluded by the trial judge. Not much evidence was offered in support of the plaintiff's punitive damage request, and nothing was asked of the defendant, who simply asked for a zero verdict. The jury incorrectly remembered an instruction on punitive damage.
described Charles Newman v. Johns Manville, a tort suit to recover for asbestosis, were researchers with the Rand Corporation’s Civil Justice Research Institute. The remaining trial descriptions are the work of journalists and free-lance writers, each of whom published in American Lawyer, and, in the case of Pennzoil v. Texaco, a Wall Street Journal reporter and a trial juror.

**Dealing with Complexity**

The first thing to notice about table 6-1 is the range of cases that are arguably complex, particularly if trial length alone indexes complexity. Corporate law violations, toxic torts, conspiracies, stock manipulations, sexual harassment allegations, claims under the antitrust laws, breaches of contract, and matters relating to trade secrets all may give rise to colorable claims of substantial complexity, and this is just a group of cases that happen to have caught the eye of courtroom observers. The point is not a small one. Even though the case for a complexity exception has been made in reference to cases that appear more complex than most of the cases summarized in the table, if a complexity exception were to be created, the potential slippery-slope problem would be substantial. The prospect of lengthy trials and conflicting expert testimony on specialized topics would make most of the cases in the table colorable candidates for the withdrawal of juries. In deciding whether to withdraw juries, substantial judicial discretion would have to be exercised. Even if a trial court stretched that discretion, an appellate court might well be reluctant to reverse given the prospect of an expensive, time-consuming retrial and the difficulties an objector would have in showing that a judge’s verdict was unreasonable.

With respect to technical difficulty, two of the cases in table 6-1 seemed to turn largely on evidence so specialized and esoteric that any nonspecialist might have considerable difficulty in understanding. These are the ABA trade secrets case, which involved highly technical testimony about subtle patent issues, and the W. R. Grace case, involving epidemiological evidence as well as conflicting expert testimony about difficult issues in hydrogeology. Next highest on this dimension is the liability issue in the Johns Manville case, which involved conflicting interpretations of pulmonary and lung-function tests as well as statistical evidence on the association between exposure to asbestos and the development of asbestosis. Cases I have categorized as moderate in their technical complexity are cases that involved unfamiliar business situations such as
the need to understand normal practice in the junk-bond market (an issue in Keating) or the need to understand economic concepts at issue in antitrust cases such as the need to understand the characteristics of a "relevant geographic market" (an issue in the C.E.I. case).

In those cases labeled low in difficulty, esoteric and unfamiliar evidence either figured less prominently in the issues the jury had to resolve or it should have been relatively easy for the jury to understand. Thus, in the sexual harassment case two psychiatrists testified for the plaintiff, but the defense called no expert witnesses and the plaintiff could have made a case for both liability and damages even without the psychiatric testimony. In the GAF stock manipulation trial, jurors had to understand how stock trading worked, a matter that most of the jurors in this case found to be difficult "new terrain," but this is the kind of information that many lay jurors would know and that should be relatively easy to explain. The fact that I have rated the evidence in a trial as low in difficulty does not, however, mean that jurors will understand it. In the C.E.I. litigation, for example, the concept "natural monopoly" figured prominently in both trials. Compared with other testimony on the issue of monopolization, this concept should have been easy for a jury to understand. Austin's interviews indicate, however, that in two trials only one juror, an alternate, adequately understood the defense's claim of natural monopoly.

One may conclude from looking at these cases that with some frequency trials confront jurors with evidence that only experts have no difficulty understanding. Moreover, even where the evidence should be comprehensible to a jury, jurors chosen in a particular case may not comprehend.

The situation may not be so bleak as this summary suggests, however. Methodologically, these case studies are reconstructions of deliberations, and after the trial (in some of these studies, weeks or months after) jurors may overstate the degree of confusion that existed in the jury room. Moreover, individual jurors who say that "no one understood" an issue may be speaking more for themselves than for others, since those who did understand may not have had the occasion to exhibit their understanding directly to their fellow jurors or, if they attempted to do so, their understanding may not have been clear to the others. In this respect the trials described in the ABA study are particularly important. The ABA study was conducted by trained researchers who both interviewed jurors from the cases they studied and observed the deliberations of alternate jurors who had been on the trial panels. The authors'
observed that in almost all cases the juries were led by their most competent members. Neil Vidmar reports a similar observation. In medical malpractice trials he has studied, the most influential jurors understood the evidence. These findings suggest that jurors with poor understandings of key evidence can be guided toward correct verdicts even if their misunderstandings of crucial evidence are never fully clarified. In *W. R. Grace* we see a perverse variant of this: the only juror who apparently understood the full legal implications of the way the jury answered a question on the verdict form kept quiet about what it meant because he liked the result the jury’s misinformed answer would yield.

**Understanding the Facts**

The jury’s problems in understanding are compounded by limited juror ability. As the table reveals, many of the juries had few or no members with college educations. Most of the jurors who worked outside the household had blue-collar or clerical jobs. Since education and occupation are correlates of juror competence, some of these juries may have had few people capable of providing intelligent leadership. However, this difficulty is not an intrinsic failure of the jury system. A number of commentators have suggested ways of dealing with this problem, including making it more difficult for the better educated to avoid jury duty or the sitting of full or partial blue-ribbon juries. Moreover, even without changing the conditions of jury duty, highly capable juries may be seated if the lawyers, perhaps urged on by the judge, cooperate and do not routinely exercise their peremptory challenges on those jurors most likely to understand the case. In the *DeLorean* trial, for example, seven college-educated individuals were on the jury, none of them missed a trial day or arrived late, and they performed at the highest level. Similarly, one of the judges interviewed by the ABA researchers reported that in a complex case where both sides wanted an intelligent jury, a highly capable jury was seated and three weeks of trial time was saved because the jurors took depositions home to read after the trial day was over.

A special problem of understanding arises when there is a conflict of expert testimony. The difficulty is that when the two sides provide different interpretations of a situation, a person who previously knew nothing about the issue may have little basis for choosing between them. This is a generic problem with expert testimony that exists almost apart
from its inherent degree of complexity. In these circumstances juries seem to rely heavily on other credibility cues (for example, does the expert seem like a hired gun?) as well as on the way in which the evidence fits in with the other elements of the parties’ stories. Thus, if as in the *Pennzoil* case a party seems to have acted improperly, the jury is less likely to believe an expert who suggests that technically there was no impropriety despite the appearance than it is to believe an expert who suggests that the action violated commonly understood business norms.

It should be recognized that juries commonly decide between conflicting nonexpert testimony on such grounds as well, and to some extent they are celebrated for this. Thus, when a jury hears an eyewitness who places the defendant at the scene of the crime and a defendant’s spouse who testifies that the defendant was home at the time of the crime, the jurors weigh the relative credibility of the two witnesses and how their testimony fits into a larger story. They may have no more rational basis for deciding between such witnesses than they do for deciding between two experts who reach opposite conclusions on matters about which the jurors were previously ignorant.

Yet the jury faces problems in dealing with conflicting but unfamiliar expert testimony, whether hard to understand or not, that it does not face to the same degree in dealing with the conflicting testimony of ordinary witnesses. The most important is that experts are selected by the parties to be convincing. In these circumstances, as Samuel Gross points out, the normal cues to credibility are likely to be misleading, for a low credibility expert who testifies in a way that is unlikely to be believed is not likely to be hired. Indeed, one might expect those who know their cases are weak to make the strongest efforts to find experts who appear credible, for if they could not they would be foolish not to settle. At the same time, one should not overdraw the distinction between conflicting expert testimony and the conflicting testimony of ordinary witnesses, for if experts are adept at appearing credible regardless of the credibility of their testimony, ordinary witnesses may appear credible or incredible for reasons that have little or nothing to do with credibility such as apparent confidence or class distinctions in speaking style.

A second problem that distinguishes battles of experts from the conflicting testimony of ordinary witnesses is that we feel that when ordinary witnesses tell conflicting stories, there is little anyone can do but decide instinctively, on the basis of credibility cues and consistency with other evidence, who is telling the truth. With experts one often has the feeling that if only the decisionmaker had sufficient expertise, a
correct judgment would be made. This feeling gives rise to proposals for such things as science courts or the use of expert panels to resolve the esoteric scientific issues that arise in litigation. To the extent that the feeling is justified, trying complex cases to juries or even to judges poses a special threat. This is a threat to legitimacy posed by the danger that clearly incorrect decisions will be reached and that these mistakes will become obvious. The threat can easily be exaggerated, however. Where experts differ in complex cases, the differences are usually sincere. Indeed, I would suggest that the more difficult the issue—that is, the harder it is for the lay person to decide which expert’s opinion is more credible—the more likely it is that both positions are reasonably maintained. Thus, in the W. R. Grace case, leading experts in hydrogeology differed on whether toxic wastes could pass beneath a river to contaminate a town’s wells. While one side is undoubtedly right, it is likely that hydrogeology today cannot tell us which one it is. In these circumstances a jury response downplaying the importance of the evidence, even though it is at the center of the proximate-cause issue, and focusing on other evidence that suggests contamination (how else can an exceptionally high incidence of leukemia be explained) and on the responsibility or irresponsibility of the defendants might be the best we can expect of a decisionmaker, even an expert decisionmaker.

Evaluating Jury Verdicts

The bottom line, and perhaps the best test of whether jury leadership or other factors alleviate individual problems of understanding in complex cases, is the quality of the verdicts that juries return. Here the case studies have an interesting tale to tell. It is summarized in the evaluation of the juries’ verdicts from the perspectives of the persons describing the cases. Note first that most jury verdicts seem defensible; they are close to the verdicts that judges would have rendered or they seem fair and reasonable. Also, in cases where the jury had to decide both liability and damages, if the verdict appears mistaken, the holding on damages is more likely to appear unsupported by the evidence than the finding of liability.

Where verdicts are of moderate or low defensibility, the complexity of the jury’s factfinding task is not ordinarily at the heart of the problem. This can be best appreciated if we look closely at cases with such verdicts.

First, consider the American Bar Association (ABA) trade secrets case. The verdict is labeled moderately defensible because the judge’s
comments indicate agreement (though not in so many words) with the jury’s finding for the counterclaim plaintiff on one count and with its findings for the counterclaim defendant on several other counts. The judge, however, would have found for the counterclaim plaintiff on two counts where the jury found for the defendant, and he would have granted greater damages than the jury did. Although the ABA trade secrets case is one of the most technically difficult of the cases in table 6-1, the jury’s failure to arrive at a more acceptable verdict does not seem to stem from that difficulty. Eleven jurors, including five of the six who decided the case and the six alternates who for purposes of the study deliberated to a decision, shared the judge’s verdict preferences. The unsatisfactory nature of the jury verdict stems from a sixth juror who appeared not to understand the judge’s instructions and forced a compromise by obstinately holding out for his preferred verdict. Apparently, therefore, the jurors did very well except for the stubborn, uncomprehending juror, a type found regardless of case complexity.

One wonders how such good performance was possible in such a complex case. A clue may be found in the lawyers’ comments about some of the most complex evidence. One attorney, suggesting that the jury probably never understood the complex chemistry underlying a zeolite production process, remarked that the plaintiff never attempted to explain the chemistry involved because the case did not require that knowledge. Another lawyer, commenting on laboratory reports and scientific progress reports filled with chemical equations that had been entered into evidence, noted that the reports may have been useful for their bulk because they made it clear how much experimental work had been done month by month. Thus one has a case that in large measure doesn’t seem comprehensible to a jury as well as jurors who admit to being mystified by some of the evidence. Yet the jurors’ verdict preferences, with one exception, seem eminently sensible because a scientist’s understanding of the evidence does not seem essential to the fair disposition of the case.

In the two C.E.I. trials that Austin observed, the jury verdicts were inconsistent, with the first jury hanging five to one in favor of the plaintiff City of Cleveland and the second jury finding for the defendant. This inconsistency might suggest jury irrationality, but the indictment is probably mistaken. As Austin points out, the defendant’s strategy changed between the two trials so that the different juries were hearing quite different presentations of the evidence. Moreover, one piece of evidence offered at the first trial, which the plaintiff had regarded as its “smoking gun,” was ruled inadmissible at the second.
Looking at the C.E.I. trials separately, I have rated the verdict at the second trial high on defensibility since the judge appears to have agreed with it, while I rated the hung-jury outcome at the first trial low on defensibility because the vote was five to one away from the direction the judge seems to have thought correct. While the jury's failure to understand some of the concepts and evidence relating to the antitrust issues may have played a part in the jurors' votes, their poor performance, if it was poor, seems attributable to less esoteric reasons. The most important is probably that the jurors failed to limit the plaintiff's smoking-gun evidence—testimony that C.E.I. had hired a lawyer to bring in his own name a suit designed to hamper the plaintiff's business—to the impeachment purposes for which it was allowed. The other factors to which Austin attributes the first verdict are the jury's failure to understand certain legal instructions and concepts such as the meaning of proximate cause and the fact that the jury took an early straw vote, after which its deliberations did not function well.

Thus, if the jury did not reach a defensible result in the first C.E.I. case, the failure does not seem to have resulted primarily from the case's complexity. Indeed, it is not clear that the case was so complex as to be beyond the ken of a jury. Rather, the first jury, having no college-educated members, seems to have been especially ill-equipped to understand what may in principle have been comprehensible evidence. The second jury, while still failing to understand certain concepts, seems to have better understood the evidence, although this jury too consisted largely of blue-collar workers, none of whom was a graduate of a four-year college. The two juries differed, however, in that the second had several members who had substantial occupational responsibility. The other failures of the first jury, the failure to give proper weight to a limiting instruction and confusion about such concepts as proximate cause, are failures that occur in simple cases as well as in complex ones.

In the W. R. Grace case, Mitchell Pacelle suggests that the jury performed poorly since a new trial was needed because of the jury's inconsistent answers to special questions and because in giving an allowed answer to one question the jury did not realize the implications of its answer. Pacelle, however, notes that the question that gave rise to the inconsistent answer was the poorly worded product of a day and a half of legal debating in the judge's chambers. Moreover, when the jury asked the court for help in understanding the question, the judge's remarks were as confusing as the initial question. The jurors, left to their own devices, came up with a reasonable, though incorrect, interpretation of
the question and their answer meant a new trial (almost immediately forestalled by a settlement) had to be ordered.

