

1999

Decisionmaking About General Damages: A Comparison of Jurors, Judges, and Lawyers

Roselle L. Wissler

Allen J. Hart
Amherst College

Michael J. Saks
University of Iowa

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



Part of the [Judges Commons](#), [Law and Society Commons](#), and the [Legal Remedies Commons](#)

Recommended Citation

Roselle L. Wissler, Allen J. Hart & Michael J. Saks, *Decisionmaking About General Damages: A Comparison of Jurors, Judges, and Lawyers*, 98 MICH. L. REV. 751 (1999).

Available at: <https://repository.law.umich.edu/mlr/vol98/iss3/3>

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

DECISIONMAKING ABOUT GENERAL DAMAGES: A COMPARISON OF JURORS, JUDGES, AND LAWYERS

*Roselle L. Wissler**

*Allen J. Hart***

*Michael J. Saks****

I. THE DEBATE OVER JURY PERFORMANCE	752
II. GENERAL DAMAGES: THE LAW AND THE EMPIRICAL RESEARCH.....	757
A. <i>The Law of General Damages</i>	757
B. <i>Empirical Research on Damages Awards</i>	758
III. THE PRESENT RESEARCH	762
A. <i>Overview</i>	762
B. <i>Design</i>	764
1. <i>Case Summaries</i>	764
2. <i>Sampling Locations</i>	766
C. <i>Respondents</i>	766
1. <i>Jurors</i>	766
2. <i>Judges</i>	767
3. <i>Lawyers</i>	768
D. <i>Procedures</i>	768
E. <i>Measures</i>	769
1. <i>Awards for General Damages</i>	769
2. <i>Injury Perceptions</i>	770
3. <i>Sociodemographic Characteristics</i>	771
F. <i>Analyses</i>	771
1. <i>Overview</i>	771
2. <i>Order Effects</i>	772

* Independent Research Consultant. Formerly, Co-Principal Investigator of the Non-Economic Damages Decisionmaking Project at the University of Iowa College of Law and Department of Psychology. Ph.D. 1986, Boston College. — Ed. This research was supported by grant No. SBR-9422789 from the Law and Social Sciences Program of the National Science Foundation. The points of view expressed here are those of the authors and do not necessarily represent the position of the National Science Foundation. The authors thank George Woodworth, David Baldus, David Evans, Will Guild, Beth Bing, Justice James Heiple, and Dean Harry Wellington for assistance with various phases of this study. Address correspondence to Michael J. Saks, College of Law, University of Iowa, Iowa City, IA 52242; michael-saks@uiowa.edu.

** Assistant Professor of Psychology, Amherst College. Ph.D. 1991, Harvard University. — Ed.

*** Edward F. Howrey Professor of Law and Professor of Psychology, the University of Iowa. Ph.D. 1975, Ohio State University; M.S.L. 1985, Yale Law School. — Ed.

- IV. RESULTS 772
 - A. *Assessments of Overall Injury Severity: The Role of Injury Attributes and Sociodemographic Characteristics*..... 773
 - 1. *Jurors’ Assessments*..... 773
 - 2. *Judges’ Assessments* 776
 - 3. *Plaintiffs’ Lawyers’ Assessments* 778
 - 4. *Defense Lawyers’ Assessments* 780
 - 5. *Comparison of Injury Severity Models Across the Four Groups*..... 782
 - B. *Predicting General Damages Awards: Translating Injury Judgments into Dollars* 783
 - 1. *Jurors’ Awards* 783
 - 2. *Judges’ Awards*..... 786
 - 3. *Plaintiffs’ Lawyers’ Awards* 788
 - 4. *Defense Lawyers’ Awards*..... 791
 - 5. *Comparison of Award Models Across the Four Groups*..... 794
 - C. *Comparing Jurors’, Judges’, and Lawyers’ Injury Severity Assessments, Awards, and Predictions of the “Average Juror’s” Awards*..... 796
 - 1. *Comparison of Overall Severity Assessments*..... 797
 - 2. *Comparison of Awards*..... 798
 - 3. *Comparison of Each Group’s Award with Its Prediction of the “Average Juror’s” Award*..... 800
 - D. *The Stabilizing Power of Aggregation into Groups*..... 801
- V. CONCLUSIONS 804
 - A. *Perceptions of Injury Severity* 804
 - B. *General Damages Awards*..... 806
 - C. *Judges’ and Lawyers’ Predictions of the Awards of the “Average Juror”* 811
 - D. *Policy Implications*..... 812
 - 1. *Whether the Role of the Jury Should Be Reduced* 812
 - 2. *Reforms Consistent with the Available Data* 816
 - E. *Postscript*..... 817
- APPENDIX A 819
- APPENDIX B 820
- APPENDIX C 825

I. THE DEBATE OVER JURY PERFORMANCE

Placing important decisions in the hands of the civil jury — made up of ordinary citizens untrained in the law — has long been criticized. For example, Erwin Griswold, law school dean and Solicitor General of the United States, asked, “Why should anyone think that 12 persons brought in from the street, selected in various ways, for their lack of

general ability, should have any special capacity for deciding controversies between persons?"¹ And Jerome Frank, law professor, aggressive legal realist, and judge, argued that juries are uncertain, capricious, and unpredictable, ignorant and prejudiced, poor factfinders, gullible, and incapable of following complex legal rules,² thus making "the orderly administration of justice virtually impossible."³ The great weakness of even the most thoughtful of these criticisms is that they are based on little more than anecdote⁴ and assertion, rather than anything resembling systematic empirical evidence.

The first major study of jury decisionmaking was designed with the awareness that any assessment of the jury ultimately had to be made in comparison with judges, the decisionmakers who would replace the jury in any cases juries would not be permitted to decide.⁵ Accordingly, Harry Kalven and Hans Zeisel asked judges in 8,000 civil and criminal cases sampled from around the United States how they would have decided each case if it were a bench trial instead of a jury trial.⁶ By comparing these responses to the actual jury verdicts, Kalven and Zeisel found that judges and juries agreed on the verdict in personal injury cases 79% of the time.⁷ They also found that when jurors and judges disagreed, the jury was no more inclined than the judge to favor plaintiffs over defendants,⁸ that the agreement-disagreement ratios remained the same for both easy and difficult cases (suggesting no lack

1. 1962-1963 HARVARD LAW SCHOOL DEAN'S REPORT 5-6, *quoted in* HARRY KALVEN, JR. & HANS ZEISEL, *THE AMERICAN JURY* 5 (1966).

2. *See* JEROME FRANK, *LAW AND THE MODERN MIND* 183-99 (6th ed., Anchor Books 1963).

3. *Id.* at 195; *see also* JEROME FRANK, *COURTS ON TRIAL* 111-25 (1949).

4. *See* Michael J. Saks, *Do We Really Know Anything About the Behavior of the Tort Litigation System — And Why Not?*, 140 U. PA. L. REV. 1147, 1161 (1992) ("The trouble with legislation by anecdote is not just that some of them are false or misleading. Even if true and accurate, anecdotes contribute little to developing a meaningful picture of the situation about which we are concerned. . . . The proportion of cases that results in one or the other error, and the ratio of one kind of error to the other, ought to be of greater interest to serious policy-makers than a handful of anecdotes on either side of the issue."); Marc Galanter et al., *How to Improve Civil Justice Policy*, 77 JUDICATURE 185, 185 (1994) ("Improving the civil justice system requires thoughtful, objective analysis based on sound empirical data. The lack of systematic, cumulative data in this area makes it possible for far-reaching policy proposals to be advanced on the basis of tendentious anecdotes and numbers. A bias in which solutions to perceived problems are developed by reference to unusual and atypical cases goes unchallenged. Not surprisingly, the effects of the resulting policies are often unanticipated."), *quoted in* Brian J. Ostrom et al., *A Step Above Anecdote: A Profile of the Civil Jury in the 1990s*, 79 JUDICATURE 233, 234 (1996).

5. *See generally* Harry Kalven, Jr., *The Dignity of the Civil Jury*, 50 VA. L. REV. 1055, 1055-56, 1056 n.3 (1964).

6. *See id.* at 1063.

7. *See id.* at 1065.

8. *See id.*

of comprehension by the juries),⁹ and that the judges rarely thought the juries' verdicts were unreasonable in light of the available evidence and applicable law.¹⁰ In fact, the only noteworthy difference was that when juries did find liability, they awarded on average 20% more than judges said they would have awarded.¹¹ Even that, however, was a more modest difference than critics of the jury would have had us believe. Perhaps because Kalven and Zeisel's research findings provided so little support to the usual criticisms, the critics fell silent for a generation.

But by the mid-1980s, criticism of the civil jury returned with vigor, part of a more wide-ranging attack on the civil justice system (in particular, the tort system).¹² "One of the most criticized aspects of jury behavior" has been juries' awards of general damages, also known as noneconomic damages or by the synecdoche "pain and suffering" awards.¹³ Unlike special or "economic" damages, the jury receives little guidance regarding what a correct general damages award might be. And, unlike punitive damages, which are similarly lacking in guidance but which are rarely awarded in accident cases, general damages are proper in every personal injury case in which awards are made.

9. See *id.* at 1066.

10. See *id.* at 1067.

11. See *id.* at 1065.

12. In general these attacks were no more grounded in empirical research than earlier assaults. Where data were mentioned, their interpretation often was rather careless. For extensive discussions and evaluations of that body of "evidence," see STEPHEN DANIELS & JOANNE MARTIN, CIVIL JURIES AND THE POLITICS OF REFORM (1995); Marc Galanter, *Jury Shadows: Reflections on the Civil Jury and the "Litigation Explosion,"* in THE AMERICAN CIVIL JURY: FINAL REPORT OF THE 1986 CHIEF JUSTICE EARL WARREN CONFERENCE ON ADVOCACY IN THE UNITED STATES 15 (1987); Robert E. Litan, *Introduction,* in VERDICT: ASSESSING THE CIVIL JURY SYSTEM (Robert E. Litan ed., 1993); Robert J. MacCoun, *Inside the Black Box: What Empirical Research Tells Us About Decisionmaking by Civil Juries,* in VERDICT: ASSESSING THE CIVIL JURY SYSTEM 137 (Robert E. Litan ed., 1993); Peter H. Schuck, *Mapping the Debate on Jury Reform,* in VERDICT: ASSESSING THE CIVIL JURY SYSTEM 306 (Robert E. Litan ed., 1993); NEIL VIDMAR, MEDICAL MALPRACTICE AND THE AMERICAN JURY: CONFRONTING THE MYTHS ABOUT JURY INCOMPETENCE, DEEP POCKETS AND OUTRAGEOUS DAMAGE AWARDS (1995); Stephen Daniels, *The Question of Jury Competence and the Politics of Civil Justice Reform: Symbols, Rhetoric, and Agenda Building,* LAW & CONTEMP. PROBS., Autumn 1989, at 269, 292-308; Marc Galanter, *The Day After the Litigation Explosion,* 46 MD. L. REV. 3 (1986); Marc Galanter, *News from Nowhere: The Debased Debate on Civil Justice,* 71 DENV. U. L. REV. 77 (1993); Marc Galanter, *Real World Torts: An Antidote to Anecdote,* 55 MD. L. REV. 1093 (1996); Saks, *supra* note 4; and Neil Vidmar, *The Performance of the American Civil Jury: An Empirical Perspective,* 40 ARIZ. L. REV. 849 (1998).

13. Neil Vidmar, *Pap and Circumstance: What Jury Verdict Statistics Can Tell Us About Jury Behavior and the Tort System,* 28 SUFFOLK U. L. REV. 1205, 1224 (1994); see also Oscar G. Chase, *Helping Jurors Determine Pain and Suffering Awards,* 23 HOFSTRA L. REV. 763, 768-69 (1995) (asserting unexplainable variation among awards of pain and suffering in wrongful death cases); Kirk B. Johnson et al., *A Fault-Based Administrative Alternative for Resolving Medical Malpractice Claims,* 42 VAND. L. REV. 1365, 1369-70 (1989) (asserting unjustified variation in awards for comparable injuries caused by medical malpractice and other tortious acts).

Moreover, general damages constitute about fifty-five to sixty percent of personal injury damages awards.¹⁴ It is no surprise, then, that general damages have been at the center of contemporary debates on juries and the civil justice system.

The critics of the civil jury assert that awards generally are unreliable, capricious, and "out of control."¹⁵ More specifically, general damages are alleged to be both excessive and influenced by inappropriate considerations, such as sympathy for the victims and the defendant's apparent ability to pay.¹⁶ By implication then, jury decisions about general damages are thought to be insufficiently responsive to proper considerations, namely, the nature of the injury and its impact on the plaintiff's life. Moreover, implicit in these allegations is a problem that Jerome Frank had made explicit: that jurors do not think about cases the same way judges and lawyers do.

Although research has begun to focus on damages, these studies have thus far left the basic issues underlying the criticisms of jury awards largely unanswered.¹⁷ These basic issues are how jurors perceive injuries and translate those perceptions into dollar awards, and whether jurors' decisionmaking processes differ fundamentally from those of judges (and lawyers).¹⁸ The research reported in this Article investigates jurors' perceptions of injury seriousness and awards of general damages, and, like Kalven and Zeisel, compares them to judges' and lawyers' responses to the same injuries. The results of this most basic comparison have significant implications for criticisms of jury awards and for legal policy concerning jury trials.¹⁹

14. See Neil Vidmar et al., *Jury Awards for Medical Malpractice and Post-Verdict Adjustments of Those Awards*, 48 DEPAUL L. REV. 265, 296 (1998).