The jurors' other problem in the case was that they responded "not determined" to a special question asking when the W. R. Grace Company had polluted certain wells—not realizing that this answer meant the plaintiffs had not proved an element of their case and so could not recover. The jurors were led to this decision by a verdict form that invited them to answer the question "not determined" without explaining the consequences of such an answer and by the failure of the plaintiff's counsel—to the defense counsel's astonishment—to explain to the jury in closing argument the implications of a not-determined verdict.

The answer "not determined" does not, however, suggest jury incompetence. The W. R. Grace case, as appears from the table, contained perhaps the most technical evidence of any of the cases discussed. It would be difficult for any lay person, judges included, to decide between the conflicting expert testimony in the case, and even an expert might have felt that it was impossible to date the start of the contamination to a specific month and year, as the question put to the jury apparently required.

The Pennzoil case involved a suit by Pennzoil against the Texaco Corporation for tortiously interfering with an "agreement in principle" that Pennzoil had reached to take over the Getty Oil Company and to purchase the stock in Getty held by the Getty museum. The jury's verdict on liability was probably wrong, and its decisions on both actual and punitive damages appear hugely excessive. Yet the three available reports make it difficult to blame these errors on the jury or on the complexity of the case, even though the jury's apparent failure to understand certain key words—"indemnities" and "agreement in principle"—fostered jury mistakes.

The jury's failure to understand the implications of certain promises of indemnity that Texaco had given the Getty interests played a major role in its decision to award Pennzoil $3 billion in punitive damages, but this failure appears attributable at least in part, however, to the trial judge's decision to exclude evidence of other litigation brought by Pennzoil. This evidence, as Stephen Adler tells us, was the "only one good way for Texaco to convince the jury that Pennzoil could have sued the Getty interests if it had wanted to—indemnities notwithstanding." The problem caused by the exclusion of the "other litigation" evidence was compounded in the deliberations on punitive damages when the jury asked a question bearing on Texaco's possible responsibility for actions
of the Getty interests and the defense was content with a reply instructing the jury to reread the instructions. Defense counsel were content with this reply because they believed the instructions contained a sentence telling the jury that the defendant was responsible only for its own actions. The evening after the jury's inquiry, however, the defense counsel discovered that the instructions contained no such sentence, but the judge the next day refused to give the jury a supplementary charge to correct this deficiency.

The $7.53 billion awarded for actual damages was exactly the amount Pennzoil sought. This may well be $7 billion more than the evidence justified, assuming Pennzoil was correct in its liability claim, but it is hard to fault the jury on this issue or to attribute the excess to the complexity of the damage issue. Rather the jury's award reflects the testimony on the damage issue of three Pennzoil witnesses, one of whom had particularly high credibility, and Texaco's decision to put forward no evidence that contradicted this testimony. Thus when the jury deliberated, the only damage evidence before them tended to establish Pennzoil's full claim, and, as a juror pointed out in deliberations, the jury had been instructed to decide the case on the evidence.

On the liability issue it is similarly difficult to attribute the verdict, assuming it was mistaken, to the case's complexity or to the jury's inability to follow the evidence. Rather, jury errors, if they occurred, can be explained by judicial rulings on the admissibility of evidence, the court's charge, which according to both Adler and Thomas Petzinger favored Pennzoil in a variety of ways, and to certain strategic errors of the defense team, such as the failure to offer expert evidence about how the crucial term "agreement in principle" was understood in the business world. 27

Micro/Vest v. ComputerLand was a contract action brought primarily against William Millard, the chief executive officer of ComputerLand, and secondarily against the corporation. It was brought to recover on a clause in a loan agreement that allowed the noteholder to convert the debt into shares of ComputerLand stock, but the plaintiff also sued under a novel theory of conversion that allowed it to claim that dividends owed on the stock allegedly due the noteholders had been tortiously converted to the defendant's own use. This theory was important because it provided a basis for claiming substantial punitive damages. In the Micro/Vest case, unlike Pennzoil, the jury's verdict on liability appears to have been well-merited; indeed before the trial began defense counsel
sent his client a letter warning that “the case was indefensible” and that if he did not settle, he could be “exposed to huge punitive . . . damages.”

The award of punitive damages, over $125 million against the two defendants, seems excessive; indeed, it is unclear whether any punitive damages were justified. While part of the reason for the award appears to be that the jury “remembered” an instruction that was never given—that punitive damages had to hurt—this failure of memory does not appear to be the result of massive evidence—the trial on punitive damages followed the return of the general verdict and lasted only a few days—or to evidence that was particularly complex. Indeed, the jury’s mistake seems to be the result more of a lack of evidence than of a surfeit. The defense attempted to set an implausibly low value on ComputerLand’s net worth and presented no evidence on what Millard was worth. More important, the trial judge excluded evidence of a partial summary judgment that left open the issue the jury had tried; this evidence tended to refute the plaintiff’s contention that Millard was using the trial in bad faith to get out of honoring his note. Defense counsel said of this ruling, “It was a devastating setback. I literally sat down stunned.” The plaintiff, like the defendant, offered little evidence on the defendant’s net worth but was nonetheless allowed to refer to Millard numerous times as a billionaire and to give an estimate of the value of ComputerLand stock. Of course, the trial judge’s questionable decision to allow the plaintiff to add a theory of tortious conversion to its contract case made the punitive damage award possible in the first place.

The final case in which the reporter seriously questions the quality of the jury’s decisionmaking is the asbestosis case, Charles Newman et al. v. Johns Manville et al., described by Molly Selvin and Larry Picus. This is a case in which three insulators and the wife of an insulator claimed they developed asbestosis as a result of exposure to the defendant’s product. These four claimants were bellwether plaintiffs for a group of thirty cases that had been consolidated for trial in a federal district court. The authors point to a number of errors that seem to have infected the verdict, but, as with most of the other verdicts that have been described, it is hard to link the authors’ criticisms to a failure to understand the admittedly complex evidence in the case. Two of the authors’ complaints concern the jury’s failure to follow judicial instructions. One of these instructions was a limiting instruction regarding evidence admissible against only one of ten defendants; another was an instruction on permissible bases of damages. These kinds of failures are found in
simple cases as well as in complex cases, and no data suggest that their likelihood increases with the complexity of the case. In the particular context of the Manville case, the first error is one that as a matter of psychology may have been difficult or impossible for the jurors to avoid; the second may be the kind of justice-oriented nullification of the law for which some might value the civil jury—akin, for example, to jury awards that take account of the fact that parties will have to pay attorneys’ fees.

The third critique that Selvin and Picus make of the Johns Manville jury is that the jurors seem to have reduced the amount awarded one plaintiff because he was a Mexican national. The authors suggest the possibility of discriminatory motives. If these existed, they are not a result of complexity, but I doubt the existence of discrimination.32

Finally, in one respect the jurors misunderstood the scientific facts. They assumed in their deliberations that the asbestosis afflicting each of the plaintiffs would progress to the point where each would become as sick as the lead plaintiff. Given the nature of asbestosis, however, even if the jurors were correct in accepting the diagnosis of asbestosis as attested by the plaintiffs’ doctors, they were almost certainly incorrect in awarding damages with the expectation that each plaintiff would inevitably become seriously disabled by the disease. Again, however, it is hard to fault the jury for the error or to blame it on the complexity of the case. The jurors had been told that asbestosis was a progressive disease, and the defendant never put on a witness to tell the jury of the wide variability in its progression.

When Verdicts Are Mistaken

Considering the group of case studies, we see that complex cases often present evidence that jurors do not fully understand and that juries do not always get things right in reaching their verdicts. At the same time complexity, particularly as operationalized by protraction or large amounts of evidence, is not necessarily confusing. Even when the evidence is quite difficult to follow, jurors may understand much of it and reach verdicts that in retrospect appear both justified and rational. Moreover, when jury verdicts seem mistaken, it is difficult to attribute the mistakes largely to the complexity of the evidence the jurors encountered and to their difficulty in understanding it. Rather, as I read the case studies, erroneous verdicts seem to have two general sources: one is the kinds of
factors that can lead jurors astray in ordinary cases and the other is the mistakes of lawyers or of judges.

The most serious problem that jurors encounter in their efforts to get things right appears to be an inability to apply instructions correctly. Yet this difficulty exists whether cases are complex or simple. Moreover, we know how to increase comprehension substantially, but little effort has been made in this direction.

The difficulty that jurors have in understanding instructions poses, however, a special problem in complex litigation that has not heretofore been recognized. A standard way of attempting to aid jurors in complex cases is to break down the jury's decisionmaking task through special verdict forms, special interrogatories, and other such devices, and there is some evidence that jurors find these helpful. These devices are themselves instructions, however, and they carry with them the danger of misunderstanding. The best example of how well-meaning efforts to simplify the jury's task can lead to error occurred when John DeLorean, after his acquittal on drug charges, was tried in Detroit for federal fraud, racketeering, and tax evasion. He was acquitted even though three of the twelve jurors said after the trial that they left the deliberations believing that DeLorean was guilty. The reason for DeLorean's acquittal was a special verdict form that read in part:

You must remember at all times that the accused cannot be found guilty... unless you unanimously find beyond a reasonable doubt that he committed at least two acts of racketeering. You must not only unanimously agree as to which of the eleven specific acts of racketeering were committed, if any, but also which specific subpart of each alleged act was committed. Without such unanimous agreement, you must find the defendant not guilty [emphasis added].

The jury, which disagreed about whether DeLorean had committed the requisite two acts of racketeering, returned a unanimous verdict of not guilty on the racketeering charge because the special verdict form seemed to mandate an acquittal if the jury members could not reach unanimous agreement.

In somewhat less dramatic fashion, several of the less defensible verdicts returned in the trials listed in table 6-1 seem to have been similarly affected by the jury's difficulty in dealing with special questions or special verdict forms. In W. R. Grace, for example, the jury never knew the consequences of concluding "not determined" on the special
verdict form, and they misinterpreted the meaning of another of the special verdict questions. In *Pennzoil*, the wording of eight special questions that the judge posed for the jury was tilted against the defendant. The jury's difficulties in dealing with such devices may, however, be largely avoidable. In several of the cases in which instructions or special questions caused problems, the jurors were aware of a problem and sought clarification by questioning the judge, but the judge usually provided no specific help, instead calling their attention to all or a portion of the instructions that the jury had previously been given. More helpful judicial responses might have avoided error.

A second difficulty the jurors had in some cases was understanding esoteric facts when the parties offered conflicting expert testimony from apparently credible sources. But this too is a problem found in cases that are not generally thought of as complex since it appears associated with a conflict of experts and not with other features that make a case complex. Jane Goodman, Edith Greene, and Elizabeth Loftus questioned trial court judges about the difficulties that jurors had in complex cases. The judges cited as recurrent problems comprehension of medical testimony and the evaluation of damages in complex personal injury cases, particularly in situations when jurors were faced with "reconciling totally conflicting expert testimony from highly qualified medical witnesses." But this kind of conflict often occurs in what are otherwise run-of-the-mill tort cases, the routine fodder of the civil jury. Ironically, conflicting expert testimony, which juries are not well equipped to handle, often does not make a jury's verdict appear irrational since a decision for either party when experts cannot agree often appears reasonable. In *W. R. Grace*, for example, regardless of how the jury decided the contamination issue, an observer cannot conclude that the jury was mistaken on the evidence. The situation is similar with respect to the jury's finding in *Johns Manville* that each of the plaintiffs was suffering from asbestosis. As a juror discussing conflicting medical testimony in another asbestos case said, "the expert testimony was not a real factor in our decision, except in the very backhanded sense that it lent medical credence to any result." Problems attributable to misunderstanding do arise, and the jury can be wrong when it credits less reliable expert testimony over more reliable testimony. At its extreme, this is the so-called "junk science" problem. In none of the cases noted in table 6-1 did the jury seem to prefer less credible scientific evidence over more credible evidence. In the *Pennzoil* case, however, the jury on the liability issue accepted
what appears to be an unrealistic view of how the business world viewed an “agreement in principle” to consummate a multibillion-dollar merger.

The jury’s action in *Pennzoil*, however, brings us to the second factor that by my reading of the case studies is particularly salient when juries go astray in complex litigation: the mistakes of lawyers and judges. In *Pennzoil*, the judge, “with some exceptions . . . didn’t permit the lawyer-witnesses to testify on their understanding of the term ‘agreement in principle,’” and Texaco did not produce any expert witness to make the point that an agreement in principle is understood not to be binding. Thus, the jury’s mistaken perspective on this issue may plausibly be attributed to the rational weighing of the evidence presented rather than to a failure to comprehend or to appreciate the evidence.

Lawyer mistakes often seem to stem from underestimating the capacity of the jury. In three cases—*Micro/Vest*, *Pennzoil*, and *Johns Manville*—lawyers did not produce important evidence on damages, probably because they feared that to do so would concede a liability issue they were contesting. In other cases jurors noted, resented, and sometimes drew adverse inferences from attempts to appeal to their emotions. In the ABA trade secrets case, for example, the jurors were not impressed by the cross-defendant’s suggestion that a verdict for the cross-plaintiffs might cost 138 New Yorkers their jobs. In *Johns Manville* the jurors were scornful of the lead plaintiff’s wife’s reference to the fact that the next day was her forty-fifth wedding anniversary. The defendant’s attempt in *Micro/Vest* to show the humble background of their client seems to have done more harm than good. As one juror in *Micro/Vest* commented: “The whole case was on the contract, but they kept going back to what he [Millard] did as a child: selling newspapers and such. When you hear seven times what school someone went to, it gets a little old. They got away from what they were really there for.”

The set of complex cases I have examined seems to include more arguably mistaken verdicts than one would expect to find in an equivalent number of simpler cases. Given the nonrandom nature of the sample and the number of celebrated cases, it is impossible to say whether judges and lawyers are more likely to err in complex cases than in simpler ones, but it is certainly reasonable to think so. Also mistakes may be more consequential in complex cases than in simple cases because juries in complex cases may be more susceptible to being misled by error. Had the Texaco case been simpler, for example, the jury might not have needed opposing evidence to appreciate the unreasonableness of Pennz-
oil's demand. Complexity may therefore pose special difficulties for juries in part because of the special difficulties it poses for judges and counsel. As most judges and lawyers realize, complexity is a sign that special care must be taken.

The case studies, examined together, suggest several additional features that bear on a jury's likely performance in complex litigation. Jurors appear to work hard and to take their jobs seriously, sometimes to the point of reading important documents in the case word by word. This observation is confirmed by Gordon Bermant and his coauthors who in examining another set of protracted trials note, "Judges and lawyers are uniformly complimentary of the diligence of the juries in these cases." 47

-Lawyers seem often to seek less well-educated juries. 48 But when they do not do so, as in the DeLorean trial, the jury may include several members with a college education. This is consistent with the findings of a study by Joe Cecil, Allan Lind, and Gordon Bermant. 49 Twenty-two percent of the jurors they interviewed who had served in trials of twenty days or more were college graduates, a proportion that is only 10 percent less than the proportion of college graduates among interviewed jurors who had served in trials of six days or less.

-Juries often include individuals who understand material that most of their fellow jurors do not. These members are particularly important because, as the ABA studies indicate, juries in complex cases tend to follow the lead of their most competent members.

-In several cases where the jury verdict was problematic and the deliberations as reconstructed seemed flawed, an early straw vote was a potential causal factor. Conversely, in several cases that appear to have been well handled by the jury, the foreperson took care to avoid early polarization.

-In some of the trials the judge's attempts to aid the jury through such means as allowing note-taking, allowing the jury to take written copies of instructions to the jury room, furnishing trial handbooks, and the like seem to have helped.

-Length alone does not seem to lead to jury confusion. Juries seem to have few problems in trials that are long but that otherwise have no special sources of complexity. The ABA sexual harassment trial and the DeLorean trial are illustrative examples. Conversely, if conflicting technical evidence is presented, a long trial is not necessary for jury confusion. These observations are consistent with the findings of Cecil, Lind, and Bermant. They interviewed ninety-nine jurors who had served in federal trials lasting twenty days or more and eighty-one jurors who
had served in federal trials lasting six days or less.\textsuperscript{50} They report that 46 percent of the jurors in long trials found that evidence was difficult, but so did 29 percent of the jurors in short trials, a difference smaller than they expected and one that could result if longer trials are more likely than shorter ones to involve issues elucidated by scientific or technical evidence. Their conclusion, like my conclusion from reviewing the case studies, is that "concerns about the unique difficulty of the evidence in protracted civil trials may have been overstated."\textsuperscript{51}

To sum up, one must be wary of drawing firm conclusions from a nonrandom sample of cases studies. This is clearly true here, particularly since half the studies I cite were done by journalists not specially concerned with showing how juries deal with complexity as a general matter. Some tentative conclusions appear reasonable, however. First, juries confronted with technical information have problems understanding that information, and if there is conflicting expert testimony, jurors may have the feeling that they do not know whom to believe. Second, juries often seem able to find their way around such confusion and to arrive at appropriate verdicts. Third, when juries make mistakes in deciding complex cases, the mistakes seem more often due to mistakes in understanding judicial instructions or to the errors of the judge or lawyers than they are to the difficulty of understanding the implications of complex or massive amounts of evidence. Complexity may exacerbate the jury's difficulties with instructions or the degree to which juries are likely to be misled by the mistakes of others, but the cases we have considered do not reveal the existence of such effects. Overall, the sample of cases I have examined provides no empirical support for the claim that there is a denial of the due-process right to a rational decision on the evidence when juries are seated in complex civil cases. The failure to find a clear link between complexity and a denial of due process is consistent with aggregated data collected by Larry Heuer and Steven Penrod.\textsuperscript{52} They report that in 160 federal cases collected so as to oversample complex cases, judges were no more likely to disagree with jury verdicts in complex cases than they were in cases that were shorter or simpler on the law or facts.

\textit{Laboratory Research}

Turning from case studies to the psychologist's laboratory, we find either more relevant studies than we can deal with or very few. Most studies that bear on the quality of jury factfinding and the influence of
particular variables should apply to juries in complex cases in the same way that they apply to juries in simpler cases, and in this sense too many studies qualify. Few studies use as stimuli the factual settings of complex cases or seek to pose for mock jurors the kinds of special problems that are associated with complex cases, and in this sense there is little to canvass. To keep things manageable, I will review only the latter, more relevant, studies.

Before I do, a word about external validity is in order. For obvious reasons, there are no studies that expose mock jurors to the trial lengths or masses of evidence that are common in complex litigation. This, however, may not be as great a threat to our ability to generalize from controlled experiments as one might assume; the case studies suggest that long trials and massive arrays of evidence are not the most important sources of the special difficulties that jurors may encounter in complex litigation. One cannot assume, however, that these factors do not matter. Even if the difference between a taped trial of two to four hours and a trial of a few days does not shake our confidence in our ability to generalize from experimental findings to juror performance in ordinary cases—and the evidence indicates that it should not—the difference between a taped trial of the same length and an actual trial of six weeks to six months may mean that what we learn from a simulation is an unreliable guide to the way an actual jury would act after a long trial. We do not know what weight to give this concern. With stimuli less rich than those of the taped trial, concerns about external validity loom even larger. At the same time, before criticizing a study for not mimicking real-life situations, one must consider the point the study is designed to reveal and the justification for generalizing from it. Realism is not always necessary to generalization.

The first set of psychological studies of potential relevance to the special problems that confront jurors in complex cases concerns the way jurors deal with statistics. Statistical evidence is often found in complex litigation, and people without statistical training frequently find statistics hard to understand. Several researchers have looked at what mock jurors do when confronted with statistical evidence. William Thompson and Edward Schumann identify two fallacies that can deceive jurors when they are confronted with statistics about incidence rates, such as blood type prevalence, that link a defendant to a crime. One, which they call the “prosecutor’s fallacy,” is to think that the probability of a defendant’s guilt can be determined by subtracting the incidence rate of a matching characteristic from one. The other, which they call the “defense
attorney's fallacy," is to treat incidence evidence as irrelevant almost regardless of the rarity of the matching characteristic because at most it shows that the defendant falls into a larger group, one of whom is guilty. \textsuperscript{55}

In an experiment with student subjects, Thompson and Schumann found that when presented with evidence relating to the probability of a hair match about 25 percent of their subjects fell into one error or the other, and they were about equally divided between the two types of errors. On a jury, particularly a twelve-person jury, such error rates would not be of great concern because it is likely that there would be people present who understood how the evidence should be weighed and that these jurors could explain the weaknesses of the fallacious approaches. Indeed, even if only people making the two types of errors were present on the jury, discussion might reveal why neither position was correct. In a second experiment, however, subjects heard advocates arguing for interpretations of blood type evidence consistent with either the prosecutor's fallacy or the defense attorney's fallacy. In these circumstances, only 22 percent of the subjects rejected both arguments. One would expect actual jurors, most of whom do not have a college education, to do worse in dealing with such statistical arguments than college students serving as experimental subjects.

The relationship between falling prey to a fallacy and reaching an incorrect decision is, however, not clear. This is the somewhat ironic lesson of another study by E. L. Schumann and W. C. Thompson. \textsuperscript{56} In this study mock jurors watched a relatively realistic trial simulation on a four-hour videotape. The closing arguments they observed either ignored the blood-type evidence, included a fallacious prosecution argument, included a fallacious defense argument, or included competing fallacious arguments. Only the argument for the prosecutor's fallacy when presented alone had an effect; it increased the conviction rate from about 50 percent to 70 percent and the average estimated probability of guilt to 85 percent. In the other conditions both the rates of conviction and the estimates of average probability of guilt fell. The irony is that given the probability of guilt as estimated by control subjects who received no blood-type evidence, the blood-type evidence should have increased the average probability of guilt to above .90. Only those subjects who fell victim to the prosecutor's fallacy were close!

The tendency of the subjects in the Schumann and Thompson study to underestimate the probative weight of statistical evidence is not surprising. There are good theoretical reasons for expecting statistical
evidence to be less influential with factfinders than intrinsically less probative nonstatistical evidence.\textsuperscript{57} Other researchers have also found evidence about the statistical incidence of matching blood types to be relatively uninfluential, even when expert statistical testimony is presented explaining its implications.\textsuperscript{58}

Some complex cases may well involve incidence statistics of the type that Thompson and Schumann and others have investigated; or perhaps cases that contain such statistics should be considered complex regardless of their other characteristics. The statistics that are involved in most cases conventionally considered complex, however, are somewhat different in nature. They usually consist, as in toxic tort cases, of epidemiological statistics which turn on comparisons between exposed and unexposed populations, or, as in many antitrust and most sex-discrimination law suits, they are the statistics associated with regression analyses. I have been able to find only two studies that deal with such statistics.

Molly Treadway has explored the adequacy of juror intuitions in dealing with evidence such as the four-fold relative-risk tables that epidemiologists use to determine whether a particular condition causes a particular disease.\textsuperscript{59} She found that the intuitions of twenty-five subjects who were members of the Baltimore city jury pool were not good. Her subjects were asked to examine two tables, one of which showed a relative risk from exposure of 2.8 and the other a relative risk of 1.01, and to determine from each table (1) whether exposure to a substance increased a person’s risk of developing an abnormality, and (2) for any particular person exposed to the substance who had the abnormality, whether it was more likely than not that the substance rather than something else had caused the abnormality. Only 41 percent of the determinations accords with the answers reached through epidemiological analysis, and only two persons, or 8 percent of the respondents, made all four determinations in accordance with epidemiological reasoning. These results are not encouraging, but they do not adequately address the issue that concerns us. The reason for their inadequacy is that in litigation jurors are not confronted with relative-risk tables and asked for their best interpretation. Rather, the tables are presented by experts who explain what they imply. One would expect instructed jurors to do far better in understanding the implications of the data presented to them than jurors not so instructed.

Shari Diamond and Jay Casper, in a particularly good simulation study, exposed 1,022 Cook County jurors to a seventy-five-minute videotape of the damage portion of a suit brought under the Sherman
The jurors were informed that in the first phase of the trial the defendants had been found to have engaged in illegal price-fixing. As one aspect of this research, they explored the reactions of jurors to competing yardstick and statistical models, two common types of models used to establish damages in such cases. In the various conditions of the experiment, the models were counterbalanced by party and by the amount of damages they implied. The jurors returned individual verdicts following their exposure to the case and then deliberated in seventy-six-person juries until verdicts were returned.

The authors did not find that the jurors ignored the statistical expert or that the effects of his testimony were dwarfed by the effects of the more concrete yardstick testimony. Overall, jurors gave somewhat higher awards ($216,515 versus $200,813) when the plaintiff's expert presented a statistical model than when he used a yardstick comparison, but the difference is not statistically significant. The overall influence of the statistical evidence relative to the yardstick evidence appeared to be the result of two competing forces: the statistical expert was seen as more expert than the expert presenting the yardstick model, and this made him more convincing; but the statistical expert's testimony was seen as less clear than the testimony presented by the yardstick expert, and this made him less convincing. Jurors who found the statistical testimony to be similar to the yardstick testimony in clarity expressed a statistically significant difference in awards. When the statistical expert testified for the plaintiff, the mean award was $220,517; when he testified for the defendant, it was $168,223.

The deliberation process had a marked effect on damage awards, as the juries' verdicts averaged about 27 percent higher than the average of their members' predeliberation judgments. This was not due to the effect of outliers, for the correlation across juries of outlier preferences with final verdicts was relatively low while the correlation with each group's mean and median predeliberation preferences was strong. Among jurors the preferences of the foreperson were particularly influential. This is important when we recall from the case studies that juries seem to be most influenced by their most capable members, and we note the tendency for juries to select particularly capable members as forepersons. For example, in the Diamond and Casper study, the 13 percent of the jurors in the pool who had had both some postgraduate education and a statistics course accounted for 36 percent of those persons chosen as forepersons. Moreover, the predeliberation preferences of forepersons with both these characteristics had considerably more influence on final
verdicts than the preferences of forepersons who had neither postgraduate education nor a course in statistics. These findings should caution researchers about generalizing from the average individual response to statistical evidence to the response of juries, and they emphasize the way in which the composition of real juries in actual cases may affect how complex evidence influences deliberations.

Overall, Diamond and Casper provide the most resounding support for the capacities of juries dealing with complex issues that can be found in the scientific literature. They write

The responses to expert testimony we observe . . . suggest that jurors play an active role in assimilating and assessing testimony. Jurors did not simply adopt the view of a witness they rated high on expertise, using apparent expertise as a peripheral cue to conclude that the expert must be correct. Rather, consistent with deeper processing of information which produces attitude change when the listener is highly involved, the jurors appeared to consider and evaluate the content of what the expert was presenting, and were less likely to be persuaded if they did not feel they understood it.

This approach not only suggests active evaluation and perhaps even subtlety in dealing with expert testimony, but it also indicates the care jurors use in evaluating evidence to reach their decisions. When presented with complex statistical testimony, jurors were not simply over-powered by material they did not understand. Rather the persuasive force of such testimony appears to depend in substantial measure on the ability of the expert to express clearly the basis for the conclusions it is being used to support. Our results thus suggest that concerns about jurors' uncritical willingness to accept statistical evidence may be overstated.

Of course, concerns about external validity may limit the real-world implications of Diamond and Casper's research. It may be difficult to make statistical evidence clear when it is embedded in weeks of other evidence or, if trial length is not a problem, when the statistical testimony itself takes hours to deliver, when it is followed by a cross-examination that may take days, when there is competing statistical testimony of a similar type, when one side is using its experts to obfuscate, and when peremptory challenges have been used to strike jurors who know too much about statistics. But even in a worst-case scenario, in which these concerns correctly identify the reality of much modern complex litigation,
one may still reasonably conclude from the Diamond and Casper study that jury weaknesses in dealing with the kinds of statistical evidence most commonly associated with complex litigation are not inherent in the institution of the civil jury but are the result of the way in which complex jury trials are managed.

A second body of laboratory studies that may have special relevance to the quality of jury performance in complex cases comprises studies that focus on joinder. These studies matter because one factor that may make for complexity is the joining of parties or counts, as in *Johns Manville*, which involved four bellwether plaintiffs—who would have litigated separately had they not been part of a “case congregation”—and ten defendants, some of whom might not have been involved in every suit had the plaintiffs’ cases been tried individually.