15. See criticisms reviewed in works at *supra* note 12.

16. Documentation of the various allegations can be found in the works cited *supra* note 12.

17. Much of the empirical research literature on compensatory damages is discussed in our review of the relevant literature, *infra*, Section II.B.

18. Despite the well known research strategy of Kalven & Zeisel, only a few studies have directly compared the performance of jurors to judges. See Edmund S. Howe & Thomas C. Loftus, *Integration of Intention and Outcome Information by Students and Circuit Court Judges: Design Economy and Individual Differences*, 22 J. APPLIED SOC. PSYCHOL. 102 (1992); Stephen Landsman & Richard Rakos, *A Preliminary Inquiry into the Effect of Potentially Biasing Information on Judges and Jurors in Civil Litigation*, 12 BEHAV. SCI. & L. 113 (1994); Gary L. Wells, *Naked Statistical Evidence of Liability: Is Subjective Probability Enough?*, 62 J. PERS. SOC. PSYCHOL. 739 (1992). None of these addressed, however, the question of damages. Another study contains a partial comparison (6 judges in a group composed mostly of lawyers). See Neil Vidmar & Jeffrey J. Rice, *Assessments of Non-economic Damage Awards in Medical Negligence: A Comparison of Jurors with Legal Professionals*, 78 IOWA L. REV. 883 (1993). All of these studies found no differences in decisionmaking between judges and jurors.

19. A more detailed explication of the research questions is presented *infra* Section III.A, and a more detailed discussion of the policy implications of the findings appears *infra* Section V.D.

To provide the context for our research, Part II of this Article briefly summarizes relevant aspects of the law and the empirical research literature. The law governing general damages offers no real guidance to jurors or anyone else called upon to award damages at trial. Empirical research shows that jurors display a substantial amount of sensitivity to the relative seriousness of injuries (vertical equity) in awarding general damages, but also a considerable amount of variation within levels of severity (horizontal inequity). This body of research, however, has some conceptual and methodological weaknesses; thus, important empirical questions remain.

Part III describes the research methods and design of our study. In brief, we sampled hundreds of jurors, judges, plaintiffs' lawyers, and defense lawyers from urban and rural regions of two different states. Through telephone interviews, we presented dozens of different injuries (two per interviewee) to the respondents and had them evaluate the seriousness of the injuries and specify what general damages award seemed appropriate for the injury suffered.

We used the survey responses to develop predictive (and arguably, explanatory) models of the decisionmaking of each group. We then used these statistical models to determine for each group which attributes of the injuries had what degree of impact on severity judgments and on awards, as well as how differences in the decisionmakers' geography, demography, and experiences affected their decisions. Part IV presents the results of these analyses. In short, the models evidence a remarkable degree of similarity among all groups of decisionmakers, plus a high degree of predictability, in regard to their judgments of injury severity. As to the translation of injury perceptions into monetary awards, however, more differences among the groups appeared, and the predictive power of the models declined.

Part V presents our conclusions with respect to both the empirical findings and their policy implications. Taken together, the findings suggest that the differences between jurors' awards and those of the other groups do not reflect fundamental differences in decisionmaking, but rather a loss of consistency in translating perceptions of severity into damages awards. That loss of consistency likely can be attributed to the fact that in assessing the value of a single case, jurors lack the frame of reference created by other cases that is readily available to judges and lawyers. If that diagnosis is correct, it suggests that some rather modest changes in procedure are likely to be sufficient to enable jurors' general damages awards to closely approximate those of judges in stability and predictability.

II. GENERAL DAMAGES: THE LAW AND THE EMPIRICAL RESEARCH

A. *The Law of General Damages*

In tort cases, general damages are awarded to compensate plaintiffs for physical pain, mental suffering, disability, disfigurement, loss of enjoyment of life, and other similar harms.²⁰ While these “noneconomic” damages may not be inherently different from “economic” damages such as lost earnings and medical expenses, markets exist to supply the values of the latter but not the former.²¹ If art can acquire a market value, in principle, so can the loss of sight. The problem is that people do not normally buy and sell the sorts of sensory, motor, cognitive, and other capacities that are injured in accidents. The result is that to compensate for noneconomic losses, the law must turn to an alternative source of values, namely, the social judgment of the community, typically supplied by juries. A task that involves assigning a value to the virtually undefinable is, by definition, challenging to perform and at least as challenging to evaluate. “One of the most difficult decisions facing the jury in a personal injury case is the size of the monetary award for pain and suffering, since there is no objective method of evaluating such damages.”²²

The law itself provides essentially no guidance to jurors in reaching a “correct” general damages award. For instance, the Restatement (Second) of Torts notes that “[t]he discretion of the judge or jury determines the amount of recovery, the only standard being such an amount as a reasonable person would estimate as fair compensation.”²³ In addition, the law requires that damages be particularized to both the details of the objective injury and its subjective effect on the

20. See generally RESTATEMENT (SECOND) OF TORTS § 905 (1979); 22 AM. JUR. 2D *Damages* § 28 (1988).

21. See Ted R. Miller, *Willingness to Pay Comes of Age: Will the System Survive?*, 83 NW. U. L. REV. 876, 896-900 (1989); W. Kip Viscusi, *Pain and Suffering in Product Liability Cases: Systematic Compensation or Capricious Awards?*, 8 INT'L REV. L. & ECON. 203 (1988).

22. AM. JUR. 2D *Damages* § 260 (1988); see also R. Perry Sentell, Jr., *The Georgia Jury and Negligence: The View from the Trenches*, 28 GA. L. REV. 1, 97-98 (1993) (summarizing a survey of Georgia plaintiffs' and defense lawyers which found that they thought the most troublesome issue for the jury is a lack of guidance on how to assess damages, especially for pain and suffering); Neil Vidmar, *Empirical Evidence on the Deep Pockets Hypothesis: Jury Awards for Pain and Suffering in Medical Malpractice Cases*, 43 DUKE L.J. 217 (1993).

23. RESTATEMENT (SECOND) OF TORTS § 912 cmt. b (1979).

individual victim.²⁴ By implication, then, the law expects awards to reflect considerable variability.²⁵

Of course, the law has as little guidance to offer judges and lawyers in this pursuit as it does to juries. Nonetheless, critics of juries seem to believe that judges and lawyers somehow acquire judgmental capabilities that juries lack, producing judgments which are somehow more "accurate," or at least more consistent, than those provided by jurors.

B. *Empirical Research on Damages Awards*

Researchers have sought to identify the determinants of damages awards. In archival studies, researchers have found that from half to almost three-fourths of the variation in awards actually made by juries (and sometimes by judges) can be accounted for by various measures of injury severity, such as medical expenses or ratings of severity and duration, as well as by other predictors.²⁶ The measures of injury type, severity, and permanence are consistently the strongest predictors of awards, with injuries of greater severity and duration tending to receive more compensation than less severe and more temporary injuries. Similarly, simulation studies, which examine mock jurors' awards in response to experimentally manipulated cases, also find the size of the award to correspond to the severity of the injury.²⁷

Despite these substantial predictive relationships, the findings suggest there are both random and systematic errors in awards. The same

24. See generally David W. Leebron, *Final Moments: Damages for Pain and Suffering Prior to Death*, 64 N.Y.U. L. REV. 256, 265 (1989).

25. Nevertheless, many trial participants and commentators wish for less variability and more predictability than the law thus far has been able to design.

26. See AUDREY CHIN & MARK A. PETERSON, DEEP POCKETS, EMPTY POCKETS: WHO WINS IN COOK COUNTY JURY TRIALS (1985); MARK A. PETERSON, COMPENSATION FOR INJURIES: CIVIL JURY VERDICTS IN COOK COUNTY (1984); David Baldus et al., *Improving Judicial Oversight of Jury Damage Assessments: A Proposal for the Comparative Additur/Remittitur Review of Awards for Nonpecuniary Harms and Punitive Damages*, 80 IOWA L. REV. 1109 (1995); Randall Bovbjerg et al., *Juries and Justice: Are Malpractice and Other Injuries Created Equal?*, LAW & CONTEMP. PROBS., Winter 1991, at 5 [hereinafter Bovbjerg et al., *Juries and Justice*]; Randall R. Bovbjerg et al., *Valuing Life and Limb in Tort: Scheduling "Pain and Suffering,"* 83 NW. U. L. REV. 908 (1989) [hereinafter Bovbjerg et al., *Valuing Life and Limb*]; Frank A. Sloan & Chee Ruey Hsieh, *Variability in Medical Malpractice Payments: Is the Compensation Fair?*, 24 LAW & SOC'Y REV. 997 (1990); Viscusi, *supra* note 21.

27. See Jane Goodman et al., *Runaway Verdicts or Reasoned Determinations: Mock Juror Strategies in Awarding Damages*, 29 JURIMETRICS 285, 289 (1989); Valerie Hans & M. David Ermann, *Responses to Corporate Versus Individual Wrongdoing*, 13 LAW & HUM. BEHAV. 151 (1989); Michael J. Saks et al., *Reducing Variability in Civil Jury Awards*, 21 LAW & HUM. BEHAV. 243 (1997); Vidmar & Rice, *supra* note 18; Roselle L. Wissler et al., *Explaining "Pain and Suffering" Awards: the Role of Injury Characteristics and Fault Attributions*, 21 LAW & HUM. BEHAV. 181 (1997). But see Corinne Cather et al., *Plaintiff Injury and Defendant Reprehensibility: Implications for Compensatory and Punitive Damage Awards*, 20 LAW & HUM. BEHAV. 189 (1996) (finding no difference in awards as a function of injury severity).

studies that find high correlations between injury severity and awards (termed "vertical equity") also report considerable variation in the amounts awarded for injuries at the same level of severity (referred to as "horizontal inequity"),²⁸ which suggests the presence of a good deal of random error.²⁹

Systematic error in awards is seen in a number of studies which have found that, despite general vertical equity, plaintiffs with small losses tend to be overcompensated and those with large losses tend to be undercompensated.³⁰ Moreover, several factors which should not have an effect on general damages awards, such as type of defendant (e.g., individual versus corporation) or type of case (e.g., auto crash versus medical malpractice)³¹ nevertheless seem to systematically influence the size of awards.³² In addition, some evidence suggests that where jurors reside influences their awards — that jurors in urban areas generally make larger awards than those in rural areas,³³ and that

28. See CHIN & PETERSON, *supra* note 26; Sloan & Hsieh, *supra* note 26; Viscusi, *supra* note 21.

29. In simulation studies, for instance, mock jurors are presented with identical case facts, yet the jurors produce an array of different awards, not one single award. See sources cited *supra* note 27. Of course, a complete absence of variability among a group of decisionmakers seldom occurs. For a glimpse of the variability of other decisionmakers, see *infra* Table 5. Accordingly, commentators who lament "horizontal inequity" cannot be suggesting that its existence demonstrates a defect in the system of damages. What they really ought to be arguing is that the amount of horizontal variability is greater than they believe it should be.

30. See ALFRED F. CONARD ET AL., *AUTOMOBILE ACCIDENT COSTS AND PAYMENTS: STUDIES IN THE ECONOMICS OF INJURY REPARATION* at 248-52 (1964); ELIZABETH M. KING & JAMES P. SMITH, *ECONOMIC LOSS AND COMPENSATION IN AVIATION ACCIDENTS* at 74-76 (1988).

31. See CHIN & PETERSON, *supra* note 26; PETERSON, *supra* note 26; Baldus et al., *supra* note 26; Bovbjerg et al., *Juries and Justice*, *supra* note 26; Bovbjerg et al., *Valuing Life and Limb*, *supra* note 26; James K. Hammit et al., *Tort Standards and Jury Decisions*, 14 J. LEGAL STUD. 751 (1985); Viscusi, *supra* note 21; David C. Baldus et al., *Children's Models for Compensatory Damages* (1992) (unpublished manuscript, on file with the authors).

32. Simulation studies, however, have added important qualifications and refinements to conclusions about the way in which these factors affect awards. Most importantly, they have found that differences in compensatory awards previously attributed to a "deep pocket" effect are not the result of differences in the defendants' wealth. See Robert J. MacCoun, *Differential Treatment of Corporate Defendants by Juries: An Examination of the "Deep-Pockets" Hypothesis*, 30 LAW & SOC'Y REV. 121 (1996); Neil Vidmar, et al., *Damage Awards and Jurors' Responsibility Ascriptions in Medical Versus Automobile Negligence Cases*, 12 BEHAV. SCI. & L. 149 (1994); Vidmar, *supra* note 22; William D. Woody et al., *Compensating Plaintiffs and Punishing Defendants: Is Bifurcation Necessary?* (1998) (unpublished manuscript, on file with the authors) (presented at the biennial meeting of the American Psychology-Law Society, Redondo Beach, Cal., 1998). Instead, the differences in awards are due to differences in perceptions of the defendants' intentionality, responsibility, recklessness, or competence to avoid the injury. See Valerie P. Hans, *The Jury's Response to Business and Corporate Wrongdoing*, 52 LAW & CONTEMP. PROBS., Autumn 1989, at 177; Hans & Ermann, *supra* note 27; see also Goodman, et al., *supra* note 27; MacCoun, *supra* note 12; Vidmar et al., *supra* note 14.

33. See PATRICIA M. DANZON, *MEDICAL MALPRACTICE: THEORY, EVIDENCE, AND PUBLIC POLICY* 74-75 (1985).

jurors in some states characteristically make smaller or larger awards than those in other states.³⁴ Thus, awards may depend on more than the characteristics of injuries and their sequelae.