Joinder can involve parties, causes of action (counts or charges), or both. Most of the research on joinder in the psychological literature involves the joinder of charges in criminal cases that would not be considered complex. Generally the research shows that a criminal defendant is disadvantaged when charges are joined, but the mechanism by which this occurs is unclear. Three possible mechanisms are: confusion of the evidence (evidence admitted on one charge is remembered as bearing on another charge); accumulation of evidence across charges (evidence admitted on one charge reinforces evidence relevant to another charge); and inferences about the defendant’s character (the jury characterizes the defendant with a criminal schema and views the evidence on each charge in that light). The studies designed to elucidate which if any of these mechanisms operate are often limited in their focus and yield inconsistent results.

These mechanisms might all operate in civil cases in which different causes of action are combined, but if they do it is not clear that they should systematically disadvantage one party or the other. In the studies dealing with charge joinder, incriminatory evidence always implicates the defendant. In civil cases, on the other hand, a plaintiff may introduce evidence on several counts suggesting that the defendant was responsible for the plaintiff’s injuries, but the defendant may offer evidence on the same counts suggesting that the harm the plaintiff suffered was the plaintiff’s fault. Also unlike those cases of criminal joinder that have been the subject of psychological research, the crucial evidence offered on one civil count may be admissible on another, as when a jury is charged with deciding whether a defendant is responsible under either a negligence theory or a strict liability theory.
To the extent that evidence in a civil case is admissible on only one count, as when evidence of a plaintiff's contributory negligence is admissible to rebut negligence liability but not strict liability, a jury's failure to apply instructions to limit the influence of the evidence to the one count would not be surprising. The difficulties posed by the added counts are not specific to complex litigation. If, however, the presence of alternative causes of action means that considerably more evidence is presented than when only a single cause of action is alleged and that the trial lasts much longer as a consequence, one might point to alternative causes of action as a source of complexity apart from the danger of legal or evidential confusion. But alternative causes of action are not usually regarded as substantially lengthening trials, and no investigation of the effects of alternative causes of action has tried to simulate what would otherwise be considered a complex case. Instead, it is party joinder that is usually seen as making potentially simple cases complex. Trying the cases of different parties together can involve numerous lawyers, dramatically increase the amount of relevant evidence, and require the jury to link different items of evidence to the cases of one or more of the parties.

Irwin Horowitz and Kenneth Bordens have attempted to study the effects of party joinder in a simulated complex case. They used as a stimulus a four-hour audio tape of a toxic tort trial in which the evidence was intentionally complicated and at times boring. The subjects were 396 jury-eligible men and women, who deliberated in 66 six-person juries. The case involved a large chemical company that had allegedly leaked effluent that entered the food chain and harmed four people who brought suit and unnamed others. Litigation of the plaintiffs' claims raised both negligence and state-of-the-art issues. The researchers were primarily interested in the difference between the verdicts when the cases brought by the four plaintiffs were tried together and the verdicts in these same cases when they were tried separately. Within these conditions, an additional variation contrasted a situation in which plaintiff A was an outlier with respect to the seriousness of her injury (suffering from a rare liver cancer) with a situation in which she was not (suffering from chloracne rather than cancer). The final variable was whether the jurors were told there were twenty-six or "many hundreds" of other victims or whether the existence of other victims was not mentioned.

Overall the data do not suggest that the aggregation of plaintiffs led to confusion. Only the plaintiff with the weakest case, a man who may have continued to eat fish after knowing the food chain was contaminated,
Civil Juries and Complex Cases

was helped by having his case aggregated with the others, and this effect may reflect rational information processing. If the question of whether the plaintiff was responsible for his illness was a close one, the fact that people who clearly weren't responsible for their condition suffered from the contamination is some reason to believe that the plaintiff's suffering was not his fault. The data also show that compensatory damages were not affected, as they should not be, by the presence of an outlier or by the number of other victims; that the greater the responsibility attributed to the defendant, the higher the awards; and that the earlier the date that the jurors thought the defendant should have known about the toxicity of the chemical it was discharging, the higher the awards. Where liability was found, punitive damages were greater if an outlier was part of the aggregate and if there were hundreds of other victims. These results also appear reasonable since punitive damages should reflect the amount and extent of the harm a wrongdoer does.

Two potentially disquieting notes in the study are first a finding that the presence of an outlier is associated with a higher proportion of findings of no liability and second that there is substantial variance across juries in their verdicts. The effect of the outlier in stimulating findings of no liability is associated in the taped deliberations with remarks attributing fault to the outlier. This appears to be a classic "just world" response. The fact that the three juries that blamed the outlier in this way also denied recovery to her coplaintiffs probably reflects the fact that evidence linking the defendant to the coplaintiff's injuries was arguably no stronger, and in one case was clearly weaker, than the evidence linking the defendant to the outlier's injuries. The variance in verdicts and awards is troubling but variance of this sort is probably not peculiar to jury trials. As the authors point out, similar inconsistencies have been found in judicial behavior when matters such as sentencing have been examined, and a recent study suggests that in medical malpractice cases jury verdicts are no less consistent than arbitrators' awards. Moreover, much of the inconsistency in liability verdicts is attributable to those cases where the irrationality of just-world thinking appears to be operating. Inconsistent damage awards may in large measure reflect the fact that the plaintiffs did not request specific amounts of money, so the juries did not have the kinds of anchors and supporting evidence that actual juries often have at trials. Whatever else is happening, the jurors seem not to be misled by the additional complexity of dealing with four plaintiffs rather than one.

Looking at the joinder studies together with the studies of how juries
use statistical evidence highlights a number of potential problems posed by the irrationality of some human decisionmaking. The studies do not, however, suggest that juries are likely to be less competent in complex cases than in other cases. Indeed, the more realistic studies, those by Diamond and Casper on statistical evidence and Horowitz and Bordens on joinder, show juries coping rather well with those features that make the decisionmaking tasks posed by the stimulus cases more complex than the decisionmaking tasks that juries encounter in more ordinary litigation.

Judges and Complex Cases

Even if the jury were substantially less able to deal with complexity than it apparently is, due process should not require the abrogation of the Seventh Amendment right to trial by jury unless judges can decide complex cases more rationally than juries can. There is, as we have seen, little systematic empirical evidence that relates to the competence of the jury in complex litigation; virtually none bears on the competence of the judge. Judges have not cooperated in studies of themselves as they decide complex cases; they seldom participate as subjects in simulation studies and have not done so in studies simulating complex litigation; and they seldom grant interviews in which they explain how they understood the evidence in complex cases. Judges do write opinions, but their opinions may provide little insight into the true bases of their decisions. Even where an opinion suggests a mastery of complex materials, the reader cannot know whether the judge has understood the subject, whether the judge after reaching a decision relied on a clerk to convey an impression of understanding, or whether the judge and/or the clerk simply copied large passages from the briefs of the side they favored. The evidence on how judges handle complexity is therefore fragmentary. What we can say is that there is no guarantee that a judge can do a better job than a jury; that there are a number of cases in which we would not want to force the parties to a bench trial; and that when the judge is unbiased and capable, parties may be particularly likely to opt for a bench trial, but if they don’t there is a better than average chance that the jury will be up to its task.

Linking Judge and Jury

In almost every case study we examined, where the defensibility of a jury verdict was evaluated as low or moderate, those who reported on the
trials noted mistakes made by the lawyers and/or the judge. Conversely, in those trials where the jury seemed to perform at its best, like the DeLorean trial, the judge seemed to have performed exceptionally well also. This has the ironic implication that were there an exception to the Seventh Amendment for complexity, it is likely that the quality of judge trials would be highest in cases where the jury was likely to do a good job as well, for competent judges can enhance the competence of juries over which they preside. To the extent that the case for a judge trial turns on the assumption of a competent judge, we are assuming the kind of judge most likely to preside over a competent jury.

The apparent link between judicial and jury competence means that if we look at cases where juries are confused or err, there is an above average chance that we will find a judge who is confused or mistaken as well. It is not necessarily true, however, that a judge who gives a jury poor guidance will perform equally poorly as a factfinder. A judge who confuses the jury with an unclear instruction might correctly interpret the law as a trier of fact. A judge whose questionable evidentiary rulings have prejudiced one side's case before the jury might nonetheless find for that side, perhaps influenced by the very evidence that he or she wrongfully excluded and the jury will not hear. By the same token, however, judicial mistakes may signal more general failings. A jury, even one hampered by judicial mistakes, may be better at factfinding than a judge who cannot adequately organize a trial or state a rule of law. Indeed, in some cases the jury's strength of numbers and experience may make it more able in finding facts than even an excellent judge. 77

Understanding the Evidence

We have no systematic evidence on the capacity of judges to deal with the kinds of evidence that make complex cases difficult, but anecdotal evidence of situations in which judges appear not to have fully comprehended scientific evidence is easy to find. To begin at the top, we have the example of the Supreme Court's opinion in Williams v. Florida, which took the results of a classic psychological experiment to mean the opposite of what they implied. 78 At the opposite end of the judicial hierarchy, Michael Saks and Richard Van Duizend, who read transcripts of a homicide case that raised definition of death problems, concluded that the prosecution was not as conversant with the medical facts as the defense. 79 Yet the judge, whom they interviewed, praised the prosecutor's preparation and was unimpressed by the defense.
Statistical evidence seems to present courts with recurring problems. Sometimes courts are too ready to receive statistical evidence because they do not realize its weaknesses. On other occasions courts are too reluctant to hear statistical evidence that is essential to understanding the facts before it. In Minnesota, for example, the state legislature had to pass a statute to overturn a Minnesota state supreme court ruling that statistical evidence relating to DNA matches was inadmissible even though such information is essential to estimate the probative value of a DNA match. Some years before in the same state, a trial court made the opposite error. In *State v. Carlson*, the court admitted misleading statistical evidence based on an inadequate scientific foundation to show the probative value of a hair match; a man was convicted and sentenced to prison for life perhaps because of it. Both a Federal District Court (*U.S. v. Massey*) and an Illinois Trial Court (*U.S. ex Rel. Di Giacomo v. Franzen*) made the same error. Other cases in which trial and appellate courts have had varying degrees of difficulty in correctly interpreting statistical evidence are documented by the National Research Council in a report it prepared on statistical evidence in courts.

The recent controversy over DNA evidence also reveals the difficulty that judges can have with scientific information. While one would expect rational decisionmakers to reach the same conclusions on the same evidence, court decisions on the admissibility of DNA evidence have differed, even though the evidence bearing on admissibility was similar. Indeed, in one case a court that refused to admit DNA evidence relied heavily on briefs that had been offered in a case that accepted it. Finally one occasionally encounters cases where judges admit that scientific evidence is beyond them.

The point of all this is not to establish the general inability of judges to cope with scientific evidence. The available anecdotal evidence does not allow us to reach any conclusion about the seriousness of the problems that judges face in dealing with the kinds of evidence that make cases complex or about the abilities of our nation's judges to understand and to decide correctly complex cases. Nor do we even know, except in the rare case, how judges go about deciding the factual issues in complex cases. The anecdotal evidence does mean that just as there are no guarantees that juries will understand the technical evidence in a complex case or decide such cases correctly, so there are no guarantees that judges will get everything right. Nor do we have any firm empirical basis for deciding whether judges will in some sense decide complex cases better than juries do over the long run. The best evidence we have on this
count is Larry Heuer and Steven Penrod's finding that judges are no more likely to disagree with jury verdicts when cases are complex than when they are not.87 This suggests that jury and judge verdicts are not likely to differ over the long run because of factors that distinguish complex cases from simpler ones.

**Making Rational Decisions**

In short we have no reliable empirical basis for saying that litigants are more likely in judge than in jury trials to receive decisions based on a rational evaluation of the evidence, nor can we reach the opposite conclusion. Perhaps we are on firmer ground when we simply note that modern psychology has demonstrated numbers of ways in which human decisionmakers act irrationally, if consistency in dealing with formally identical problems or scientifically rational models such as Bayes' Theorem are valid standards of rational decisionmaking.88 Moreover, people sometimes will be influenced by actions or conditions they are not aware of, and they may misattribute their decisions to factors that have not influenced them at all.89 Even knowing about these dangers does not necessarily provide protection from them. All people have limited capacities, which can lead to problems in understanding difficult, unfamiliar material. Both judges and jurors are human. As factfinders in complex cases, judges and juries are probably more like each other than they are different in dealing with the problems they confront.

Juries, however, may maintain one advantage over judges. This is that in some number of complex cases they are likely to be the fairer decisionmaker, both in appearance and in reality. The *Pennzoil* trial provides a vivid example of why a jury may appear to be and perhaps is a fairer decisionmaker. In *Pennzoil*, two days after Judge Farris was chosen to supervise pre-trial proceedings, Joe Jamail, plaintiff's lead counsel, donated $10,000 to the Farris reelection campaign. Until that time Jamail had given Farris $100.00. Jamail also gave $10,000 to the reelection campaign of the judge who was Farris's administrative superior. Farris, in turn, though one can never know for certain what influenced him, made a number of questionable pre-trial evidentiary rulings that severely hampered the case Texaco wanted to present.90

Consider also the work of Bermant and his coauthors. They interviewed lawyers in eleven protracted cases that could have been tried to either a judge or a jury. Their basic finding is not at all surprising: lawyers prefer to try their cases before the factfinder that gives them a
better chance of winning. All six attorneys who chose jury trials and said why they did so listed the identity of the judge as the main or the only reason. Four of these attorneys specifically referred to the biases of the judges assigned to hear the case, and bias may have motivated the two who simply gave the judge's identity as their reason for choosing a jury trial. In two of the four cases in which the parties could have forced a jury trial but agreed to a bench trial, the competence and fairness of the judge was mentioned as the most important reason, with one attorney specifically noting that he looked on jury trials as "buffers against incompetent judges." 91

Several lawyers interviewed by Bermant and his coauthors noted reasons other than bias or judicial competence for wanting or avoiding a jury trial. Chief among these other reasons was a sense of how a client or witnesses would appear to the jury. Two attorneys, both of whom had been forced by their opponents to try their cases to juries, said that they preferred bench trials when their cases were strong and jury trials when their cases were weak, implying that judges are more accurate factfinders than juries.

The comments reported by the Bermant team present a mixed picture. It seems that lawyers view jury trials in protracted cases (and perhaps in all cases) primarily as a protection against judicial bias. But they also see juries as responding to certain "human elements" that judges might ignore, and as less predictable than judges in the sense that a weak case does not necessarily mean defeat. At the same time, where a case is protracted and both lawyers have confidence in the competence and integrity of the trial judge, agreement to waive a jury trial may not be difficult to achieve. Whether the views these lawyers have of the relative merits of juries and judges is accurate is difficult to determine. A considerable body of research indicates that even when aspects of a case might appeal to the prejudices of jurors, unless the case is otherwise close on the facts, the evidence dominates. 92 One might similarly expect judicial biases to be tempered by the weight of the evidence. If so, the choice of a bench trial or a jury trial is more likely to affect the outcome when a case is close on the facts than when the evidence clearly favors one side. In these cases it is most important to avoid a biased judge, or, for that matter, a biased jury.

The right to trial by jury also provides some degree of protection against the managerial judge, a figure that has been both praised and damned in the scholarly literature. 93 The power of a judge bent on settling a case in accord with a personal sense of justice is considerable
even if the exercise of this kind of judicial authority may be questioned. In some cases managerial judges have used tactics that parties might perceive as coercive to achieve settlements the judges regarded as fair. It is simply a fact of modern litigation that judges who desire to mold cases to their preferences have substantial power to do so. If a judge who favored a certain result were sure to try the case if a suggested settlement was declined, the judge's power to coerce settlements would truly know no bounds. Thus a right to jury trial in complex cases may be necessary to ensure that the more basic due-process right of a fair trial remains.

To sum up, even if juries cannot resolve complex cases rationally, one cannot be sure that bench trials would improve the situation. Judges, too, make mistakes in dealing with scientific and specialized evidence. Examples of judges who get things wrong when confronted with difficult, unfamiliar evidence are easy to find. We currently have no empirical basis for saying that judges will ordinarily do better than juries in deciding complex cases, nor do we have an adequate empirical basis for reaching the opposite conclusion. In the situation where the judge is most likely to do well—that is, where the judge is known to be both fair and competent—limited anecdotal evidence suggests that the right to a jury trial is most likely to be waived. If there is a jury trial in such instances, other anecdotal evidence suggests that the jury system will function at its best.

If it is difficult to choose between bench and jury trials when rational factfinding is taken as a due-process requisite, it seems easier when the right to an impartial decisionmaker or the right to meaningful trial are the values that due process protects. Here it appears that situations arise where the right to a jury trial in complex litigation is an important guarantor of due process rights.

Procedural Reform

The third point that I made in my 1981 article was that even if juries as now constituted were incapable of rationally deciding complex cases, one should not read a complexity exception into the Seventh Amendment unless the jury's incapacities were inherent in the institution. Before giving up on the jury, I argued, we should be sure that we cannot increase jury competence through changes in the way we handle jury cases. Included among the possible changes I mentioned were rewriting jury instructions to make them more comprehensible, furnishing jurors
with written copies of the court’s instructions, allowing jurors pre-instructions, allowing jurors to deliberate on various issues as the case progressed, routinely seating twelve-person juries, furnishing jurors with daily transcripts (as an alternative to another possible reform—juror note-taking), dividing issues for decision, refusing party joinder, providing appointed experts to help the jury understand the testimony of the parties’ expert witnesses, removing factual controversies by encouraging stipulations, using masters to clarify particularly difficult issues, sitting partial blue-ribbon juries, and allowing jurors to ask questions. Many commentators have suggested similar reforms as ways of helping juries cope with complexity, and the new Manual for Complex Litigation endorses some of these reforms as well.96 Others have noted the potential for simplifying complex litigation through partial summary judgment and for using special masters to promote coherent case organization or to actively encourage factual stipulations.97

The implications of such reforms for jury competence may all, in principle, be examined empirically, but little rigorous social scientific investigation of any of them has been done in the years since I wrote.98 Among the few suggested reforms that have been the subject of carefully controlled research are pre-instructions, the use of written jury instructions, jury note-taking and jury question-asking.

Pre-instructing the Jury

Research in cognitive psychology suggests that informing a person how to frame information he or she is about to receive can enhance recall and aid in the interpretation of ambiguous material. Thus pre-instructing a jury on case-specific matters might aid it in understanding how evidence fits into a case.99 Vicki Smith reports partial support for this view. She found that the timing of instructions did not affect the recall of mock jurors for either relevant or irrelevant facts in a homicide case, nor did the timing of instructions affect interpretations of the evidence. Mock jurors instructed both before and after the trial were, however, better able to apply the law to the facts of the case than jurors instructed only after the trial.

Pre-instruction on general matters such as the presumption of innocence, burdens of proof, or the limited admissibility of certain evidence could in theory affect how evidence is encoded and in this way could affect verdicts.100 Heuer and Penrod report a field study in which Wisconsin trial judges agreed to assign cases randomly to a “pre-
instruction" or a "no pre-instruction" condition. It appears that pre-instructions did not hurt the trial process and may have helped. Jurors who received pre-instructions reported that the pre-instructions generally helped them in accomplishing such tasks as evaluating the evidence during the trial, applying the law to the facts, and remembering the judge's instructions, but a comparison of their responses with the responses of jurors who had not been pre-instructed does not support their belief that they were better off in these ways than if they had not been pre-instructed. Pre-instructed jurors were, however, more satisfied with the way their trials had been conducted than jurors who had not been pre-instructed, and judges were both less surprised by the verdicts of pre-instructed juries and more satisfied with them. Moreover, lawyers and judges saw pre-instructions as having virtually no disruptive or adverse affects on the jury and as tending to increase the fairness of trials. Thus pre-instructions seemed to have had only positive effects, if they had any effects at all. Since Heuer and Penrod do not report what pre-instructions were given, further work must be done to see whether their findings apply to both case-focused, substantive pre-instructions and to instructions on matters, like burdens of proof, that are common across trials.

Giving Written Instructions

Heuer and Penrod's study of pre-trial instructions also randomized the presentation of written instructions to juries for use during their deliberations. Jurors who received written instructions reported that these instructions helped them in a variety of ways, but their ratings of the helpfulness and understandability of the judge's instructions was no higher than the ratings of those who did not receive written instructions, and jurors who had received written instructions performed no better on a multiple choice test designed to measure comprehension of the judge's instructions than did jurors who received only oral instructions. Neither jury deliberations nor the trial seemed to be affected in other significant ways by written instructions, nor did written instructions seem to function differently in longer cases than in shorter ones. Judges and lawyers generally approved of written instructions, however, and did not see them as causing any significant problems. Because all parties seemed to like written instructions and they seemed to cause no harm, Heuer and Penrod recommend written instructions as a jury reform.
Note-taking and Question-asking

Heuer and Penrod also provide us with the best research to date on the effects of allowing jurors to take notes or to ask questions. They did two field studies of these procedures, one in Wisconsin and one with cooperating federal judges across the nation. The Wisconsin study, like the study of pre-instructions and of written instructions, found that no great benefits but little harm came from these procedures. The national study, however, is more relevant to our concerns since the authors looked specifically at how note-taking and question-asking functioned in complex and in simpler cases. This study involved 103 federal judges who agreed to assign one or two cases on an experimental basis. The authors randomly assigned cases to the experimental conditions, and the judges implemented these assignments either in their next case or, if they were willing to accept two assignments, in both their next case and their next complex case. All told, the authors collected data on seventy-five civil and eighty-five criminal trials.

Based on judicial responses, three dimensions of complexity were defined, one relating most closely to the complexity of the evidence, a second to the quantity of the evidence, and a third to the complexity of the law. On average, the judges seemed satisfied with the jury’s performance, rating the jury verdicts in both criminal and civil cases above seven on a nine-point scale with respect to both legal correctness (nine indicating strong agreement that the verdict is legally correct) and their own satisfaction with the verdict (nine indicating very satisfied). A composite of these and two other items reveals that judicial satisfaction with the jury’s performance did not vary with the complexity of the case.

Jurors who were allowed to ask questions had significantly higher scores than those not allowed to ask questions with respect to feeling well informed, the perceived ease of reaching verdicts and understanding the law, the perceived helpfulness of the prosecutor and defense counsel, and the certainty that they felt that the verdict was correct. Surprisingly, jurors allowed to take notes reported feeling less well informed than jurors who were not allowed to take notes, and they reported it was more difficult to reach a verdict. Judicial approval of jury verdicts was, however, not affected by either of the two experimental procedures, and this finding does not change with the complexity of the case. Excluding hung juries in criminal cases, judges agreed with the juries’ verdicts in about 75 percent of the criminal trials and in about 63 percent of the
In civil cases. Interestingly, in civil cases judges were considerably more likely to disagree when juries returned defendants’ verdicts (52 percent disagreement) than when juries returned plaintiffs’ verdicts (29 percent disagreement). Disagreement did not seem to be related to case complexity, to trial procedures, or to their interaction.

A test of interaction effects revealed that jurors found the ability to ask questions increasingly helpful (in the sense that they felt well informed and found it easy to understand the law) as the law became more complex. Permission to ask questions also added more to jurors’ confidence in their verdicts as the evidence became more complex. As the amount of information offered at trial increased, however, the chance to ask questions seemed to backfire. With heavy information loads, jurors allowed to ask questions were less likely to feel well informed than jurors not allowed to ask questions and more likely to report difficulty in understanding the law and reaching a verdict. For note-taking the two-way interactions with the legal dimension of complexity are significant but directionally inconsistent; in some situations increasing legal complexity is associated with positive juror evaluations, and in other cases it is associated with negative reports. The most important findings seem to be that as legal complexity increases, note-taking jurors are less likely to feel satisfied with their verdicts and are less certain they have reached the right result than jurors not allowed to take notes.

Heuer and Penrod also examined certain procedures that varied because responding judges handled cases in different ways. They found that an initial juror orientation lecture contributed to juror satisfaction and that the use of special verdict forms increased not only juror satisfaction but also verdict confidence and feelings among jurors that they understood the judge’s instructions and were well informed. Judges who commented on the evidence or summarized, on the other hand, left jurors feeling that it was harder to reach a verdict and harder to understand the law, and the closer a judge hewed to pattern verdict instructions, the less confidence jurors had in the verdicts they reached. Judicial satisfaction with jury performance was not, however, affected by the use of any of these four procedures.

These data suggest that jurors feel that they perform better when they can ask questions, when they are given verdict forms, and when they receive an initial orientation. But if juries do perform better with these aids, judges seem not to notice. One might argue that as long as jurors feel they are being helped by certain measures, such measures should be used, and this is what Heuer and Penrod conclude. It is not clear,
however, that the juror responses that Heuer and Penrod treat as positive indicators of competence are always that. For example, confidence in a verdict rendered in a close, complex case may be illusory since there may be no verdict in which a factfinder should feel confident. A juror who felt less confident in such a case might have done a more honest job of grappling with the evidence and returned a verdict more likely to be correct. The same could be said about the sense that a verdict is easy. Indeed, certain juror attitudes toward their experiences may have a different normative status depending on whether cases are close on the evidence or easy, and whether they are simple to comprehend or complex. Thus while the Heuer and Penrod study is the best effort to date to shed an empirical light on the usefulness of reforms designed to increase jury competence, their results are inescapably ambiguous, for the normative status of their dependent variables is not completely clear. Perhaps this is the reason that judicial evaluations do not confirm juror reports.

Evaluating Informal Experiments

If there have been few rigorous scientific experiments evaluating proposed jury reforms, there has been no lack of experimentation in a more colloquial sense. We saw, for example, in our case studies that most of the trial judges involved “experimented” with one or more of the procedural reforms that have been commonly suggested. Indeed, some reforms, such as note-taking, may be more the rule in complex cases than the exception. Assessments indicate that when judges try innovative procedures, they seldom find the harms they feared and often perceive benefits.

Leonard Sand and Steven Reiss, for example, persuaded one or more of the twenty-eight judges serving on the Second Circuit in June of 1983 to try one or more of seven “novel” procedures. These procedures included ten minutes of attorney participation in voir dire; individual, private voir dire; pre-instructing the jury; allowing jurors to ask questions; informing jurors they could take notes; providing jurors with a written copy of the charge; and providing jurors with a tape recording of the charge. Sand and Reiss found in almost every instance that if lawyers and judges could be induced to go along with a procedure, a majority, sometimes an overwhelming majority, reacted favorably to it. These findings are consistent with the later work by Heuer and Penrod, who report that those Wisconsin judges, lawyers, and jurors who participated in their field experiments invariably liked the innovative procedures they
experienced, and they liked them more and saw fewer problems with them than judges, lawyers, and jurors who had not experienced the procedures but were asked how they thought they would have reacted to them.\textsuperscript{113}

\textit{Reforms Worth Doing}

Having reviewed the evidence, I cannot say that there is an adequate empirical basis for concluding that reforms like those that have been proposed will make juries better factfinders in complex cases. Yet given a wealth of unsystematic experience and some few experiments, it appears likely that most proposed reforms will not be harmful. Indeed, there is reason in common sense, scientific theory, and occasional experience to believe that certain of the suggested reforms will do some good. Where the likelihood of harm is low and the possibility of benefits exists, there is a case for action, as there is when some harm is possible but the likelihood of significant benefits is substantial. I believe that a number of the proposed procedural reforms fit one or the other of these two circumstances. Rather than wait for definitive research, some of the proposed procedural reforms should be instituted now.

Since we are acting without full knowledge, however, we should, in innovating, act to gain knowledge. Ideally, courts would institute any reforms experimentally, but this is unlikely to happen. In default of experimental implementation, there should be a systematic attempt, perhaps under the auspices of the Federal Judicial Center and the National Center for State Courts, to monitor reforms and to systematically canvass reactions to them. Indeed these agencies and other funding sources should make the continued experimental evaluation of such reforms a priority. Many reforms are plausible; I will discuss ten I would like to see.

NOTE-TAKING. Opposing note-taking seems futile because it is already often allowed in complex cases. Note-taking was permitted in a number of our case studies, and in no case in which it was allowed was there a suggestion of any problems. Moreover, a theoretical argument can be made that note-taking will increase juror involvement in the trial and juror performance.\textsuperscript{114} The experiment by Heuer and Penrod provides some cause for concern, however. In this experiment note-taking was associated with feeling less well informed and finding it harder to decide on a verdict. These feelings are too consistent with a hypothesized danger of note-taking—the possibility that jurors who take notes will miss
information as they try to copy other information down—to be summarily dismissed. Thus, in trials where daily transcripts are being prepared, jurors might be barred from taking notes and instead be furnished with transcripts to use as they see fit. The reactions and performance of jurors furnished with transcripts should be compared with those of jurors in similar trials who instead may take notes.