For answering questions of how jurors respond to the attributes of the injuries in making their general damages awards, the archival studies are insufficient in one or more ways. Typically, most jury awards are reported as single aggregate awards, instead of being separated into their component parts. Thus, researchers generally use the total, undifferentiated award, not the specific general damages component.³⁵ And efforts by some researchers to disaggregate the awards after the fact³⁶ tend to produce erroneous estimates of the general component.³⁷

Additionally, the degree of horizontal inequity suggested by many of the studies may reflect the limitations of the measurements used as much as any random error in jurors' decisions. For example, the severity of injuries usually has been measured by grading a vast array of injuries on a scale of no more than eight or nine points.³⁸ Each point on such a scale necessarily captures a range of injuries having different degrees of severity.³⁹ That jurors would rate the injuries more finely, and award different amounts to injuries that researchers have treated as identical, is hardly a failing of juror decisionmaking.

Further, these scales are unidimensional measures, though injuries, like most social objects and events, occupy a multidimensional space. For example, on the National Association of Insurance Commissioners' ("NAIC") Severity of Injury Scale, burns, surgical material left in a patient, drug side effects, and brain damage all are scored as Level 4 injuries. The Procrustean compression by researchers into a single dimension of what the decisionmakers almost certainly saw as a multidimensional injury squanders explanatory potential.⁴⁰ Evidence of this

34. See DANIELS & MARTIN, *supra* note 12.

35. See CHIN & PETERSON, *supra* note 26; PETERSON, *supra* note 26; Bovbjerg et al., *Juries and Justice*, *supra* note 26; Sloan & Hsieh, *supra* note 26.

36. See Bovbjerg et al., *Valuing Life and Limb*, *supra* note 26; Viscusi, *supra* note 21.

37. See Neil Vidmar, *Making Inferences About Jury Behavior from Jury Verdict Statistics: Cautions About the Lorelei's Lied*, 18 LAW & HUM. BEHAV. 599 (1994); Vidmar et al., *supra* note 14.

38. Some researchers use the Severity of Injury Scale developed by the National Association of Insurance Commissioners, see NAIC MALPRACTICE CLAIMS: FINAL COMPILATION (1980), which ranges from one (emotional injury only) to nine (death). See DANZON, *supra* note 33; Bovbjerg et al., *Juries and Justice*, *supra* note 26.

39. See Oscar G. Chase, *Helping Jurors Determine Pain and Suffering Awards*, in REFORMING THE CIVIL JUSTICE SYSTEM 339, 354 (Larry Kramer ed., 1996); Gary T. Schwartz, *Proposals for Reforming Pain and Suffering Awards*, in REFORMING THE CIVIL JUSTICE SYSTEM 416, 419 (Larry Kramer ed., 1996).

40. See Allen J. Hart et al., *Multidimensional Perceptions of Illness and Injury*, 2 CURRENT RES. SOC. PSYCHOL. 30 (Sept. 11, 1997) <<http://www.uiowa.edu/~grpproc/crisp/crisp.html>>; Andrew J. Slain et al., *Multidimensional Perceptions of Injury: Implications for*

problem is provided by data from simulation jurors, whose assessments of injuries on four dimensions (disability, mental suffering, disfigurement, pain) better predicted their general damages awards (accounting for 74% of the variance) than did a single predictor based on the same jurors' assessments of the overall severity of injury (61% of the variance in awards), which in turn was a better predictor than the NAIC scale (44%).⁴¹

In addition to poor measurement of injuries, the studies finding substantial horizontal inequity also fail to take into account legally relevant differences among injury victims.⁴² For example, the exact same permanent loss of hearing *should* produce a larger general damages award for a musician than for a novelist. Moreover, a permanent injury should result in a larger award to compensate a 20-year-old plaintiff for 60 years of future pain and suffering than to compensate a 60-year-old plaintiff for 20 years.⁴³ All of these are examples of vertical equity mistaken for horizontal inequity.

Of course, to demonstrate that horizontal inequity has been overestimated in the literature is not to say that real sources of horizontal inequity, both random and systematic, do not exist. But we should not mistake one problem for another. And we certainly should not mistake the limitations of the studies for the limitations of juror decision-making.

In all of these studies, it is impossible to know how well or poorly the jurors have done because there are almost no data for answering the inescapable question of "compared to what?" As Kalven and Zeisel noted long ago, in the study of juries, the usual answer to that question is: compared to the decision a judge would make in the same case, because judges are the most likely alternative to jurors.⁴⁴ More recently, the answer might include alternative dispute resolution forums in which the decisionmakers might be lawyers (in the role of arbitrators) or other professionals (as members of compensation boards). Studies using a straightforward comparison of damages awarded by judges and by juries suffer from case selection confounds because, owing to the strategic choices of lawyers, different mixes of cases go to juries than go to judges.⁴⁵ Two recent studies overcame

the "Adversary Culture" (1998) (unpublished manuscript, on file with the authors) (presented at the Biennial Meeting of the American Psychology-Law Society, Redondo Beach, Cal., 1998).

41. See Wissler et al., *supra* note 27, at 195-202 (Experiment 2).

42. Jurors evaluating general damages may and should take into account such things as pre-injury health, post-injury recovery, lifestyle, occupation, and social circumstances. See RESTATEMENT (SECOND) OF TORTS § 905 cmt. i (1979).

43. See RESTATEMENT (SECOND) OF TORTS § 905 (1979).

44. See HARRY KALVEN, JR. & HANS ZEISEL, *THE AMERICAN JURY* (1966).

45. See Kevin M. Clermont & Theodore Eisenberg, *Trial by Jury or Judge: Transcending Empiricism*, 77 CORNELL L. REV. 1124 (1992).

this problem by presenting the same case facts to arbitrators⁴⁶ and to senior lawyers,⁴⁷ as well as to jurors. Both studies found no significant differences in the general damages awards given by the different types of decisionmakers.⁴⁸

Significantly, prior studies tell researchers and reformers little about how jurors decide awards. As Goodman notes, “[l]ittle is known about the strategies or cognitive processes used by jurors in determining damages.”⁴⁹ Such research is needed because “although critics tend to focus on the *outcomes* of jury decisionmaking — liability verdicts and damages awards — effective evaluation and implementation of policy also require an understanding of the *process* of jury decisionmaking.”⁵⁰

Thus, fundamental questions about jurors’ decisionmaking about general damages remain unanswered. What is the content and organization of people’s perceptions of injuries? Given that most people have little systematic experience evaluating injuries, but generally respond with intuitive, “gut” reactions, we might expect considerable random variation across individuals. How much random variation is there? By contrast, are systematic variations to be found between social groups? Are major differences to be found in perceptions of injuries or in the special challenge of translating those perceptions into dollar awards?⁵¹ Does this translation step reveal its own systematic differences among groups? Most importantly, do judges and lawyers see or value injuries differently, or show less variability, than the ordinary citizens who serve on juries?

III. THE PRESENT RESEARCH

A. Overview

The present study asks how the core attributes of injuries affect decisionmakers’ perceptions of the injuries, how those perceptions are translated into dollar awards, and whether (and if so, how) these patterns of response differ among jurors, judges, and lawyers. This re-

46. See Vidmar & Rice, *supra* note 18.

47. See Vidmar, *supra* note 12, at 229-234.

48. See Vidmar & Rice, *supra* note 18.

49. See Goodman et al., *supra* note 27; see also Jane Goodman et al., *Money, Sex, and Death: Gender Bias in Wrongful Death Damage Awards*, 25 L. & SOC’Y REV. 263, 264 (1991) (“Little is known about factors that influence jury damage awards in civil cases.”).

50. MacCoun, *supra* note 12, at 137 (emphasis in original).

51. See Cass R. Sunstein et al., *Assessing Punitive Damages (With Notes on Cognition and Valuation in Law)*, 107 YALE L.J. 2071 (1998); Daniel Kahneman et al., *Shared Outrage and Erratic Awards: The Psychology of Punitive Damages*, 16 J. RISK & UNCERTAINTY 49 (1998) (a somewhat more technical presentation of the same study as Sunstein et al.); Wissler et al., *supra* note 27.

search took the form of a fractional factorial survey,⁵² in which numerous and diverse injury scenarios were presented to judges, lawyers, and jurors. The respondents were asked to evaluate the injuries on several dimensions, to indicate their judgment of the injury's overall severity, and to provide an appropriate award. Respondents were selected from two different states, which were chosen on the basis of previous research suggesting that the damages awards in one were characteristically high and the other characteristically low relative to other states.⁵³ Respondents also were stratified within each state by their urban or rural location, another factor previous research suggested was related to award size.⁵⁴

The present design avoids the confounding inherent in purely observational research. By presenting an identical set of injuries to jurors in both states, and in both urban and rural areas, the factorial survey permits a test of whether awards are a function of where the jurors reside (that is, economic and cultural differences between states or between rural and urban dwellers within states) or of the mix of cases to which they are responding. Similarly, by having jurors, judges, and lawyers respond to exactly the same group of injuries, the present study eliminates the confounding of the decisionmaker with the case mix. Thus, unlike passive observational studies, the present design can distinguish the attributes of the decisionmakers from the attributes of the cases they are deciding. The comparative nature of the present research allows us to see how similar or different the decisionmaking groups are in their decision processes and responses. Asking respondents to evaluate injuries on a separate set of scales and to make general damages awards allows us to build decision models of both the injury-evaluation phase of their thinking and the award-making phase, and to compare these decision processes.

The pattern of findings obtained will imply different diagnoses and different policy solutions to improving the process of awarding general damages. For example, if jurors' injury severity evaluations or damages awards were not highly predictable, but no less predictable than those of judges and lawyers, that would suggest that the problem is not properly ascribed to *jurors*, but is attributable to the nature of the decisionmaking task, and replacing jurors with judges offers no solution. Alternatively, if all groups were equally predictable, but jurors relied exclusively on one dimension of injury, such as mental suffering, while judges and lawyers relied on a definable mix of all injury attributes, then differences between decisionmaker groups would imply system-

52. See PETER H. ROSSI & STEVEN L. NOCK, *MEASURING SOCIAL JUDGMENTS: THE FACTORIAL SURVEY APPROACH* (1982).

53. See DANIELS & MARTIN, *supra* note 12.

54. See generally DANZON, *supra* note 33.

atic differences in decisions and outcomes. Similarly, if jurors showed a high degree of unpredictability in both injury evaluations and awards, but judges and lawyers showed a high degree of predictability in both, then it becomes plausible to think that replacing jurors with judges would produce more stable patterns of awards.

B. *Design*

This study is a fractional factorial survey — that is, a fractional experimental design conducted in the manner of a survey.⁵⁵ The study utilized 62 injury case summaries, grouped into 31 pairs. Each pair was heard by a subset of the respondents; each respondent heard one injury pair. Thus, injury was both a between- and within-subjects factor. The respondents were drawn from three populations — jury-eligible citizens, judges, and lawyers — stratified by rural and urban counties in the states of Illinois and New York. Respondents rated each injury on several dimensions and provided damages awards.

1. *Case Summaries*

The 62 case summaries were adapted from actual personal injury cases⁵⁶ and represented a broad spectrum of injuries. The injuries encompassed different types: nerve damage and paralysis (e.g., quadriplegia, temporary paralysis of two fingers); head injuries and brain damage (e.g., epilepsy, permanent vegetative state); sensory loss (e.g., vision, smell); cuts (e.g., on face, legs); harm to internal organs (e.g., loss of kidney, spleen); illnesses (e.g., leukemia, lung cancer); burns (e.g., second degree on lower body, third degree over 80% of body); and orthopedic injuries (e.g., fracture, sprain, amputation). We did not include cases in which the plaintiff had died or in which the plaintiff suffered only emotional distress unaccompanied by physical injury.

The case summaries included a description of the nature, severity, and location of the injury; the nature, painfulness, duration, and restrictions associated with the treatment and healing process; any anxiety, fear, embarrassment, or depression about the injury, its impact, treatment, or prognosis; the current status of the plaintiff's pain, disfigurement, mental state, and physical impairment; and the likelihood and extent of future improvement or deterioration and treatment or care needs.⁵⁷ These injury characteristics were suggested by the legal

55. See generally ROSSI & NOCK, *supra* note 52.

56. The cases were selected from those identified in a study by Baldus et al., *supra* note 26.

57. See Appendix A for several examples of case summaries.

literature,⁵⁸ workers' compensation schedules,⁵⁹ British compensation schedules for personal injury cases,⁶⁰ and empirical research in this area.⁶¹

The case summaries noted the age of the plaintiff, who was said to be 37 years old in one case in each pair and 33 years old in the other case. The plaintiff's age was specified in order to reduce the variability in awards that would result if compensation for future pain and suffering for permanent injuries were based on different life expectancies (e.g., if some respondents thought they were compensating a 20-year-old for 60 years of pain and suffering while others thought they were compensating a 60-year-old for 20 years).⁶² The mid-thirties were chosen as representative of a fairly large range of ages and unlikely to involve special considerations that might arise for young or elderly plaintiffs. In order to discover the dimensions along which people view injuries and determine awards, independent of other contextual factors, no details were supplied about the plaintiff (other than age),⁶³ about the defendant, or about the circumstances that caused the injuries.⁶⁴

58. See RESTATEMENT (SECOND) OF TORTS §§ 905, 912 cmt. b (1979); 22 AM. JUR. 2D *Damages* §§ 240-241, 252-256, 269, 271, 275, 293-387 (1988); 1 MARILYN MINZER ET AL., *DAMAGES IN TORT ACTIONS* (1993); JEROME MIRZA, *ILLINOIS TORT LAW & PRACTICE* (2d ed. 1989); 36 NY JUR. 2D *Damages* §§ 56-66, 96-99 (1984); MICHAEL J. POLELLE & BRUCE L. OTTLEY, *ILLINOIS TORT LAW* (1985); Robert W. Phillipott, Jr., ed., *Quantum Study*, 42 LOY. L. REV. 817 (1997); Carl T. Drechsler, Annotation, *Excessiveness or Adequacy of Damages Awarded for Injuries to Arms and Hands*, 12 A.L.R. 4th 96 (1981); Carl T. Drechsler, Annotation, *Excessiveness or Adequacy of Damages Awarded for Injuries to Legs and Feet*, 13 A.L.R. 4th 212 (1981).