QUESTION-ASKING. This is the reform most strongly supported by the Heuer and Penrod field experiment, and it seems to be the one most desired by jurors as well. The experimental and anecdotal evidence both suggest that jurors do not abuse question-asking privileges, and questions may give lawyers an idea of how well jurors are following the evidence. It should be noted, however, that the Pennzoil judge reversed himself and ceased to allow juror questions part way through the trial. Problems with questioning may have arisen in that trial because the judge permitted jurors to interject questions orally at any point in the testimony simply by raising their hands. When jurors submit written questions after a witness has finished direct or cross-examination, problems are not reported.

PRE-INSTRUCTION. Somewhat less confidently, I suggest that at the outset of complex cases jurors receive a composite of pre-instructions and an orientation lecture. The case for this change is better grounded in theory than in experimental evidence, although some apparently positive effects appeared in the Heuer and Penrod experiment. Theory suggests that if the jurors are aware of the issues they have to resolve, they may be better able to understand the evidence they receive and to put otherwise confusing evidence in an understandable perspective. Theory coupled with some evidence tells us that instructions concerning how evidence should be coded are better given before evidence is encountered than after the fact. In accordance with these views, jurors should be informed before the trial begins of the legal issues they are expected to resolve and of the major factual disputes between parties. While this is also the stuff of opening statements, an introduction unskewed by partisanship might be a helpful prelude to an exposition by the parties. There is, however, a danger of reversible error here, and a judge should develop such comments carefully in consultation with the parties.

It also makes sense for the judge to pre-instruct the jury, as judges now often do, on what is and is not evidence, on how they should respond to objections, on why discussions occur out of the jurors'
hearing or with the jury escorted out of the courtroom, and on similar matters. A potentially important pre-instruction that is commonly given emphasizes that every case has two sides and that the jurors should not discuss the case or make up their minds until they have heard all the evidence. While some people regard such instructions as futile, anecdotal evidence from jurors serving in lengthy trials shows that the temptation to discuss a case as it progresses can be and often is resisted. The effectiveness of the instruction to maintain an open mind is more problematic since jurors admit to leaning toward one side or the other as a case progresses. Perhaps the goal of the instruction to keep an open mind would be better achieved if jurors were told at the outset that when they began deliberations, they should not vote immediately but should together review the evidence in the case. Some of the case studies, consistent with mock-jury research suggest that early voting can lock people into positions and diminish the productivity of deliberations.119

It is probably also wise to explain the burden of proof to jurors at the outset and, in a criminal case, to acquaint jurors with the presumption of innocence. The theoretical reason for this is that jurors may engage in “on line” decisionmaking as evidence is received, and if they do, they should at least do so knowing the correct standard. Despite the Saul Kassin and Lawrence Wrightsman study that purports to have shown the virtues of this instruction experimentally, this recommendation is based more on common sense and theory than on experimental evidence.120

WRITTEN INSTRUCTIONS. Each juror should receive written copies of the instructions for reference during deliberations.

SIMPLIFICATION. Complex cases should be simplified through pre-trial stipulations, severance of joined claims or counts, partial summary judgment when facts are indisputable, and other methods of limiting issues in dispute. One promising technique is to allow the parties only a limited time to present their cases. Two of the courts in our case studies used this technique to apparently good effect. Where, however, one party’s case is inherently more time consuming than the other’s, difficult problems of fairness may arise. Also, to promote simplification a court should not allow parties to present technical documents largely to impress jurors with their bulk, as may have occurred in the ABA trade secrets case. If documents are redundant or unnecessary, they should be ruthlessly pared down according to the provisions of Rule 403 of the Federal Rules of Evidence to what is new and essential. If jurors are
unlikely to understand certain documents, those documents should not be admitted unless the party offering them offers evidence to aid in their understanding.

TRIAL NOTEBOOKS. Steps should be taken, as they often are today, to ensure that jurors can easily follow documentary evidence. One aid in this respect is the loose-leaf trial notebook that can be added to as exhibits are offered. Such notebooks can include at the outset information about important facts the parties have agreed on. Documentary evidence should be distributed to the jurors so that they can follow it as it is discussed, or visual aids should be used to the same effect. Those documents and exhibits admitted into evidence should be organized and indexed for easy retrieval once the jury has retired to the jury room. The jury should not be given boxes of unorganized exhibits, as has occasionally happened.

SIMPLIFIED INSTRUCTIONS. Legal instructions should long since have been rewritten to foster ease of understanding. In jurisdictions where this has not happened (most places), rewriting should occur.

IMPROVED COMMUNICATIONS. During the deliberations jury questions should be answered in plain language. Juries should not simply be referred back to the instructions. The case studies reveal that such references seldom helped and that without judicial guidance a reasonable interpretation of an instruction could be wrong. Indeed, in several of the cases where the jury verdict seemed to reflect a legal misinterpretation, the jury had realized it was having trouble understanding what the law required and had sought the court's aid in understanding but had not received any. As a corollary to this recommendation, as long as a court's answers to juror questions appear fair and helpful, appellate courts should not reverse trial courts simply because precise legal language was not tracked and an unlikely legal misinterpretation is possible.

ENLARGED JURIES. We saw in a number of cases that jurors tend to follow the lead of their most competent members when cases are complex. Increasing the number of jurors increases the likelihood that one or more jurors will have a good understanding of the law and evidence. In addition, the larger the jury, the better the jury's collective memory is likely to be. The recent amendment to Rule 48 of the Federal Rules of Civil Procedure, which eliminates alternate jurors, should help achieve larger juries since no court will seat only six jurors in a complex case and risk a mistrial should one juror be unable to continue. Courts should routinely seat twelve jurors when cases are likely to be complex. Indeed, in long complex cases where substantial juror attrition is a danger, the
court should seek the parties' consent to ignore the limitations of the new rule and seat more than twelve jurors, on the understanding that if more than twelve jurors remain at the end of the trial, excess jurors will be treated as alternates. Juror attrition over the course of a long trial might otherwise dissipate the strengths that larger juries afford or even reduce the jury to fewer than six members, allowing one party to force a mistrial. Twelve-person civil juries should be allowed to return verdicts over the dissent of one or two members. This will minimize hung juries, which are especially costly in lengthy cases, and the excessive influence, as in the ABA trade secrets case, of individual and probably incorrect dissenters. Accepting nonunanimous verdicts in federal civil cases requires the consent of the parties, which a judge should try to secure in advance; allowing nonunanimous verdicts in civil cases without party consent is a matter for Congress's agenda.

JUROR SELECTION. Courts should encourage the selection of competent jurors. The tactical desire to gain a jury that can be fooled deserves no legal respect. The issue is, however, a delicate one, for judicial intervention to ensure juror competence can interfere with a party's right to exercise peremptory challenges. But some interference is tolerable if, as I have argued elsewhere, the peremptory challenge is justified largely as a device to eliminate prejudice that escapes the sieve of the challenge for cause. Encroachments on the peremptory challenge, however, have already been made to forestall racial prejudice, and further encroachments, albeit for good reasons, threaten to reduce the peremptory challenge to a challenge for cause. I do not, therefore, support legal change in this area. Rather, I recommend judicial jawboning. The trial judge should make the desirability of securing a capable jury clear and should urge the attorneys not to challenge those jurors who seem most likely to understand the issues in the case. Where such a juror is challenged, the judge might ask an attorney why the challenged juror was thought undesirable. While an attorney would be free to tell the court there was no reason, few attorneys would want to admit, even implicitly, that they wished to avoid jurors who seemed likely to understand the case. Provisions should also be made to compensate jurors generously in trials that exceed a certain length so that fewer people would seek to be excused where cases are likely to last more than a few weeks.

I could suggest more changes that might be made now, but I suspect these ten are more than are likely to happen in the near future. I believe that each of these suggestions can be plausibly justified based on what we know now, but none is as firmly rooted in reliable research as I
would like it to be. For this reason it is essential that the institution of such changes should be the occasion for more study and not a signal that research is no longer needed because reform has occurred.

A New Perspective

I have now canvassed the research done since my first article that bears on the three issues that in my view had to be resolved to support a due-process-based complexity exception to the Seventh Amendment right to jury trial. There has, however, been another development since 1981 in the psychological understanding of jury trials that relates to the likely capacity of juries in complex cases. This is the development of a new model of jury decisionmaking that Nancy Pennington and Reid Hastie, its principal proponents, call the "story model." The story model suggests that jurors try to make sense out of the evidence they are offered by constructing the story that best explains it. It is an exciting and plausible development in the effort to understand jury decisionmaking both in its own right and because it fits in nicely with recent developments in cognitive psychology, where stories are a kind of schema linking people, motives, and actions.

At present the story model has been used largely to explain jury decisions in cases like homicides, where such alternative stories as the premeditated assault and the self-defense story can be plausible explanations for the same array of evidence. The model seems to work well in these settings, and it illustrates how jurors can take evidence and, using various criteria for consistency, try to fit it to one among a stock of stories with which they are already familiar. We also see in the experimental evidence that preexisting story structures can cause jurors to misperceive certain evidence or to recall evidence that was never presented. This is to be expected if stories are a type of schema.

Now consider the complex case. From a story perspective there are two types. First, there are cases in which the evidence that makes a case complex can fit a slot, by which I mean a place that holds a story element, in a story. For example, consider a case like W. R. Grace. The plaintiff's main story line is not complex and is readily available to most jurors: two companies carelessly dumped poisonous effluents that reached the water table and percolated to the town's wells that supplied the plaintiff's water. The plaintiffs drank the water and became seriously ill.
The easy conclusion is that the defendants are wrongdoers who should pay for the harm their effluents caused.

What makes the case difficult is not the unfamiliarity of the basic story line. Instead, at several points the defendants draw on scientific evidence to argue that connections assumed by the story do not exist. For example, the defendant companies denied that their effluents could reach the town’s wells and claimed there was no proof that the effluents were the source of the increased incidence of leukemia in the plaintiff population. Thus, what is complex in the trial involves two story slots that must be filled in for the plaintiffs to recover: there must be a way for the defendants’ effluents to reach the town’s wells, and there must be a reason to conclude that the effluents caused the plaintiffs’ leukemia. The trial is not limited to these issues—the plaintiffs, for example, must prove that the defendants discharged the effluents—but the trial, particularly what is complex about the trial, turns on them.

We can think of the jury’s decision regarding how to fill each slot as itself involving a competition among stories. But these competing stories are specialized stories that only scientific experts know well. Invariably, in a case like W. R. Grace, the scientists whom the jury hears disagrees. How are jurors to decide between opposing scientific stories if such stories are unfamiliar and the jurors have no set of plausible stories with which to compare the scientific evidence? To what extent will juror decisions reflect the ways in which the parts of the story that they can follow pull them along? Is, for example, the very fact that the plaintiffs in Grace are suffering from a feared and esoteric disease likely to lead jurors to believe that the defendants’ effluents must have seeped into the town’s wells and must have been capable of causing the disease (because otherwise there is no explanation for the evidence before them)? Would a jury that accepted the plaintiffs’ expert’s scientific story for this reason be acting irrationally? Suppose instead of relying on an unfamiliar scientific story about how effluents travel, a juror instead relies on schemas with which he or she is familiar—like the stereotype that nervousness is associated with lying—to decide which expert’s story to accept. Is this an irrational basis for decision?

Is there any reason to believe that the judge in a case like W. R. Grace would do a better job than the jury? The judge is likely to have a stock of pollution stories much as the jury does, but the judge, like the jurors, is probably unfamiliar with the stories the scientists tell. Moreover, the trial judge’s stories may be different from the jury’s stories or at least
may be differentially accessible. The unscrupulous plaintiff's attorney story (plaintiff's attorney is known for drumming up litigation) may figure in the judge's evaluation of the evidence when it hardly enters into the jury's. Or the judge may have easy mental access to a story in which potential polluters are more responsible or pollution is less likely to be harmful—access not shared by the jurors. In this situation, whose stock of stories should be brought to bear to help make sense of those aspects of the story—scientific stories within the story—that neither the judge nor the jury understands well? I can not answer questions like these here, but they are worth raising; they suggest a new approach to thinking about how judges and jurors respond to one kind of evidentiary complexity.

Other complex cases viewed from a story perspective are ones in which evidence is not necessarily hard to understand, but jurors are unlikely to be acquainted with the set of stories that most plausibly fit the evidence. The *Pennzoil* and *Micro/Vest* trials furnish examples. The jurors in these cases did not know various story lines that might make sense of facts relating to competing tender offers or business understandings about the terms of loans. In cases like these it is not surprising if jurors are tempted to make sense of the evidence by fitting it to stories they are more familiar with, such as a story of people making ordinary agreements. This tendency can be exacerbated when one side encourages the jury to use its ordinary stock of stories to interpret the facts of a case. This occurred in *Pennzoil*. It was perhaps best revealed in an interchange between the plaintiff's attorney, Joe Jamail, a personal-injury lawyer who ordinarily did no business litigation, and Marty Lipton, a well-known New York takeover lawyer. Jamail asked Lipton whether he was "saying that you have some distinction between just us ordinary people making contracts with each other, and whether or not it's a ten-billion-dollar deal? Is there a different standard in your mind?" Lipton replied, "Yes, indeed." Jim Shannon, who may have been the jury's most influential member recalls, "At that point my jaw just dropped." For a Wall Street lawyer or for most judges (not necessarily the judge in the *Pennzoil* trial), an answer that the standard was the same might have caused jaws to drop.

In these circumstances, are we better off with trial to a jury or to a judge? Despite the *Pennzoil* case, people can learn that stories they have not heard of before are plausible. The entrapment story in *DeLorean* is an example. And in *Pennzoil*, there is no reason to believe that the judge would have decided differently; indeed, a number of his pre-trial
Civil Juries and Complex Cases

Evidence rulings prevented Texaco from presenting a convincing, gapless version of the story it wanted to tell. Moreover, having access to a story that appears to fit a body of evidence is not always a virtue, for when a story is familiar, evidence may be distorted, overlooked, or even invented where the evidence recalls the story but does not fit it perfectly. The jury’s virtue is that different jurors have different preferred stories so the jury collectively is less likely than individual decisionmakers to be misled by the kinds of natural distortions and gap-filling that an available story may stimulate.