59. See ALBERT J. MILLUS & WILLARD J. GENTILE, *WORKERS' COMPENSATION LAW AND INSURANCE* 93-194 (2d ed. 1980); JEFFREY V. NACKLEY, *PRIMER ON WORKERS' COMPENSATION* (2d ed. 1989).

60. See PETER CANE, *ATIYAH'S ACCIDENTS, COMPENSATION AND THE LAW* (4th ed. 1987); JUDICIAL STUDIES BOARD, *GUIDELINES FOR THE ASSESSMENT OF GENERAL DAMAGES IN PERSONAL INJURY CASES* (1992); DAVID A. KEMP & MARGARET S. KEMP, *QUANTUM OF DAMAGES* (1992); JOHN MUNKMAN, *DAMAGES FOR PERSONAL INJURY AND DEATH* (8th ed. 1989); David A. Kemp, *Damages for Non-Pecuniary Loss, in DAMAGES FOR PERSONAL INJURY AND DEATH* (David Kemp ed., 1993).

61. See CHIN & PETERSON, *supra* note 26; Baldus et al., *supra* note 26; Hart et al., *supra* note 40; Viscusi, *supra* note 21; Wissler et al., *supra* note 27.

62. See Baldus et al., *supra* note 26; Bovbjerg et al., *Valuing Life and Limb, supra* note 26.

63. Respondents who requested more information about the plaintiff were told to assume that the person was of average good health and had an average lifestyle before the injury.

64. A potential criticism of our methodology is that our findings, based on relatively brief injury descriptions, might not be consistent with those that would be obtained with a fuller presentation of information, such as at trial. See Vladimir J. Konecni & Ebbe B. Ebbesen, *External Validity of Research in Legal Psychology*, 3 LAW & HUM. BEHAV. 39 (1979); Wayne Weiten & Shari S. Diamond, *A Critical Review of the Jury Simulation Paradigm: The Case of Defendant Characteristics*, 3 LAW & HUM. BEHAV. 71 (1979). Few respondents commented that they did not have enough information about the injuries themselves; those who requested additional information sought contextual information that we had deliber-

2. Sampling Locations

Our sample was stratified by state and, within state, by rural and urban counties. Illinois was selected to represent a characteristically low-award state and New York a high-award state.⁶⁵ For sampling purposes, population density figures were used to designate counties as rural or urban.⁶⁶ In Illinois, Cook County (population density of about 5,000 people per square mile) was designated urban and counties with fewer than 115 people per square mile were designated rural; counties with an intermediate population density were not included. In New York state, Queens, Bronx, Kings, and New York Counties (population densities ranging between 17,000 and 53,000 people per square mile) were designated urban. For the jury-eligible sample, New York counties with fewer than 200 people per square mile were designated rural, and respondents were not drawn from counties with an intermediate population density. In order to obtain an adequate sample of judges and personal injury lawyers, however, New York counties with an intermediate population density were included among the rural counties.⁶⁷

C. Respondents

We drew representative samples from three populations: jury-eligible citizens, judges, and lawyers.

1. Jurors

Random-digit dialing was used to obtain the sample of jury-eligible (i.e., 18 years or older) citizens (hereafter referred to as "jurors").

ately withheld in order to eliminate its potential confounding effects. As discussed in Wissler et al., *supra* note 27, short as opposed to longer injury descriptions did not produce differences in injury ratings, awards, or the relationship between the two.

65. The median 1988-90 jury award in Cook County (Chicago), Illinois, was \$32,550, compared to from \$242,500 to \$417,200 in several urban New York counties. See DANIELS & MARTIN, *supra* note 12, at 70-71. When the awards were adjusted for differences in the amount in controversy subject to the jurisdiction of courts in the two states, Cook County's median award increased to \$162,316, but was still well below that of New York urban counties. Median awards in rural counties ranged from \$24,868 to \$39,180 in Illinois and from \$66,000 to \$192,500 in New York. *See id.*

66. The population density figures were obtained from the U.S. Bureau of the Census at its internet site at <www.census.gov/population/censusdata/90den_stco.txt>. These designations were used for sampling purposes only; the rural/urban variable used in the analyses relied on the respondents' self-report of where they live (jurors), where jurors in their courts live (judges), or where personal injury plaintiffs in their cases live (lawyers).

67. The population density (number of people per square mile) of these intermediate counties was as follows: twelve counties between 200 and 1000, four counties between 1000 and about 2000, and two counties between approximately 4,500 and 6,500.

Five hundred fifty-eight jurors completed an interview.⁶⁸ Accordingly, each injury pair was heard by 18 jurors. Half of the jurors (279) were from each state; within each state, two-thirds (186) were from urban counties and one-third (93) were from rural counties.⁶⁹

The jurors who completed the survey comprised a variety of education, income, and age groups. Five percent of jurors had not completed high school, 24% were high school graduates, 32% had attended college, 24% were college graduates, and 15% had done some post-college study or held an advanced degree. Eleven percent of jurors had a household income less than \$15,000; 23% from \$15,000 to \$30,000; 31% from \$31,000 to \$50,000; 20% from \$51,000 to \$75,000; and 15% over \$75,000. The jurors ranged in age from 18 to 92 (mean = 41, median = 39). We over-sampled male respondents in order to obtain relatively equal numbers of male (49%) and female (51%) respondents. Only 12% of the jurors knew someone who suffered from the condition in one of the scenarios they heard.

2. Judges

We obtained the names of state general jurisdiction civil trial judges from directories and from the court system's administrative office in each state. Two hundred forty-four judges completed an interview.⁷⁰ Accordingly, most injury pairs were heard by eight judges, but four pairs were heard by only seven judges.⁷¹ Approximately half of the judges were from each state (124 from Illinois, 120 from New York). In Illinois, approximately half of the respondents (52%) were

68. The response rate for each area was as follows: Illinois urban, 19%; Illinois rural, 28%; New York urban, 20%; New York rural, 26%. These are conservatively estimated response rates. Because the computer-assisted telephone interviewing system used to conduct the interviews counted calls, not people, we are unable to obtain more exact response rates than these.

The essential question, however, is whether the respondents are representative of the populations from which they are drawn. We found the age, income, and education distributions of our respondents to approximate fairly closely the distributions of those demographic characteristics in each state, as shown by census data. Census data were obtained through the United State Census Bureau's world wide web pages: venus.census.gov/cdrom/lookup/cmd=tables/db=C90STF3A/lev=state/FO=FIPS.state/Fl=stab.geo/sel=36,new [for New York] and venus.census.gov/cdrom/lookup/cmd=tables/db=C90STF3A/lev=state/FO=FIPS.state/Fl=stab.geo/sel=17,new [for Illinois]. The exception is that the lowest income and education groups are somewhat underrepresented among our respondents, a pattern typical of telephone survey research.

69. When asked where they live, 48% of jurors said a large city, 19% said the suburbs of a large city, 7% said a medium-sized city, and 26% said a small town or rural area. For purposes of analysis, the first two categories were designated "urban" and the latter two were designated "rural."

70. The response rate for each area was as follows: Illinois urban, 68%; Illinois rural, 91%; New York urban, 54%; New York rural, 62%.

71. We were unable to get an eighth judge for each of these four pairs due to the relatively small number of civil trial judges in New York courts and their limited availability.

from urban and half (48%) from rural counties. In New York, due to the limited number of judges in rural counties, 71% of the respondents were from urban counties and 29% were from rural counties.⁷²

Personal injury cases composed from 1% to 100% of the judges' caseload (mean = 52%, median = 50%). The judges had been hearing personal injury cases from 1 to 34 years (mean = 9, median = 7). Most of the judges (87%) were male. Fifty-seven percent of the judges either personally knew someone with the condition or had heard a case involving the injury presented in one of the scenarios they were given.

3. *Lawyers*

We obtained the names of lawyers from several different directories.⁷³ Two hundred forty-eight lawyers completed an interview.⁷⁴ Thus, each injury pair was heard by eight lawyers. Half of the lawyers (124) were from each state and, within each state, half primarily represented plaintiffs and half primarily represented defendants. In each state, over half of the respondents (56%) were from urban counties and almost half (44%) were from rural counties.

Personal injury cases composed from 1% to 100% of the lawyers' caseload (mean = 69%, median = 80%). The lawyers had been handling personal injury cases from 2 to 61 years (mean = 22, median = 20). Sixty percent of the lawyers had never served as an arbitrator in a personal injury case, 14% had done so rarely, 20% occasionally, and 6% often. Although we over-sampled women lawyers, 90% of those interviewed were men. Seventy percent of the lawyers either personally knew someone with the condition or had been involved in a case in which the plaintiff had the injury presented in one of the scenarios they were given.

D. *Procedures*

Prior to calling the judges and lawyers, we sent each a letter explaining the purpose of the study, the time involvement, and the confidentiality of their responses, and advising them of the upcoming phone call. The mailing included a letter from a prominent member of

72. When asked where most of the jurors in their courts live, 45% of all judges said a large city, 17% said the suburbs of a large city, 9% said a medium-sized city, and 29% said a small town or rural area. Thus, across both states, 62% of judges were from urban courts and 38% were from rural courts.

73. These were the Membership Directory of the Defense Research Institute, the Desk Reference of the Association of Trial Lawyers of America, the Martindale-Hubbell Law Directory, and West's Legal Directories.

74. The response rate for each area was as follows: Illinois urban, 83%; Illinois rural, 86%; New York urban, 70%; New York rural, 74%.

their state's legal community urging their participation in our research. Jurors did not receive a letter in advance.

All respondents were contacted by telephone and were asked to participate in a study of people's decisions about how much monetary compensation an injured person should receive for general damages. Respondents were told that they would hear summaries of two cases and would be asked to give their reactions to each. Respondents also were told that, in an earlier trial, a jury had found the defendant to be responsible for the plaintiff's injury and had compensated the plaintiff for medical expenses and lost earnings. We emphasized that the respondent was to focus on how much pain and suffering the injury has caused the plaintiff and how much money would provide fair compensation for those general damages.

The interviewer read the respondent the first case summary, followed by jury instructions that were adapted from standard instructions used in each state.⁷⁵ Although instructions for the two states differed somewhat and therefore created a slight confounding of instructions and population, using instructions that correctly reflected the law in each state was necessary to have credible instructions for the judges and lawyers.

Respondents then were asked to indicate how much money they would award the plaintiff for general damages, how much they thought the average juror would award the plaintiff, and to rate the plaintiff's injury on five dimensions. This entire procedure was then repeated for the second case. Finally, the respondents were asked sociodemographic questions.⁷⁶

E. *Measures*

1. *Awards for General Damages*

A review of the awards for general damages for each injury revealed that the highest award for some injuries was more than five times as large as the second highest award.⁷⁷ Of the 62 injuries, these

75. See Appendix B. The instructions used for Illinois respondents were adapted from ILLINOIS PATTERN JURY INSTRUCTIONS — CIVIL, §§ 30.00-30.05 (3d ed. 1994). The instructions used for New York respondents were adapted from NEW YORK PATTERN JURY INSTRUCTIONS - CIVIL § 2:280 (2d ed. 1994).

76. Appendix B contains the interview protocol used.

77. Respondents who gave an award of zero were dropped from the study if the reason for their award was (a) they felt the defendant was not responsible (despite the fact that the stimulus materials had stated that a prior jury already had established the defendant's liability), or (b) they did not believe in the concept of compensation for general damages. If a potential juror in an actual trial expressed the latter attitude during jury selection, the person would be excluded from the jury. During the trial, the judge would admonish a juror expressing this attitude that the law had established the plaintiff's right to compensation for general damages and that one's duty as a juror was to determine a fair amount. Respondents who gave a zero dollar award were retained if they gave reasons such as: the pain was

outliers were found in 23 injuries for jurors, 5 for judges, and 6 for lawyers. Each outlier was recoded to the value at three standard deviations above the mean of the distribution (calculated without the outlier). If that recoded value would have altered the rank order of the awards, however, the outlier was recoded instead to the value at the next quarter-standard-deviation increment above the second highest award. The recoding of outliers to 3 or 4 standard deviations above the mean reduced their undue influence on the distribution without the loss of data.⁷⁸

After this adjustment, the lawyers' and judges' awards for each injury showed no serious departures from normality. The jurors' awards for some injuries, however, were moderately positively skewed. Accordingly, we applied a square root transformation to the awards of all three groups.⁷⁹ The square root transformation adequately reduced the skewness of the jurors' awards without overcorrecting the judges' and lawyers' awards, providing a comparable award measure for the three groups.⁸⁰

2. Injury Perceptions

Respondents were asked to give their impressions of the plaintiff's injuries on several dimensions, each rated on a five-point, Likert-type scale. Specifically, respondents were asked to indicate how much physical pain, mental suffering, disability, and disfigurement the given injury will have caused the plaintiff. For each of these dimensions, respondents were asked to consider both its amount and duration. Respondents also were asked to indicate how severe the injury was overall.

only temporary; the injury was not disabling, not serious, or caused no permanent damage; or the plaintiff was unconscious and, thus, did not experience pain and suffering.

78. See generally BARBARA G. TABACHNICK & LINDA S. FIDELL, USING MULTIVARIATE STATISTICS (1983).

79. The purpose of transforming non-normal distributions is to make them more nearly normal in order to use statistical procedures that require normality; to use highly skewed distributions could lead to misleading results. It is important to realize that correlations among variables depend most heavily on the rank order of cases along each measure and much less on the relative distances between the cases. The transformations do not alter the rank order. When looking for *relationships* among variables, the use of such transformations advances the effectiveness and validity of the search. Nonetheless some authors have criticized the use of such data transformations on the grounds that "defendants and plaintiffs live in a world of real dollars, not of log dollars." Sunstein et al., *supra* note 51, at 2077 n.21.

80. A log transformation was not used because it somewhat overcorrected the skewness in jurors' awards and substantially overcorrected the judges' and lawyers' awards.

3. *Sociodemographic Characteristics*

For all members of the three respondent groups, we recorded their gender, state, and the rural/urban status of their locale (for jurors, where they lived; for judges, where their court's jurors lived; and for lawyers, where plaintiffs in their cases lived). In addition, all groups were asked whether they knew of someone who had the condition presented in either of the scenarios they were given. Jurors also were asked their age as well as their education and income level. Judges and lawyers were asked what percentage of their caseload was comprised of personal injury cases and for how many years they had been hearing or handling personal injury cases. The lawyers also were asked whether and how frequently they had served as an arbitrator in personal injury cases.

F. *Analyses*

1. *Overview*

Two sets of analyses were conducted to examine the factors that predict respondents' (a) perceptions of the overall *severity* of the injury, and (b) *awards* for general damages. For each set of analyses, two sets of factors were used as predictors: the respondents' subjective assessments of the injuries on the dimensions of pain, disability, disfigurement, and mental suffering; and the respondents' sociodemographic characteristics. All measures were standardized in order to provide coefficient estimates that could be used to interpret the relative contribution of the predictors within each model and to facilitate the screening of interactions. The analyses were conducted separately for jurors, judges, plaintiffs' lawyers, and defense lawyers.

Each set of analyses consisted of the following. To accommodate both the within- and between-subjects aspects of the design, a series of mixed linear model analyses was conducted to assess the unique contribution of each predictor. First, all predictors were entered simultaneously in a single analysis. Next, interactions among the predictors were tested. Finally, nonsignificant terms were sequentially eliminated from the model until a final model was obtained that consisted of significant interactions and main effects that either were significant or were a component of a significant interaction. To assess the collective predictive strength of the items, multiple regression analyses were conducted on the full and final models. In addition, the zero-order correlation of each predictor with awards and overall severity was examined to assess each predictor's total contribution (i.e., both unique and shared variance).⁸¹ Comparisons of the structure and relative

81. See Appendix C for a brief summary of some concepts of regression analysis.

strength of the models were made across the four respondent groups and between the award and severity models.

2. *Order Effects*

The order in which the injuries in each pair were presented was counterbalanced. Nonetheless, we examined whether the order of presentation affected respondents' awards and ratings. The principal effect of order was that jurors' awards were higher for the second injury they heard than for the first one.⁸² Jurors' ratings of the overall severity of the injury were somewhat higher for the first injury they heard than for the second one.⁸³ For judges, plaintiffs' lawyers, and defense lawyers, the order in which the injuries were heard did not affect awards or overall severity ratings.⁸⁴ Finding an order effect for jurors but not for judges or lawyers is not surprising, given the jurors' lack of experience with determining personal injury awards.

Ideally, we would have included order as a predictor in the mixed models, but we did not so that the sets of analyses across the four respondent groups would be parallel. Order could not be included in the models for lawyers because the order in which the injuries were heard was confounded with their rural/urban location. The effect of excluding order from the jurors' analyses is to increase the error variance and, thus, to produce a more conservative estimate of the predictive strength of the models.

IV. RESULTS

The results will be presented in four Sections. The first presents the empirically derived models of the factors that account for judgments of injury severity. Separate models were developed for each group of decisionmakers: jurors, judges, plaintiffs' lawyers, and defense lawyers. Each model consists of respondents' ratings of four core attributes of injuries — physical pain, mental suffering, disability, and disfigurement — plus various demographic factors as predictors. The second Section presents models, using the same factors as predictors, to account for respondents' awards of general damages. The third Section compares the various decisionmakers' severity assessments, awards, and predictions of how much the "average juror" would award for the same injuries. The final Section demonstrates the

82. $F(2,556) = 4.44, p < .05$.

83. $F(2,554) = 2.54, p = .08$.

84. For each respondent group, order interacted significantly with several of the predictors when predicting overall severity. When predicting awards, order interacted with pain ratings for the jurors, but did not interact with any predictors for the judge or lawyer groups.

variability-reducing power of combining individual decisionmakers into groups, such as juries.

A. *Assessments of Overall Injury Severity: The Role of Injury Attributes and Sociodemographic Characteristics*

1. *Jurors' Assessments*

First, we examined the extent to which jurors' ratings of the injuries on four core attributes and jurors' sociodemographic characteristics predicted their assessments of the overall severity of the injuries. Combined in a multiple regression analysis, these factors together accounted for 72% of the variance in jurors' assessments of the injuries' overall severity. The four injury attribute ratings combined accounted for 71% of the variance. The sociodemographic characteristics together accounted for only 3% of the variance in judgments of overall severity.⁸⁵

To examine which particular factors were related to jurors' assessments of the overall severity of the injuries, jurors' injury ratings and sociodemographic characteristics were entered in a mixed linear model analysis. Both the full model and the final model are shown in Table 1A. In the full model, all items were entered simultaneously. The final model, which resulted from testing the interactions among the predictors and sequentially eliminating nonsignificant terms, included only those items that had a significant influence on judgments of overall severity or that were a component of a significant interaction.

The final model also accounted for 72% of the variance. The ratings of each of the four core injury attributes were significantly related to assessments of overall injury severity. Perceptions of disability and mental suffering had the largest effect on overall severity judgments, almost twice as large as the influence of perceptions of pain and disfigurement. Higher ratings of the injuries on each of these attributes were associated with greater perceived overall injury severity. Comparing the relative magnitude of the Betas with the relative magnitude of the bivariate correlation coefficients,⁸⁶ we see that the contributions of pain and disfigurement in the multivariate analysis were reduced

85. The variance accounted for by individual predictors, or groups of predictors, does not simply add together to equal the amount of variance accounted for by all of them combined because the components often are somewhat redundant. As here, the total variance accounted for seems to be less than the sum of its component parts. Other times, *see infra* Section IV.B.1, the combined variables can account for more than the sum of their parts.

86. *See* Table 1A. The Betas reflect only the unique contribution of each item, whereas the *r*'s reflect both unique and shared contributions combined. For an explanation of the statistical terms used in this Section, *see* Appendix C.

because of the variance in severity they shared with disability and mental suffering ratings.⁸⁷

None of the sociodemographic characteristics had a significant effect on perceptions of overall injury severity.⁸⁸ The main effects of state, income, and age were included in the final model because each was a component of a significant interaction, not because they themselves were significant predictors.

Several of the injury ratings interacted with other predictors to affect judgments of overall severity: pain and state of residence, pain and disability, pain and income, and disfigurement and age. For each of the interactions, higher ratings of the injury attribute led to significantly higher overall severity ratings at all levels of the other injury rating or sociodemographic characteristic, but the size of the effect varied depending on the level of the other factor. Specifically, perceptions of pain had a larger influence on the overall severity judgments for jurors in New York than for jurors in Illinois. In addition, pain ratings' contribution to overall severity tended to decrease as the level of disability increased and as the jurors' household income increased. Disfigurement ratings made a smaller contribution to the severity judgments of younger jurors than of other jurors. Although statistically significant, the contribution of each interaction was small relative to that of each of the four injury ratings and did not change the interpretation of the injury ratings' contribution to judgments of overall severity.

87. The inter-item correlations among the four injury ratings ranged from .49 to .67; the correlations among the sociodemographic characteristics ranged from .01 to .18, with the exception of $r = .45$ for income and education; and the correlations between the two sets of factors ranged from .01 to .13.

88. Familiarity with the injury might not have been significantly related to overall severity ratings for several reasons. First, only a relatively small portion (12%) of the jurors knew someone with the particular injury they were asked to judge, making it more difficult to detect an effect. Second, it is plausible that knowing someone who had a serious injury has the opposite effect on severity ratings as does knowing someone who had a minor injury. That is, familiarity with a minor injury might confirm its minimal impact and produce lower severity ratings than if one were not familiar with that injury. Conversely, an understanding of the truly devastating impact of major injuries may produce higher severity ratings than if one were not familiar with the injury. These countervailing responses might cancel each other out.

TABLE 1: OVERALL INJURY SEVERITY ASSESSMENTS

A. Jurors' Assessments

Variable	Bivariate <i>r</i>	All items entered simultaneously		Final model			Interactions
		Beta	<i>F</i>	Beta	<i>F</i>	df	
mental suffering	.74	.34	195.76***	.34	193.65***	(1,478)	
pain	.59	.16	57.27***	.19	53.86***	(1,478)	
disability	.75	.37	221.73***	.35	209.17***	(1,478)	
disfigurement	.61	.15	47.06***	.17	57.58***	(1,478)	
state	-.01	.05	2.27	.05	2.29		
rural/urban	.09	-.06	2.29	--	--		
gender	.03	.04	1.23	--	--		
income	-.10	.03	2.10	.03	2.21	(1,504)	
education	-.12	-.03	1.89	--	--		
age	-.03	-.02	1.83	-.03	2.96 ⁺	(1,504)	
familiar with injury	-.02	.01	.22	--	--		
			(<i>R</i> ² = .72)	-.08	5.68*	(1,478)	pain x state
				-.05	7.48**	(1,478)	pain x disability
				-.04	5.24*	(1,478)	pain x income
				.04	5.01*	(1,478)	disfigurement x age
					(<i>R</i> ² = .72)		

Note: In the final model, the standard error of the Betas ranged from .02 to .03.

*** *p* < .001

** *p* < .01

* *p* < .05

+ *p* < .10

2. Judges' Assessments

The same approach was used to model and account for judges' assessments of overall injury severity. Ratings of injuries on the four core attributes, together with the judges' sociodemographic characteristics, accounted for 70% of the variance in their assessments of the injuries' severity. The four injury attribute ratings accounted for 69% of the variance; the sociodemographic characteristics combined accounted for only 2% of the variance in judgments of overall severity.

In the final model, which accounted for 69% of the variance, ratings of each of the four core injury attributes were significantly related to assessments of overall injury severity (*see* Table 1B). Ratings of the extent of disability made the largest contribution to overall severity judgments, followed by ratings of mental suffering. Ratings of disfigurement and pain also made significant, although substantially smaller, contributions to perceptions of severity. Higher ratings of the injuries on each of these attributes were associated with greater perceived overall injury severity. Comparing the relative magnitude of the Betas with that of the bivariate correlation coefficients, we can see that the contributions of mental suffering, disfigurement, and pain ratings were reduced because of the variance in overall severity that they shared with disability ratings.⁸⁹

None of the sociodemographic characteristics made a significant contribution to perceptions of overall injury severity. However, the state in which the judges presided interacted with disability ratings to affect overall severity judgments. Specifically, disability made a larger contribution to overall severity ratings for New York judges than for Illinois judges. Because higher disability ratings led to significantly higher overall severity ratings for both New York and Illinois judges, the interaction did not alter the interpretation of disability's influence on severity judgments. The contribution of the interaction was similar in magnitude to the contribution of disfigurement and pain ratings.

89. The inter-item correlations among the four injury ratings ranged from .39 to .60; the correlations among the sociodemographic characteristics ranged from .02 to .33, with the exception of $r = .55$ for the percentage of personal injury cases in the judges' caseload and rural/urban location; and the correlations between the two sets of factors ranged from .01 to .14.

TABLE 1: OVERALL INJURY SEVERITY ASSESSMENTS (CONTINUED)

B. Judges' Assessments

Variable	Bivariate <i>r</i>	All items entered simultaneously		Final model			Interactions
		Beta	<i>F</i>	Beta	<i>F</i>	df	
mental suffering	.71	.34	88.56***	.33	87.21***	(1,206)	
pain	.46	.10	8.78**	.09	9.07**	(1,206)	
disability	.76	.48	176.33***	.53	177.71***	(1,206)	
disfigurement	.49	.08	6.52*	.10	10.16**	(1,206)	
state	.07	-.03	.33	-.01	.05	(1,235)	
rural/urban	.05	.08	1.53	--	--		
gender	.04	.05	.33	--	--		
familiar with injury	-.06	-.01	.13	--	--		
% caseload that is personal injury	.11	.02	.37	--	--		
# years hearing pi cases	.03	-.02	.63	--	--		
			($R^2 = .70$)	-.13	5.72*	(1,206)	disability x state
					($R^2 = .69$)		

Note: In the final model, the standard error of the Betas ranged from .03 to .05.