I do not pretend to have answers to the questions I pose, nor do I now suggest that the story model provides a new or more adequate answer to the question of whether judges or juries are better situated to decide complex cases. But story-model research does provide a new perspective on the issue; one that I think is worth further thought and exploration.

Conclusion

I began this paper by recalling my earlier article suggesting that three kinds of showings had to be made to justify a Supreme Court holding that due process required a complexity exception to the Seventh Amendment right to jury trial: first, a showing that juries did not decide such cases rationally; second, a showing that changes in the management of juries and cases could not cure identifiable jury deficiencies; and third, a showing that judges would find facts in such cases more rationally than juries.

I then examined the empirical evidence relating to these three issues. Much of the available evidence is unfortunately anecdotal. It is the kind of evidence on which one does not want important policy decisions to turn. To the extent that there is systematic social science knowledge that elucidates these three issues, none of it resolves an issue, and much of it is unclear in its import even if concerns about external validity are discounted. In short, not enough social science research has been done to date to allow us to reach firm conclusions about the capacities of juries and judges to handle complex cases or about the potential improvement in jury capacity afforded by possible reforms.

Thus, I could simply end this paper as I ended my first one, with a call for additional empirical research, an argument that the burden of proving the incapacity of juries in complex cases is on those who would
limit the Seventh Amendment right, and a judgment that that burden has not been met.

But there is more to be said. Throughout this review, strengths of the jury emerge. A close look at a number of cases, including several in which jury verdicts appear mistaken, does not show juries that are befuddled by complexity. Even when juries do not fully understand technical issues, they can usually make enough sense of what is going on to deliberate rationally, and they usually reach defensible decisions. To the extent that juries make identifiable mistakes, their mistakes seem most often attributable not to conditions uniquely associated with complexity, but to the mistakes of judges and lawyers, to such systematic deficiencies of the trial process as battles of experts and the prevalence of hard-to-understand jury instructions, and to the kinds of human error that affect simple trials as well. The anecdotal evidence should also remind us that it is difficult to predict which complex cases will trouble juries and which they will handle well.

The import of the experimental literature is similar. Experiments show that some factors that make for complexity, such as the joinder of charges or statistical evidence, can lead to jury mistakes. Yet in the most realistic studies, juries perform surprisingly well. In the study by Diamond and Casper, for example, contrary to what one might predict, regression models were as influential as the more concrete and intuitively understandable yardstick models. In the study by Horowitz and Bordens, joined plaintiffs and scientific evidence caused few problems. Moreover, juries in the Diamond and Casper study, like some of the case-study juries, showed an extraordinary capacity for identifying their most capable members and then letting them lead. This suggests that the challenges that complex cases pose might be met best by the simple expedient of getting more capable people to serve on juries.

The empirical evidence also provides no reason to believe that judges will fare better in the face of complexity than juries, for we have little basis for deciding how judges will do at all. What we have instead are anecdotes that make the point that judges dealing with unfamiliar, technical information can be as confused as we fear similarly situated juries are. The most reasonable conclusion about the relative capacity of juries and judges is probably one that I reached in my earlier article without substantial empirical support and am willing, still without substantial empirical support, to reiterate here: in complex cases we can expect that some judges will be more capable than the average jury, and we can expect that the average jury will be more capable than some
judges. But in many cases we will not know in advance whether judge or jury is likely to be the more rational decisionmaker.

With regard to reforms in the management of complex cases and the conduct of jury trials, we live in a world of constantly experimenting judges. The problem is that the experimentation most judges do is uncontrolled, hardly visible, and unsystematic, so we learn almost nothing from it except what we can learn from the fact that great outrages over various novel procedures seem not to have arisen. Given this, I have suggested that we should apply the little we have learned from systematic experimentation, consider what social science theory suggests, add a good dose of common sense, and make those limited changes we think will improve the jury system. We should, however, resolve to study the changes we make so that we will learn if our hunches are right.

For more than a decade now, the fitness of juries to hear complex cases has been on trial. Twelve years ago it was only possible to say that the case against the civil jury had not been proven. While we still do not have a wealth of evidence, there is enough to support a more positive verdict. Based on what we know today, there is no empirical case for a complexity exception to the Seventh Amendment. Instead the weight of the evidence indicates that juries can reach rationally defensible verdicts in complex cases, that we cannot assume that judges in complex cases will perform better than juries, and that there are changes that can be made to enhance jury performance. These conclusions will not necessarily be sustained as research proceeds. But they are the best that empirical knowledge today can offer.

Notes


3. The Supreme Court has never applied the Seventh Amendment to the states so, as a matter of federal constitutional law, state courts are free to eliminate jury trials in civil actions. Many states, however, are bound by their own constitutions to offer jury trials in actions at law.


5. In a situation where neither judges nor juries could be expected to resolve a matter rationally, the resolution of the issue, as some commentators have suggested, could be vested in some other institution, such as a panel of experts. See Luneberg and Nordenberg, “Specially Qualified Juries,” p. 887; and David U. Strawn and G. Thomas Munsterman, “Helping Juries Handle Complex Cases,” *Judicature*, vol. 65 (March–April 1982), pp. 444–47. In the setting of a lawsuit, such an alternative institution might be confined by the Constitution to the limited role that due process requires; it might, for example, resolve difficult technical questions that the decisionmaker, whether judge or jury, would then take as proven in deciding a case. Outside that class of actions that are “legal” within the meaning of the Seventh Amendment, a broader role might be accorded expert decisionmakers, as is often done in administrative agencies.

6. The cases included in the table are a sample selected for convenience. Some of them had come to my attention before I wrote this chapter; others were uncovered by a research assistant. When I started writing the first draft of this chapter, I knew of no other case studies that met the criteria for inclusion. Since completing the first draft, I have learned of several additional case studies that meet my criteria, but including them would lengthen the chapter without changing any of the conclusions supported by the current sample. It is important to recognize that my sample of cases is a nonrandom selection containing a disproportionate number of high-profile cases. The cases were not, however, selected because of their tendencies to prove or to disprove any of the hypotheses examined in this chapter. Indeed, not until I selected and read the cases was I aware of their implications for the issues explored in this chapter. Lisa Bernt, a 1992 graduate of the University of Michigan Law School, checked my summaries of key case characteristics.

7. Of course confusion may also be understated after memories have faded.


Civil Juries and Complex Cases


17. This is not, I should note, what the jury did in Grace. Their verdict is discussed below. And the evidentiary dispute recounted in the text involved pollution attributed to the Beatrice Company, a codefendant in the case.

18. More precisely, I am giving my subjective view of another observer's subjective view of the quality of the verdict.

19. Since mistakes on damages can be corrected through remittitur either at the trial level or on appeal, such mistakes are not as costly to the efficient administration of justice as mistakes on liability that, if caught, are more likely to require a new trial.

20. This is why I argued in my 1981 article for nonunanimous juries in civil cases.

21. This limitation strikes me as legally questionable, so the jury's performance might be better than the judge's in this case.


24. One juror, the jury's only college-educated member, understood the implications of the answer. This juror kept quiet because he was the defendant's strongest proponent within the jury room.

25. This does not mean that the jurors would have been acting irrationally had they realized the consequences of an answer of "not determined" and reported a date. The jurors might reasonably have been confident that by a certain date the contamination had begun and might have noted this date. The question, however, called for the earliest date on which the contamination could, by a preponderance of the evidence, be shown to have happened.


27. Adler, "How to Lose the Bet-Your-Company Case"; and Thomas Petzinger, Jr., Oil and Honor: The Texaco-Pennzoil Wars (Putnam, 1987).


29. Ibid.


31. The authors' description of the deliberations is based on one collective interview held several months after the trial had ended. Research on limiting instructions suggests that the admitted role of the forbidden evidence may have loomed larger at the time of the interview, when the limiting instructions may have been forgotten, than it would have had the interview immediately followed the trial. The reason is not that jurors are unaffected by evidence that they are instructed not to consider or to consider only for a limited
purpose, but rather that jurors often act during their deliberations as if they will not be improperly influenced. Thus, if we may rely on studies of mock juries, when a juror mentions evidence the jury has been instructed not to consider, another juror often points out that the evidence is not to be considered, and the information receives no further overt attention. Nevertheless, such evidence may have the effects the instructions were designed to eliminate. Valerie P. Hans and Anthony N. Doob, "Section 12 of the Canada Evidence Act and the Deliberations of Simulated Juries," Criminal Law Quarterly, vol. 18 (March 1976), pp. 235–53; and Jonathan D. Casper, Kennette Benedict, and Jo. L. Perry, "Juror Decisionmaking, Attitudes and the Hindsight Bias," Law and Human Behavior, vol. 13 (September 1989), pp. 291–310. What may have happened over time is that the jurors forgot they were not supposed to consider certain evidence and accurately reported the effects that evidence had on them, even though at the time these effects would have seemed insubstantial to one observing the jury deliberation. This possibility seems more plausible with respect to the documents that were improperly used against defendants other than the one charged with their production than it does in the case of the jury’s apparent defiance of the judge’s instructions not to make awards for past medical damages. On the damage issue the jurors reported a formula they used, something they seem unlikely to misremember, and the formula awarded money for past medical expenses contrary to the judge’s instructions.

32. Not only did the jurors have reason to believe that the plaintiff’s future earnings might be diminished by the likelihood he would return to Mexico, but the adjustment made was not great and avoided a situation in which this plaintiff would have been awarded more than a plaintiff who was far sicker than he was, albeit with a shorter life expectancy. Moreover, when punitive damages were apportioned, the Mexican plaintiff received $1 million, the same as each of the other plaintiffs.


37. Ibid., p. 101, n. 27.

38. Adler, “How to Lose the Bet-Your-Company Case.”


45. The fear is not irrational. Haney has shown that jurors who have observed the death-qualification process have attitudes toward the case that are more consistent with a
willingness to convict than the attitudes of those who have not witnessed it. Craig Haney, "The Biasing Effect of the Death-Qualification Process," *Law and Human Behavior*, vol. 8 (June 1984), pp. 121-32. At least part of the explanation for this is that the jurors assumed that unless the judge and lawyers thought the defendant was guilty of a capital offense, there would be no death qualification. A defense counsel might similarly fear that a jury would regard the presence of evidence contesting damages as an acknowledgement that the plaintiff’s position on liability deserved to prevail. Nevertheless such an attitude on the part of jurors would be only one factor that would affect their weighing of the evidence on liability, and it is unlikely to matter as much as evidence more directly related to liability. The potential costs of not presenting damage evidence are often so substantial that it is foolish to withhold such evidence for strategic reasons.

48. Ell, “Right to an Incompetent Jury.”
50. Ibid.
51. Ibid., p. 38.
52. Heuer and Penrod, “Trial Complexity.”
54. If the defendant has the same blood type as the perpetrator of the charged crime and that blood type is found in 10 percent of the population, a juror caught in the prosecutor’s fallacy would conclude that since there is a 10 percent probability that the defendant would have the blood type if he were innocent, there must be a 90 percent chance that the defendant is guilty. William C. Thompson and Edward L. Schumann, “Interpretation of Statistical Evidence in Criminal Trials: The Prosecutor’s Fallacy and the Defense Attorney’s Fallacy,” *Law and Human Behavior*, vol. 11 (September 1987), pp. 167–88.
55. If the defendant and perpetrator have a blood type that they share with 1 percent of the population, one who falls prey to the defense attorney’s fallacy would reason that in a city of 1,000,000 people probably 10,000 people share the blood type in question, so the evidence has almost no incriminatory value.
61. Yardstick models are nonstatistical. They use comparative data from similar firms which did business in competitive markets at the time of the defendant’s anticompetitive activity and are based on the premise that the difference in prices paid or profits made by
the benchmark firms and the plaintiff is a good measure of the excess costs imposed on or profits lost by the plaintiff company. While valid comparisons are difficult to obtain, the approach is the kind of concrete, common-sense approach that one would expect a jury to understand intuitively and to appreciate easily. Regression models involve time series analyses of pricing patterns before, during, and sometimes after the price-fixing agreement. The model used in the experiment involved a statistical projection, based on performance before price-fixing, of the price the plaintiff would have paid had there not been a price-fixing agreement. Jurors are not accustomed to dealing with regression models, and one would expect that they would find statistical modeling difficult to understand.

The other aspect of this study, which I shall not discuss, concerned the implications of informing jurors that awards in suits brought under the Sherman Act are trebled.

62. Whether he offered a statistical model or a yardstick model, the plaintiff's expert presented data that suggested the plaintiff's damages had amounted to $490,000, while conceding that they might be as low as $420,000. The defendant's expert, regardless of the model used, presented data suggesting that the most likely figure for damages was $35,000, but he conceded that the model could not rule out damages of up to $105,000.

63. There was no statistically significant difference in the ratings given the experts on persuasiveness or trustworthiness.

64. The correlation between the juries' mean predeliberation awards and their final verdicts was .54; for median predeliberation awards it was .62. For a jury's highest predeliberation award it was .24; and for the lowest predeliberation award it was .20.

65. The correlation between the foreperson's predeliberation awards and final jury awards was .44 across jury verdicts, while the correlation of final verdicts with the awards made by other members of the jury was .22.

66. The correlation between the preferences of the foreperson and final jury verdicts is .57 for forepersons with some postgraduate education and a course in statistics and .36 for forepersons who had neither of these qualifications.

67. The contrary implication that one may draw from the Schumann and Thompson study may result from the fact that their simulated jurors were a relatively homogenous group or that they had few subjects who had studied statistics. Thompson's description of that study, on which I have relied, does not describe the demographics of their simulated jurors. See "Are Juries Competent to Evaluate Statistical Evidence?"