*** $p < .001$

** $p < .01$

* $p < .05$

3. *Plaintiffs' Lawyers' Assessments*

Together, the plaintiffs' lawyers' injury ratings and their sociodemographic characteristics accounted for 78% of the variance in their judgments of the injuries' overall severity. The four injury attribute ratings accounted for 78% of the variance; the sociodemographic variables together accounted for only 3% of the variance in judgments of overall severity.

In the final model, which also accounted for 78% of the variance, ratings on each of the four core injury attributes made a significant contribution to assessments of overall severity. Ratings of the extent of disability made the largest contribution to judgments of overall severity, followed by mental suffering (*see* Table 1C). Ratings of pain and disfigurement made significant but considerably smaller contributions. Higher ratings of the injury on each of these dimensions contributed to greater perceived overall injury severity. The contribution of mental suffering, pain, and disfigurement ratings in the multivariate analysis was reduced because of the variance in overall severity ratings shared with disability.⁹⁰

None of the sociodemographic characteristics made a significant contribution to perceptions of overall severity.⁹¹ Rural/urban setting interacted with ratings of mental suffering to influence judgments of overall injury severity, although it did not alter the interpretation of the relationship of mental suffering and severity. Mental suffering made a slightly greater contribution to severity ratings for plaintiffs' lawyers whose clients came mostly from rural settings than for those whose clients came mostly from urban settings. The contribution of this interaction was comparable to that of the pain and disfigurement ratings.

90. The inter-item correlations among the four injury ratings ranged from .40 to .53; the correlations among the sociodemographic characteristics ranged from .01 to .33; and the correlations between the two sets of factors ranged from .01 to .16.

91. The number of years that the plaintiffs' lawyers had been handling personal injury cases was not included in the multivariate analyses because of missing data. But it appears from the very small bivariate correlation between number of years in personal injury practice and overall severity ratings, $r(147) = -.01$, that no relationship exists between these variables.

TABLE 1: OVERALL INJURY SEVERITY ASSESSMENTS (CONTINUED)

C. Plaintiffs' Lawyers' Assessments

Variable	Bivariate <i>r</i>	All items entered simultaneously		Final model			Interactions
		Beta	<i>F</i>	Beta	<i>F</i>	df	
mental suffering	.80	.40	67.74***	.32	70.15***	(1,106)	
pain	.57	.11	7.28**	.12	8.85**	(1,106)	
disability	.80	.44	86.45***	.45	99.72***	(1,106)	
disfigurement	.52	.10	6.47*	.11	7.74**	(1,106)	
state	-.04	-.03	.23	--	--		
rural/urban	.08	.01	.02	.01	.01	(1,121)	
gender	.07	-.07	.46	--	--		
familiar with injury	-.07	.05	1.88	--	--		
% caseload that is personal injury	.05	-.01	.11	--	--		
how often serve as pi arbitrator	-.02	-.01	.01	--	--		
			($R^2 = .78$)	.13	4.07*	(1,106)	mental suffering x rural/urban
					($R^2 = .78$)		

Note: In the final model, the standard error of the Betas ranged from .04 to .06.

*** $p < .001$

** $p < .01$

* $p < .05$

4. *Defense Lawyers' Assessments*

Defense lawyers' injury attribute ratings and their sociodemographic characteristics together accounted for 75% of the variance in their assessments of the injuries' overall severity. The four injury ratings accounted for 75% of the variance; the sociodemographic characteristics together accounted for only 1% of the variance.

The final model accounted for 75% of the variance. Perceptions of mental suffering made the largest contribution to assessments of overall severity in the final model (*see* Table 1D), followed by disability ratings. Ratings of disfigurement made a significant though substantially smaller contribution to overall severity. Pain ratings made no significant independent contribution to overall severity judgments, but instead influenced overall severity indirectly through pain's relationship with the other injury attributes. Comparing the relative magnitude of the Betas with the relative magnitude of the bivariate correlation coefficients, we see that the effect of pain, and to a lesser extent disfigurement, in the multivariate analysis was reduced because of shared variance.⁹²

None of the sociodemographic variables made a significant contribution to perceptions of overall injury severity,⁹³ and none of the predictors interacted significantly to influence severity ratings.

92. The inter-item correlations among the four injury ratings ranged from .31 to .46, with the exception of $r = .72$ for disability and mental suffering; the correlations among the sociodemographic characteristics ranged from .01 to .21; and the correlations between the two sets of factors ranged from .01 to .13.

93. The number of years that the defense lawyers had been handling personal injury cases was not included in the multivariate analyses because of missing data. But it appears from the very small bivariate correlation between number of years in personal injury practice and overall severity ratings, $r(165) = .04$, that no relationship exists between these variables.

TABLE 1: OVERALL INJURY SEVERITY ASSESSMENTS (CONTINUED)

D. *Defense Lawyers' Assessments*

Variable	Bivariate <i>r</i>	All items entered simultaneously		Final model			
		Beta	<i>F</i>	Beta	<i>F</i>	df	Interactions
mental suffering	.79	.48	101.66***	.47	105.19***	(1,110)	
pain	.46	.05	1.77	--	--		
disability	.79	.37	56.01***	.39	69.54***	(1,110)	
disfigurement	.47	.10	8.04**	.11	10.36**	(1,110)	
state	.02	-.05	.57	--	--		
rural/urban	.03	.05	.42	--	--		
gender	-.06	.08	.31	--	--		
familiar with injury	-.07	.01	.01	--	--		
% caseload that is personal injury	-.01	-.04	.88	--	--		
how often serve as pi arbitrator	-.09	.04	1.13	--	--		
			(<i>R</i> ² = .75)		(<i>R</i> ² = .75)		

Note: In the final model, the standard error of the Betas ranged from .03 to .05.

*** *p* < .001

** *p* < .01

5. Comparison of Injury Severity Models Across the Four Groups

By comparing the final models for each group of decisionmakers, we can learn about the factors, the weights of those factors, and the predictability of the judgments of overall severity made by jurors, judges, and lawyers. The similarities among these diverse decisionmakers are far more striking than their differences. The judgments of all four groups were highly predictable, and the factors that had great influence — as well as those that had little or no influence — were very similar among all groups.

The predictability of severity judgments, as measured by the total proportion of variance accounted for by each of the final models, was large and similar for each of the groups: 72% for jurors, 69% for judges, 78% for plaintiffs' lawyers, and 75% for defense lawyers. For all groups, variation in severity ratings was to a far greater extent the product of differences in the injuries being responded to than it was the product of individual differences among the respondents. The variation of responses of different people to the *same* injury was small and similar across the four groups.⁹⁴ That is, members of each group responded similarly in assessing the severity of the injuries involved in a given case. And, as one should expect, the variation in responses of the same person to *different* injuries was many times larger than that of different people to the same injury.⁹⁵

For all four groups, the only main effects that remained in the final models as significant predictors of overall severity were the core injury attributes — none of the socio-demographic characteristics made a significant contribution. For jurors, judges, and plaintiffs' lawyers, disability made the largest contribution, followed by mental suffering; pain and disfigurement made significant but smaller contributions. The model for defense lawyers differed from the other three in that mental suffering was the most powerful influence while disability was second,⁹⁶ and pain did not make a significant unique contribution to overall severity. Interactions involving sociodemographic characteristics played only a small role in the models: the magnitude of their influence was fairly small, and none changed the interpretation of the injury ratings' influence on overall severity assessments.

94. The between-subjects covariance parameter estimate was .006 for jurors, -.008 for judges, -.003 for plaintiffs' lawyers, and .06 for defense lawyers.

95. The within-subjects covariance parameter estimate was .27 for jurors, .31 for judges, .23 for plaintiffs' lawyers, and .17 for defense lawyers.

96. Shared variance could explain the reduced impact of disability for defense lawyers compared to the other groups. For them, the correlation between disability and mental suffering was $r = .72$, while it was smaller for the other groups.

B. *Predicting General Damages Awards: Translating Injury Judgments into Dollars*

The same predictive data and the same modeling techniques as above were used to account for the general damages awards. In contrast to the high degree of predictability of perceptions of injury severity, models for general damages awards show less predictability and greater decisionmaking complexity. Even so, across the groups of decisionmakers, a considerable degree of similarity remains in the patterns of variables that influence the damages awards.

1. *Jurors' Awards*

We examined the extent to which jurors' perceptions of the injuries on the four core attributes⁹⁷ and their sociodemographic characteristics predicted awards (which had been square-root transformed). Combined in a multiple regression analysis, these factors accounted for 21% of the variance in awards. Ratings of the four core injury attributes accounted for 17% of the variance; the sociodemographic characteristics accounted for only 2% of the variance in awards.

The final model accounted for 23% of the variance in awards for general damages. Of the four injury ratings, perceptions of disability had the greatest impact on awards, more than twice as large as ratings of either mental suffering or disfigurement (see Table 2A). Higher ratings of the injury on each of these three attributes were associated with larger awards. Pain ratings did not have a significant unique effect on awards when combined with the other predictors. Given the bivariate correlation between pain ratings and awards, pain apparently affected awards through its relationship with the other injury ratings. The influence of pain ratings, and to a lesser extent disfigurement and mental suffering, was reduced in the multivariate analysis due to the variance in awards shared with disability.

Of the sociodemographic characteristics, gender was the most strongly related to awards. In fact, the magnitude of gender's influence on awards was second only to that of disability. Income also had a significant but smaller influence on awards. Male jurors gave larger awards than female jurors, and jurors with higher household incomes gave larger awards than those with lower incomes.

Several of the predictors interacted to affect awards in a manner that altered the interpretation of the injury attributes' influence on awards. Examining the interaction of disfigurement ratings with rural/urban residence showed that as perceptions of disfigurement in-

97. Because the four specific injury ratings were highly correlated with overall severity ratings, overall injury severity was not included as a predictor of awards in order to avoid multicollinearity.

creased, awards increased for urban jurors but not for rural jurors. The nature of the interaction of pain ratings and education was such that as pain ratings increased, awards increased for those jurors who were college graduates but not for jurors with other levels of education. Finally, the interaction of disfigurement and disability ratings showed that as the perception of disfigurement increased, awards increased only at the higher levels of disability, while higher disability ratings were related to larger awards at all levels of disfigurement. The influence of each of these interactions was similar in magnitude to the impact of mental suffering and disfigurement ratings and household income.

In sum, jurors gave larger awards when they saw the injuries as involving a greater degree of disability and mental suffering and when the jurors were male or had a higher household income. Greater perceived disfigurement led to giving larger awards only when it was associated with higher levels of disability and only for urban jurors. Greater perceived pain led to giving larger awards only for college graduates.

TABLE 2: PAIN AND SUFFERING AWARDS (SQUARE-ROOT TRANSFORMED)

A. Jurors' Awards

Variable	Bivariate <i>r</i>	All items entered simultaneously		Final model			Interactions
		Beta	<i>F</i>	Beta	<i>F</i>	df	
mental suffering	.34	.13	10.61**	.13	10.69**	(1,481)	
pain	.24	.02	.24	.02	.38	(1,481)	
disability	.38	.28	46.54***	.31	57.06***	(1,481)	
disfigurement	.32	.12	10.86**	.15	4.59*	(1,481)	
state	.05	-.05	.61	--	--		
rural/urban	.08	-.05	.47	-.05	.45	(1,505)	
gender	-.12	.26	15.79***	.26	15.74***	(1,505)	
income	.07	.11	8.25**	.11	8.41**	(1,505)	
education	.07	.05	2.21*	.06	2.63	(1,505)	
age	.02	.02	.40	--	--		
familiar with injury	-.01	.02	.44	--	--		
			(<i>R</i> ² = .21)	-.13	5.28*	(1,481)	disfigurement x rural/urban
				.10	12.88***	(1,481)	pain x education
				.08	7.23***	(1,481)	disfigurement x disability
					(<i>R</i> ² = .23)		

Note: In the final model, the standard error of the Betas ranged from .03 to .07.

*** *p* < .001

** *p* < .01

* *p* < .05

2. Judges' Awards

Judges' perceptions of the injuries and their sociodemographic characteristics accounted for 44% of the variance in square-root awards. The four injury attributes accounted for 40% of the variance; the sociodemographic variables accounted for only 4% of the variance in awards.

The final model accounted for 42% of the variance in judges' awards for general damages. Of the four injury ratings, disability had the greatest influence on awards, almost twice as large as that of mental suffering (*see* Table 2B). Disfigurement ratings had a smaller, but significant, impact on awards. Higher ratings of the injury on each of these three dimensions were associated with larger awards. Perceptions of pain did not have a significant unique influence on awards when combined with the other predictors, although its bivariate correlation with awards was moderate. Pain apparently does not affect awards independently but does so through its relationship with the other ratings. The effect of perceptions of pain, and to a lesser extent perceptions of mental suffering and disfigurement, was reduced because of shared variance.

Of the sociodemographic characteristics, only the state in which the judges presided affected their awards. New York judges gave larger general damages awards than did Illinois judges. Additionally, the state in which the judges presided interacted with disability ratings to affect awards. Specifically, disability had a greater impact on awards for New York judges than for Illinois judges, although higher disability ratings led to significantly larger awards in both states. The influences of the state and of the interaction were similar in magnitude to the effect of perceptions of mental suffering.

In sum, judges gave larger awards when they saw the injuries as involving a greater degree of disability, mental suffering, and disfigurement, and when the judges presided in New York courts rather than in Illinois courts. Perceived disability had a greater effect on awards for New York judges than for Illinois judges.

TABLE 2: PAIN AND SUFFERING AWARDS (CONTINUED)

B. Judges' Awards

Variable	Bivariate <i>r</i>	All items entered simultaneously		Final model			Interactions
		Beta	<i>F</i>	Beta	<i>F</i>	df	
mental suffering	.53	.24	26.41***	.25	27.85***	(1,212)	
pain	.36	.04	.75	--	--		
disability	.58	.34	51.95***	.49	68.06***	(1,212)	
disfigurement	.38	.13	8.66**	.12	8.02**	(1,212)	
state	.14	-.22	8.42**	-.23	10.40**	(1,236)	
rural/urban	.12	-.12	1.69	--	--		
gender	.02	.14	1.41	--	--		
familiar with injury	-.04	-.03	.49	--	--		
% caseload that is personal injury	.11	.01	.07	--	--		
# years hearing pi cases	-.01	-.04	1.30	--	--		
			($R^2 = .44$)	-.21	9.56**	(1,212)	disability x state
					($R^2 = .42$)		

Note: In the final model, the standard error of the Betas ranged from .04 to .07.

*** $p < .001$

** $p < .01$

3. *Plaintiffs' Lawyers' Awards*

Plaintiffs' lawyers' perceptions of the injuries and their sociodemographic characteristics accounted for 37% of the variance in their square-root awards. The four injury attributes accounted for 34% of the variance; the set of sociodemographic characteristics accounted for only 1% of the variance in awards.

The final model accounted for 48% of the variance in awards for general damages. Of the four injury attributes, disability had the greatest impact on awards in the final model, followed by mental suffering (*see* Table 2C). Higher ratings of the injury along these dimensions were associated with larger awards. Perceptions of pain and disfigurement did not have a significant effect on awards in the multivariate analysis, despite their moderate bivariate correlations with awards. Thus, pain and disfigurement did not affect awards independently but through their relationship with the other ratings.

Of the sociodemographic characteristics, only the state in which plaintiffs' lawyers practiced had even a marginally significant influence on awards. Specifically, New York plaintiffs' lawyers tended to award more than did Illinois plaintiffs' lawyers. None of the other demographic characteristics were associated with awards.⁹⁸

Several of the predictors interacted, having a substantial effect on awards and altering the interpretation of the effect of injury attributes and demographic characteristics on awards. First, the state in which the lawyers practiced interacted with how often they had served as an arbitrator in personal injury cases. New York plaintiffs' lawyers who had frequently served as arbitrators gave larger awards than those who had served less often, while the pattern was reversed for Illinois plaintiffs' lawyers. Second, the interaction of familiarity with the type of injury presented and rural/urban practice indicated that familiarity with the injury led to significantly larger awards for urban plaintiffs' lawyers but had no effect on rural lawyers' awards. Third, the interaction of disfigurement and disability ratings showed that perceptions of greater disfigurement were significantly related to larger awards only at the highest level of disability, while higher disability ratings were related to larger awards at all levels of disfigurement. Fourth, disability perceptions and state interacted such that although higher disability ratings were related to larger awards for plaintiffs' lawyers in both states, the magnitude of the relationship between disability ratings and awards was almost twice as large in New York as in Illinois. Finally,

98. The number of years that the plaintiffs' lawyers had been handling personal injury cases was not included in the multivariate analyses because of missing data, but it appears from the very small bivariate correlation between number of years in personal injury practice and awards, $r(148) = .05$, that no relationship exists between these variables.

perceived disability interacted with the frequency with which the lawyers had served as arbitrators, but no systematic, interpretable pattern appeared.

In sum, plaintiffs' lawyers gave larger awards when they saw the injuries as involving a greater degree of disability and mental suffering. If they practiced in New York rather than in Illinois, they were marginally more likely to give larger awards. Moreover, five different interaction effects emerged, consisting of complex relationships involving either the perception of disability or the location (state as well as urban/rural) of their law practice.

TABLE 2: PAIN AND SUFFERING AWARDS (CONTINUED)

C. *Plaintiffs' Lawyers' Awards*

Variable	Bivariate <i>r</i>	All items entered simultaneously		Final model			Interactions
		Beta	<i>F</i>	Beta	<i>F</i>	df	
mental suffering	.50	.26	10.94**	.28	17.16***	(1,106)	
pain	.35	-.03	.18	--	--		
disability	.53	.28	13.93***	.46	27.28***	(1,106)	
disfigurement	.40	.11	2.81 ⁺	.05	.63	(1,106)	
state	.02	-.18	3.18 ⁺	-.19	3.89 ⁺	(1,118)	
rural/urban	.05	.05	.17	.06	.42	(1,118)	
gender	-.05	.20	1.73	--	--		
familiar with injury	-.01	.09	2.8 ⁺	.20	2.46	(1,118)	
% caseload that is personal injury	.02	-.02	.17	--	--		
how often serve as pi arbitrator	.06	.06	1.23	.30	1.97	(1,118)	
			(<i>R</i> ² = .37)	-.44	17.03***	(1,118)	how often pi arbitrator x state
				-.25	6.88*	(1,106)	familiar with injury x rural/urban
				.23	17.05***	(1,106)	disability x disfigurement
				-.21	4.97*	(1,106)	disability x state
				.14	8.60**	(1,106)	disability x how often pi arbitrator
					(<i>R</i> ² = .48)		

Note: In the final model, the standard error of the Betas ranged from .05 to .11.

*** *p* < .001

** *p* < .01

* *p* < .05

+ *p* < .10

4. *Defense Lawyers' Awards*

Defense lawyers' perceptions of the injuries and their sociodemographic characteristics accounted for 52% of the variance in their square-root awards. The four specific injury ratings combined accounted for 50% of the variance in awards; the sociodemographic characteristics together accounted for 4% of the variance in awards.

The final model accounted for 58% of the variance in general damages awards. Perceptions of mental suffering had the greatest impact on awards (*see* Table 2D). Disability and disfigurement ratings also had a significant effect on awards. Higher ratings of the injury on each of these dimensions were associated with larger awards. Pain ratings had no significant independent effect on awards, despite a moderate bivariate correlation. Thus, pain influenced awards indirectly through its relationship with the other injury ratings. The influences of disfigurement and disability also were somewhat reduced because of the variance in awards they shared with the other injury ratings.

The state in which the defense lawyers practiced was the only sociodemographic characteristic to have a significant effect on awards.⁹⁹ Defense lawyers practicing in New York awarded more than those in Illinois. The magnitude of the impact of state on awards was comparable to that of disability ratings.

Several of the predictors interacted to affect awards. The significant interaction of mental suffering and perceived disability had an effect on awards comparable in magnitude to that of disability ratings and the state in which the lawyers practiced. The nature of this interaction was such that greater perceived disability resulted in larger awards only at the highest level of mental suffering, while greater perceived mental suffering led to larger awards at each level of disability, with an increasing impact as the level of disability increased. The interaction of pain and disfigurement ratings showed that greater perceived pain led to larger awards only at the highest level of disfigurement, and greater perceived disfigurement led to larger awards only at moderate and severe levels of pain. Finally, the interaction of perceived disfigurement with the proportion of the defense lawyers' caseload that was personal injury work showed that higher disfigurement ratings were significantly related to larger awards only for those defense lawyers for whom personal injury cases comprised half or less

99. The number of years that the defense lawyers had been handling personal injury cases was not included in the multivariate analyses because of missing data, but it appears from the very small bivariate correlation between number of years in personal injury practice and awards, $r(166) = -.08$, that no relationship exists between these variables.

of their caseload. The influence of these latter two interactions on awards was comparable in magnitude to that of disfigurement ratings.

In sum, defense lawyers gave larger awards when they saw the injuries as involving a greater degree of mental suffering, disability, and disfigurement and if they practiced in New York rather than in Illinois. The influence of disability, disfigurement, and pain was qualified by three significant interactions with other predictors.

TABLE 2: PAIN AND SUFFERING AWARDS (CONTINUED)

D. *Defense Lawyers' Awards*

Variable	Bivariate <i>r</i>	All items entered simultaneously		Final model			Interactions
		Beta	<i>F</i>	Beta	<i>F</i>	df	
mental suffering	.66	.38	31.21***	.40	39.32***	(1,104)	
pain	.37	.03	.23	.03	.46	(1,104)	
disability	.60	.23	10.26**	.23	11.77***	(1,104)	
disfigurement	.47	.18	11.85***	.14	8.53**	(1,104)	
state	.10	-.24	7.54**	-.24	8.28**	(1,120)	
rural/urban	.03	-.02	.04	--	--		
gender	-.01	-.15	.59	--	--		
familiar with injury	-.10	-.01	.01	--	--		
% caseload that is personal injury	-.06	-.05	1.00	-.06	1.82	(1,120)	
how often serve as pi arbitrator	-.14	-.03	.40	--	--		
			(<i>R</i> ² = .52)	.20	15.19***	(1,104)	mental suffering x disability
				.13	8.58**	(1,104)	pain x disfigurement
				-.12	8.18**	(1,104)	disfigurement x % caseload pi
					(<i>R</i> ² = .58)		

Note: In the final model, the standard error of the Betas ranged from .04 to .08.

*** *p* < .001

** *p* < .01

5. Comparison of Award Models Across the Four Groups

Earlier, we had seen that the regression models predicting overall injury severity ratings for all four groups of decisionmakers showed the groups to be remarkably similar in nearly every respect. The regression models in this Section aim to account for variation in the general damages awards. With these models, we see a decline in predictability for all four groups and greater divergence among them. At the same time, important parallels continue to exist.

The final models for each group accounted for a moderate proportion of the variance in awards: 23% for jurors, 42% for judges, 48% for plaintiffs' lawyers, and 58% for defense lawyers. Although it is more difficult to predict the awards made by all of these decisionmakers, the translation from perceptions of injury severity to damages awards has led to the greatest decline in predictability for the model of jurors' decisions. Of course, this may be as much a statement about the limitations of the predictive model to capture what jurors are doing as it is a statement about the jurors' decisions.

Nonetheless, confirming the preceding impressions, the variation of responses of different people to the *same* injury was greater for jurors than for the other groups.¹⁰⁰ These patterns of predictability and intra-injury variability in awards are not surprising, given that jurors have essentially no experience assigning a dollar value to injuries while the other groups do. In addition, had we been able to include an order term in the jurors' model to reflect the fact that there was a significant order effect for jurors (which was not found for the other groups), the error variance for jurors likely would have been reduced.¹⁰¹

For all groups, the variation in responses of the same person to *different* injuries was much larger than that of different decisionmakers to the same injury.¹⁰² The somewhat lower variability in awards for defense lawyers when the same person was reacting to different injuries suggests they may be less responsive to case details. Thus, the greater predictability of awards for defense lawyers does not necessarily mean they are "better" at this task, but instead that their responses might be less nuanced or more mechanical.

For all groups, some of the variability in awards might be due to the fact that, in order to avoid potentially confounding contextual factors, the case summaries did not describe the plaintiffs or their pre-

100. The between-subjects covariance parameter estimate was .21 for jurors, compared to .06 for judges, .01 for plaintiffs' lawyers, and .02 for defense lawyers

101. See *supra* note 82 and accompanying text.

102. The within-subjects covariance parameter estimate was .56 for jurors, .47 for judges, .48 for plaintiffs' lawyers, and .35 for defense lawyers.

injury lifestyles. Depending on the extent of changes in a plaintiff's work, family, and recreational life, the same injury could be seen to have a different impact on different plaintiffs and, thus, merit different awards. Numerous judges and lawyers requested information about the plaintiff's gender, occupation, lifestyle, family, and other background details.¹⁰³ For instance, some wanted to know the plaintiff's gender because they felt a scar on a female plaintiff's face would warrant more compensation than the same injury for a male plaintiff. Thus, award variability might be greater because respondents brought different assumptions about the plaintiff to bear on the task.¹⁰⁴

The respondent groups showed important similarities in the role that injury attributes played in their awards, but also some differences. For all four groups, pain ratings had no significant unique influence on awards. For jurors, judges, and plaintiffs' lawyers, perceptions of disability had the greatest impact on awards — about twice as large as the impact of the second strongest predictor — while for defense lawyers, perceptions of mental suffering had the largest effect on awards. Mental suffering had a significant effect on awards for all four groups. Perceived disfigurement had a significant impact on awards for all groups except plaintiffs' lawyers. For all groups except judges, the interpretation of the effect of some of the injury attributes was qualified by significant interactions with the sociodemographic characteristics.

In all instances, higher ratings on the injury attributes were associated with larger awards. This suggests considerable vertical equity in awards — that all groups gave larger awards to plaintiffs with injuries involving a greater degree of disability, mental suffering, and (except for plaintiffs' lawyers) disfigurement. Additional data provide evidence of considerable vertical equity for all groups. For each injury, the mean award and mean assessment of overall injury severity were calculated within each respondent group. These items then were correlated, using the injury scenario as the unit of analysis ($n = 62$). The correlation between overall severity and awards was .83 for jurors, .88

103. Recall that, if they inquired, respondents were told that the plaintiff was of average good health and had an average lifestyle before the injury, which seemed to be sufficient information for most of the respondents.