69. The possibility that jurors, when faced with a conflict of experts, simply endorse the views of one of them is a suggestion made by Raitz and others, who looked at the implications of expert economics testimony in a mock age-discrimination case. Allan Raitz and others, "Determining Damages: The Influence of Expert Testimony on Juries' Decision Making," Law and Human Behavior, vol. 14 (August 1990), pp. 385-95. As a general matter, the findings of Raitz and his colleagues simply do not hold up in the face of the methodologically superior work of Diamond and Casper. See "Blindfolding the Jury." Raitz and his coauthors presented subjects with a 150-word summary of a trial and a 200-word "transcript" of testimony relating to damages in their "no expert" condition. The "plaintiff's expert only" condition added 200 words and included the examination and cross-examination of the plaintiff's expert, while the two-expert condition added an additional 400 words and the examination and cross-examination of a defense economics expert.

One cannot, however, on the basis of the work by Diamond and Casper rule out the possibility that jurors use an "endorsement strategy" when conflicting experts use the same methodology, although the W. R. Grace case shows jurors reacting in the opposite way by acknowledging that they are unable to decide. Nor is there necessarily anything wrong with an endorsement approach if the endorsement is based on a reasonable assessment of the credibility of the opposing experts, as may have occurred in Johns Manville. Also, if
only one side presents expert evidence, endorsement of the expert's views may be a natural response, as seems to have occurred in *Pennzoil*.


73. Researchers who have investigated the "just world" paradigm report that the greater the harm suffered by a victim, the greater the tendency to blame the victim. This is attributed to a psychological need to feel in control, which is undermined if people are seen to be suffering excessively through no fault of their own. Melvin J. Lerner, *The Belief in a Just World: A Fundamental Delusion* (Plenum Press, 1980).


76. In a later article using the same stimulus tape, the authors investigated the effect of bifurcating and trifurcating trials on damage awards. Irwin A. Horowitz and Kenneth S. Bordens, "An Experimental Investigation of Procedural Issues in Complex Tort Trials," *Law and Human Behavior*, vol. 14 (June 1990), pp. 269–85. The article is important for those interested in how juries deal with complex cases because it studies a variation that exists in such cases and suggests that the decision to bifurcate or trifurcate a verdict can have substantial consequences. It adds little, however, to what the first study suggests about the jury's capacity to handle complex litigation. Assuming that the jury's task is easier when it hears evidence on one issue and then decides that issue than when it hears evidence on all issues together and decides the issues seriatim, the study does not suggest that juries do worse when their task is more complex. First, there is no a priori reason to suppose that one pattern of verdicts represents a more rational outcome than another pattern. To the extent that verdict consistency across juries is such a standard, unitary trials produce more consistent verdicts. To the extent that a normative judgment is possible, the verdicts of the unitary juries appear more normative, for they are more likely to find for the plaintiffs on the liability issue, and an expert panel of two law professors and a lawyer who reviewed the evidence thought the liability evidence favored the plaintiff. Second, when a trial lasts only four hours, even if some evidence is complex and boring, presenting evidence on only one issue may not simplify the jury's task. Indeed, the longer trial may present a simpler decision problem if evidence allocated to one issue has some relevance on another, for additional evidence may make a case less close and a decision
easier rather than harder. An overlap in the implications of evidence may be one reason why the unitary juries studied by Horowitz and Bordens were more consistent with one another and more pro-plaintiff in deciding the issue of causality. The causality evidence was the only evidence subset that the expert panel saw as ambiguous rather than pro-plaintiff. Perhaps when the implications of other evidence for causality were considered, the task of deciding the causality issue became simpler, and this led to greater consistency. Alternatively, one might agree that sympathy or other factors that should not affect decisions on causality motivated pro-plaintiff decisions on this issue. But sympathy at least does not seem to have motivated the unitary jurors in this study. Sympathy should be a response to the plaintiff's damage claim; separated juries that heard liability evidence first, rendered a verdict, and then heard causal but not damage evidence, had a verdict pattern much like that rendered by the unitary jury deciding the liability and causal issues in the same order. In both cases the causal decision may reflect a desire to punish the defendant, but it is at least as plausible to suppose that the causal link between the presence of a chemical and the illness of a plaintiff becomes more likely when one knows that a defendant has been discharging the chemical where the plaintiff can encounter it. The authors also observed that once a jury decided for a side, consistency on the verdict among the members of a jury was "remarkable."


78. *Williams v. Florida*, 399 U.S. 78 (1970). Lempert, "Uncovering 'Nondiscernible' Differences." The experiment was Asch's line experiment in which experimental subjects were induced by group pressure to misidentify the larger of two lines. Misidentifications disappeared if one group member gave a true answer. The Court suggested that the proportion of group members giving true answers was important.


80. See, for example, *People v. Collins*, 438 P.2d 33 (1968).


86. See, for example, *Ethyl Corp. v. Environmental Protection Agency*, 541 F.2d 1 at 67 (D.C. Cir. 1976); and *International Harvester Workers v. Ruckelhaus*, 478 F.2d 615 at 651 (D.C. Cir. 1973). Many of the text's examples of misunderstanding come from appellate judges, whose errors and bewilderment are more accessible because their opinions are generally published and they cannot disguise their difficulties by simply finding facts without describing the basis of their reasoning. While there is some reason to believe that trial judges are structurally better able than appellate judges to gain an understanding of difficult scientific facts, there is no reason to believe that their structural advantages eliminate problems of understanding. As individuals, most appellate judges are likely to be more capable than most of the judges who sit in the nation's trial courts. See Richard O. Lempert, "'Between Cup and Lip': Social Science Influences on Law and Policy," *Law and Policy*, vol. 10 (April-July 1988), pp. 167–200.

87. Heuer and Penrod, "Trial Complexity."


90. In the day-to-day evidentiary rulings at trial, Farris’s rulings, particularly at the beginning of the trial, favored the defense on such mundane matters as objections to leading questions and hearsay. See Petzinger, *Oil and Honor*.


95. I do not mean to suggest that the jury will always be impartial and the judge not. An inadequate voir dire or prejudices too subtle to spot may lead to a biased jury in situations where the judge is unbiased. The right to jury trial itself, however—at least in criminal cases—is an element of due process, which suggests that the possibility of benefiting from certain kinds of jury biases is in some measure an element of due process.

I also recognize that due process is not considered to be violated in settings where it is permissible to dispense with a jury—suits in equity are an example. As a formal matter this is certainly true since due process is in large measure historically defined. Yet even in equity, if bias could be proved, due process would be violated by using a partial judge. Bias can be hard to demonstrate, however. In *Pennzoil*, for example, Judge Farris was challenged on the basis of the $10,000 campaign contribution from Jamail, which was the largest single contribution he had received and amounted to about 10 percent of his reelection funds. The judge who heard the challenge, however, ruled there was no cause for disqualification. Totally apart from the issue of bias, the point of the Seventh Amendment is that in civil actions at law brought in the federal courts the process due includes the right to a jury trial if one party requests it.


98. Convincing experimental research on rewriting jury instructions had been done by 1981. See, for example, Charrow and Charrow, “Making Legal Language Understandable.” It indicated that legal instructions could be rewritten to enhance jury comprehension considerably. Most states have not acted to revise their jury instructions; among the states that have done so, only a few have rewritten their jury instructions with this research in mind. See Tanford, “Law Reform by Courts, Legislatures, and Commissions.”


100. Kassin and Wrightsman found that mock jurors who were pre-instructed on the presumption of innocence and the need to prove guilt beyond a reasonable doubt in a criminal case gave lower estimates of guilt both during the presentation of the evidence and after all the evidence was in than did mock jurors who were instructed only after the evidence was in or who received no instructions at all. The latter jurors were in their judgments very much like those jurors who were instructed only after all the evidence was in. Kassin and Wrightsman provide an information-integration model for their findings, arguing that the pre-instructed mock jurors began deliberations with a lower prior probability of guilt, which affected their judgments of the defendant’s likely guilt as they assimilated the evidence offered at trial. Saul M. Kassin and Lawrence S. Wrightsman, “On the Requirements of Proof: The Timing of Judicial Instructions and Mock Juror Verdicts,” *Journal of Personality and Social Psychology*, vol. 37 (October 1979), pp. 1877–87. Recent work by Pennington and Hastie on “story models” calls the adequacy of information integration models into question, however. Nancy Pennington and Reid Hastie, “Explaining the Evidence: Further Tests of the Story Model for Juror Decision Making,” *Journal of Personality and Social Psychology*, vol. 62 (February 1991), pp. 189–206; Pennington and Hastie, “A Cognitive Theory of Juror Decision Making: The Story Model,” *Cardozo Law Review*, vol. 13 (November 1991), pp. 519–57; Pennington and Hastie, “Explanation-Based Decision Making: Effects of Memory Structure on Judgment,” *Journal of Experimental Psychology: Learning, Memory and Cognition*, vol. 14 (July 1988), p. 521; and Pennington and Hastie, “Evidence Evaluation in Complex Decision Making,” *Journal of Personality and Social Psychology*, vol. 51 (August 1986), pp. 243–58. An alternative explanation consistent with more general findings in cognitive psychology is that the early instructions affected the way in which the mock jurors encoded the information they received (as more consistent with innocence) so that at each point they remembered the evidence as less probative of guilt than did the jurors who had not been pre-instructed. From an encoding perspective, instruction after all the evidence was in would not be expected to have a substantial effect because the evidence received would have been already encoded and so would be remembered in a way more consistent with guilt. This explanation is consistent with research suggesting that instructions to disregard or limit evidence will be relatively ineffective and with research on the timing of opening statements. Casper, Benedict, and Perry, “Juror Decisionmaking, Attitudes, and the Hindsight Bias”; Hans and Doob, “Section Twelve of the Canada Evidence Act”; and

It was apparently left to the judge to decide what to pre-instruct on, and they did not report on this. Thus, we do not know for a given judge or for the mix of cases whether the pre-instructions given contained general instructions, case-specific substantive instructions, or a mixture of the two. Larry Heuer and Steven D. Penrod, "Instructing Jurors: A Field Experiment with Written and Preliminary Instructions," *Law and Human Behavior*, vol. 13 (December 1989), pp. 409–30.

There was no difference between the effects of pre-instruction in more or less complex cases, where complexity is operationalized by the length of the trial. My hunch, however, is that the Wisconsin trials that Heuer and Penrod studied probably included few if any that would qualify as complex in the sense that I have been using that term in this paper.

As Diamond points out in her chapter in this volume, Heuer and Penrod only tested comprehension of general instructions so they could use their instrument across trials. This performance measure may not have been sensitive to the advantages of trial-specific written instructions.

A true random sample was not achieved in the Wisconsin field study because attorneys refused to go along with the experimental assignment in some cases. This happened most frequently with respect to written instructions, which, at the time of the study, were required in Wisconsin. Defense attorneys in particular were reluctant to waive their right to have the jury furnished with written instructions. This is the best testimony to their popularity.

Heuer and Penrod, "Increasing Jurors' Participation in Trials," and Heuer and Penrod, "Trial Complexity."

Two thousand judges were originally contacted. While two-thirds of the 314 judges who returned a questionnaire saying why they declined to participate indicated they were no longer hearing jury trials, the participation rate is still so low that the sample does not allow one to be certain that the results of this study would generalize to the federal courts as a group. The most important reason to have reservations about generalizing from Heuer and Penrod's sample is that participating judges probably had more interest in the experimental procedures than nonparticipating judges. Without an interested judge, the effects that the authors observe might be different. Furthermore, participating judges may be more capable than nonparticipants, and the capacity of a jury in a complex case seems in part to be a function of the capacity of the judge presiding at the trial.

The length of the trial in hours was related to both the complexity of the evidence and, somewhat less strongly, to the quantity of the evidence as measured by the number of pages of documents in evidence, the number of items of evidence submitted, and the number of parties. The study indicates that long trials tend to involve masses of evidence and that at least some of the evidence is more complex than that found in most trials.

Questioning occurred in only fifty-one of seventy-one trials in which question-asking was allowed, according to a preliminary report of the data. American Judicature Society, *Toward More Active Juries: Taking Notes and Asking Questions* (Chicago, 1991). Heuer and Penrod, "Increasing Jurors’ Participation in Trials," report seventy-four cases in which question-asking was allowed. While Heuer and Penrod do not address the issue specifically, they seem to have compared trials based on whether question-asking was
allowed and not on whether it actually occurred. This analytical decision is in keeping with the design of the experiment, which is concerned with the effects of allowing judges to ask questions rather than with the effects of questions asked.

109. With respect to both comments and special verdict forms, I have ignored a significant association with the perceived helpfulness of the plaintiff-prosecutor because I do not think it is as important as the dimensions I cite, and I do not know precisely what to make of it. I am also ignoring significant interactions involving those variables that were not experimentally varied. There are a number of such interactions, but they are at times directionally inconsistent, and given the lack of experimental control, their interpretation is problematic.

110. Since there are fewer responses from judges than from jurors, one possible reason why juror responses are significantly associated with procedural variations while judicial responses are not is the relative lack of power in the analyses that deal with judges’ reactions. This possibility is difficult to evaluate since comparable statistics are not presented for judges and jurors. Given the small amount of variance that the procedural measures uniquely explain when judicial satisfaction is dependent, however, it is probably a mistake to attribute the lack of significance of these measures to a problem of power. Relative differences in power are more plausible as an explanation of why certain interactions are significant when the jurors’ views are dependent but not when the judges’ attitudes are.


115. In their Wisconsin field experiment, Heuer and Penrod did not find any evidence that this occurred. Heuer and Penrod, “Increasing Jurors’ Participation in Trials.”


117. Shannon, Texaco and the $10 Billion Jury.

118. Heuer and Penrod, “Trial Complexity.”


122. I support the creation of a fund that would drastically increase the payment jurors receive in any trial that extends past fourteen trial days. To pay for such a fund, I would “tax” the payments made to lawyers in long cases where money damages are sought or where litigation involves a commercial matter. For example, in long cases the parties might be charged 1 percent of all lawyer’s fees paid for pre-trial work and 3 percent of the fees paid for work during the trial. This proposal should not unduly discourage contingent-fee
representation since there would be no fees to be taxed if the plaintiff lost. It would also not discourage litigation over noncommercial matters disputed as issues of principal since no fees would be taxed in such cases. The exact trigger for payment could be adjusted to take account of party responsibility for lengthening the trial. In some small measure the tax should encourage pre-trial settlements and the shortening of cases that go to trial.


126. Diamond and Casper, "Blindfolding the Jury."

127. Horowitz and Bordens, "Experimental Investigation of Procedural Issues in Complex Tort Trials."