104. It is interesting to note that, during the interviews, lawyers and judges requested information that, legally, should have no relevance for awards for general damages, but which revealed their own search for shortcuts to estimating general damages. For instance, they wanted to know the amount of the plaintiff's medical expenses and lost wages so that they could use a formula to calculate awards for general damages. That judges and lawyers sometimes grasped at such shortcuts suggests that they did not have much better bearings than jurors in the chartless seas of determining general damages awards. They also wanted to know the effect of the injury on employment status. Relatedly, Vidmar & Rice, *supra* note 18, at 896-97, found that experienced lawyers serving as arbitrators tended to make more improper use of information than lay jurors had, such as using evidence of liability to fix damages and increasing the size of awards to cover attorneys' fees.

for judges, .77 for plaintiffs' lawyers, and .88 for defense lawyers.¹⁰⁵ Thus, all groups gave larger awards to plaintiffs with injuries seen as more severe and smaller awards to plaintiffs with injuries viewed as less severe.

One or more demographic variables had a significant and large effect on the awards for each of the four groups. Judges and lawyers in New York gave larger awards than did their counterparts in Illinois.¹⁰⁶ Comparing the moderate multivariate influence of state on awards with its smaller bivariate relationship in each of these groups suggests that state suppressed irrelevant variance in the other variables in the equation. This may suggest that a constellation of subtle differences exists between New York and Illinois judges and lawyers, which in the multivariate analysis is captured by the single variable of state. For jurors, the state in which they lived was not related to awards, but gender and household income were, with male jurors and jurors with higher incomes giving larger awards.¹⁰⁷

C. *Comparing Jurors', Judges', and Lawyers' Injury Severity Assessments, Awards, and Predictions of the "Average Juror's" Awards*

Next, we compare the general tendencies of jurors, judges, plaintiffs' lawyers, and defense lawyers, as groups, to perceive injuries as more or less serious and to give awards that are relatively higher or lower. In addition, we report a number of analyses relating to respondents' estimates of what the "average juror" would award.

105. Calculating these correlations without first averaging across people, that is, using the respondent instead of the injury as the unit of analysis, produces these relationships: .38 for jurors, .65 for judges, .61 for plaintiffs' lawyers, and .71 for defense lawyers. The lower correlation for jurors no doubt reflects the greater variability noted above.

106. Larger awards by New York judges and lawyers may reflect the fact that New York instructions explicitly include compensation for "loss of enjoyment of life," while the Illinois instructions do not specify that this component should be considered in determining general damages. Compare NEW YORK PATTERN JURY INSTRUCTIONS - CIVIL § 2:280.1 (2d ed. 1994) with ILLINOIS PATTERN JURY INSTRUCTIONS - CIVIL, §§ 30.00 - 30.05 (3d ed. 1994). Cutting the other way, New York instructions direct that compensation be given only for *conscious* pain and suffering, while Illinois instructions do not address this issue. Compare NEW YORK PATTERN JURY INSTRUCTIONS - CIVIL § 2:280 cmt. (2d ed. 1994) with ILLINOIS PATTERN JURY INSTRUCTIONS - CIVIL, § 30.05 (3d ed. 1994). Although this difference would tend to produce smaller awards in New York, it was unlikely to have much effect in the present study because the plaintiff was not conscious of her or his condition in only two of our 62 injuries.

107. Gender might not have been related to awards for judges or lawyers because only a relatively small portion of each group was female (13% of judges, 13% of plaintiffs' lawyers, and 7% of defense lawyers), making it more difficult to detect an effect. Or perhaps because they have learned to see damages in similar ways regardless of their gender.

1. Comparison of Overall Severity Assessments

Do the four different groups of decisionmakers, in the two different states, vary in their perceptions of the severity of the injuries, so that some see the same injury as more severe and others see it as less severe? To answer this question, for each respondent group, the mean overall severity rating was calculated for each injury, separately within each state. Table 3 presents these data, averaged across the 62 injuries. A repeated-measures analysis of variance ("ANOVA") was then conducted to compare the severity judgments across the groups, using injury as the unit of analysis.

TABLE 3: RATINGS OF OVERALL INJURY SEVERITY

	<i>Illinois</i>	<i>New York</i>
Jurors	3.64	3.63
Judges	3.31	3.46
Plaintiffs' Lawyers	3.52	3.42
Defense Lawyers	3.12	3.17

Note: Each cell represents a mean taken over all 62 injuries.

Ratings were made on a five-point scale, with 5 designated as "extremely severe."

In each state, the four respondent groups differed in their assessments of overall injury severity.¹⁰⁸ For Illinois respondents, follow-up analyses showed that jurors rated the injuries as more severe than did judges¹⁰⁹ and defense lawyers,¹¹⁰ but did not differ from plaintiffs' lawyers. Plaintiffs' lawyers rated the injuries as more severe than did judges.¹¹¹ Defense lawyers rated the injuries as less severe than did all other groups.¹¹² For New York respondents, follow-up analyses showed that jurors rated the injuries as more severe than all other groups,¹¹³ and defense lawyers rated the injuries as less severe than all other groups.¹¹⁴ Plaintiffs' lawyers' severity ratings did not differ significantly from judges' ratings. Thus, jurors tended to rate the injuries as more severe than the other groups, and defense lawyers rated them as less severe than the other groups.

108. Illinois: $F(3,183) = 14.78, p < .001$; New York: $F(3,183) = 11.70, p < .001$.

109. $F(1,61) = 27.75, p < .001$.

110. $F(1,61) = 34.02, p < .001$.

111. $F(1,61) = 6.33, p < .05$.

112. Versus jurors: $F(1,61) = 34.02, p < .001$. Versus judges: $F(1,61) = 5.12, p < .05$. Versus plaintiffs' lawyers: $F(1,61) = 16.66, p < .001$.

113. Versus judges: $F(1,61) = 7.47, p < .01$. Versus plaintiffs' lawyers: $F(1,61) = 7.13, p < .05$. Versus defense lawyers: $F(1,61) = 30.88, p < .001$.

114. Versus judges: $F(1,61) = 15.63, p < .001$. Versus plaintiffs' lawyers: $F(1,61) = 6.88, p < .05$. Versus jurors: $F(1,61) = 30.88, p < .001$.

To what extent did the different groups agree on which injuries were more severe and which less? Correlations were computed using the mean overall severity ratings calculated for each group, separately for each injury. Jurors' overall severity assessments were highly correlated with those of the other groups. In Illinois, the correlation of jurors' severity ratings with those of judges was $r = .84$; with those of plaintiffs' lawyers, $r = .72$; and with those of defense lawyers, $r = .74$.¹¹⁵ In New York, the correlation of jurors' severity ratings with those of judges was $r = .90$; with those of plaintiffs' lawyers, $r = .82$; and with those of defense lawyers, $r = .81$.¹¹⁶ The intercorrelations of overall severity ratings among judges, plaintiffs' lawyers, and defense lawyers ranged from .69 to .78 in Illinois and from .76 to .85 in New York. Thus, injuries seen as more severe by one group tended strongly to be seen as more severe by the other groups.

2. Comparison of Awards

Do the four different groups of decisionmakers, in the two different states, vary in the damages awards they offer, such that some award more and others less for the same injury? To answer this question, for each respondent group, the mean square-root award was calculated for each injury, separately within each state. Table 4 presents these data, averaged across the 62 injuries. A repeated-measures ANOVA was conducted to compare the awards across the groups, with injury as the unit of analysis.

TABLE 4: GENERAL DAMAGES AWARDS

	<i>Illinois</i>	<i>New York</i>
Jurors	665	746
Judges	567	716
Plaintiffs' Lawyers	780	817
Defense Lawyers	561	659

Note: Each cell represents a mean taken over all 62 injuries.
The damage awards are square-root transformations of dollar distributions.

In Illinois, awards differed by respondent group.¹¹⁷ Follow-up analyses showed that jurors' awards did not differ from those of plaintiffs' lawyers, but both gave awards that were larger than those of

115. For each of the preceding correlations, $n = 62$ injuries and the significance level was $p < .001$.

116. For each of the immediately preceding correlations, $n = 62$ injuries and the significance level was $p < .001$.

117. $F(3,183) = 7.55, p < .001$.

judges¹¹⁸ and defense lawyers.¹¹⁹ Defense lawyers' awards did not differ significantly from judges' awards. In New York, awards differed marginally by respondent group.¹²⁰ Follow-up analyses showed that jurors' awards were marginally larger than those of defense lawyers¹²¹ and did not differ from those of judges and plaintiffs' lawyers. Plaintiffs' lawyers' awards were significantly larger than those of defense lawyers¹²² and marginally larger than those of judges.¹²³ Defense lawyers' awards were marginally smaller than judges' awards.¹²⁴ Thus, in both states, plaintiffs' lawyers and jurors tended to give the largest awards, while defense lawyers tended to give the smallest awards. Given that jurors tended to rate the injuries as more severe than the other groups (and the defense lawyers rated them as least severe), and that the relationship between severity assessments and awards is strong, it would not be unreasonable for jurors' awards to be commensurately larger (and defense lawyers' awards smaller).

To what extent did the different groups agree on which injuries deserved higher general damages awards and which were entitled to lower awards? Correlations were computed using the mean square-root award for each injury within each group, using injury as the unit of analysis. Jurors' awards were highly correlated with those of the other groups. In Illinois, the correlation of jurors' awards with judges' awards was $r = .73$; with plaintiffs' lawyers' awards, $r = .63$; and with defense lawyers' awards, $r = .71$.¹²⁵ In New York, the correlation of jurors' awards with judges' awards was $r = .69$; with plaintiffs' lawyers' awards, $r = .60$; and with defense lawyers' awards, $r = .66$.¹²⁶ The inter-correlations among the judges' and lawyers' awards ranged from .73 to .83 in Illinois and from .66 to .83 in New York. Thus, injuries that were given larger awards by one group also tended to be given larger awards by the other groups.

118. Versus jurors: $F(1,61) = 6.59, p < .05$. Versus plaintiffs' lawyers: $F(1,61) = 11.01, p < .01$.

119. Versus jurors: $F(1,61) = 6.64, p < .05$. Versus plaintiffs' lawyers, $F(1,61) = 12.69, p < .01$.

120. $F(3,183) = 2.56, p = .06$.

121. $F(1,61) = 2.88, p = .095$.

122. $F(1,61) = 5.14, p < .05$.

123. $F(1,61) = 2.97, p = .09$.

124. $F(1,61) = 2.88, p = .095$.

125. For each of the preceding correlations, $n = 62$ injuries and the significance level was $p < .001$.

126. For each of the immediately preceding correlations, $n = 62$ injuries and the significance level was $p < .001$.

3. *Comparison of Each Group's Award with Its Prediction of the "Average Juror's" Award*

Do the four different groups of decisionmakers, in the two different states, vary in their estimates of how much the "average juror" would award in general damages? Do some groups expect jurors to give awards that are more different from their own awards than other groups do? Are some groups' estimates of jurors' awards closer to the awards actually given by jurors in this study than other groups' estimates? Respondents in each of the four groups were asked, after indicating their own award, what they thought the "average juror" would award for the same injury. Each group's median award and median predicted "average juror" award were calculated for each injury, separately within each state.¹²⁷ Dependent-group t-tests then were conducted.

The estimated jurors' awards given by New York judges were larger than the awards they themselves gave,¹²⁸ while those estimated by Illinois judges did not differ from their own awards. Defense lawyers in Illinois¹²⁹ and New York¹³⁰ both gave estimated jurors' awards that were larger than the awards they themselves gave. Plaintiffs' lawyers' awards did not differ from their predicted jurors' awards in either state. And jurors themselves gave estimated "average juror" awards that were larger than the awards they themselves gave.¹³¹

Despite differences in the award amount predicted, all groups gave larger predicted jurors' awards to the injuries for which they themselves gave larger awards, and gave smaller predicted awards to the injuries for which they themselves gave smaller awards. In Illinois, the correlations between each group's own awards and their predicted "average juror" awards were $r = .91$ for jurors, $r = .49$ for judges, $r = .82$ for plaintiffs' lawyers, and $r = .93$ for defense lawyers.¹³² In New York, the correlations between each group's own awards and their predicted "average juror" awards were $r = .88$ for jurors, $r = .95$ for judges, $r = .96$ for plaintiffs' lawyers, and $r = .89$ for defense lawyers.¹³³ Thus, all groups expect jurors' vertical equity to be quite similar to their own.

127. Because the data for this group of analyses consisted of raw damages estimates, medians were used to reduce the influence of outliers.

128. $t(61) = 3.60, p < .01$.

129. $t(61) = 3.32, p < .01$.

130. $t(61) = 2.67, p < .05$.

131. Illinois jurors: $t(61) = 3.17, p < .01$. New York jurors: $t(61) = 3.60, p < .01$.

132. For each of the preceding correlations, $n = 62$ injuries and the significance level was $p < .001$.

133. For each of the immediately preceding correlations, $n = 62$ injuries and the significance level was $p < .001$.

How well could judges and lawyers predict what awards jurors would give? Within each state, we compared judges' and lawyers' "average juror" awards with the actual awards given by the jurors in our study. In Illinois, the awards actually given by jurors were smaller than what plaintiffs' lawyers¹³⁴ and defense lawyers¹³⁵ predicted they would be, but did not differ from the judges' predictions. In New York, the awards given by jurors were smaller than what judges¹³⁶ and defense lawyers¹³⁷ predicted they would be, but did not differ from the plaintiffs' lawyers' predictions. Generally, then, the awards jurors actually gave tended to be smaller than the awards judges and lawyers predicted they would give.

Correlations between jurors' actual awards and the awards predicted for them by judges and lawyers tell us how well those legal professionals could anticipate which injuries would evoke larger and which would evoke smaller awards from jurors; that is, the extent to which judges and lawyers could predict the vertical variation in jurors' responses to injuries. These correlations varied considerably. In Illinois, the correlations between jurors' actual awards and the awards predicted for them by both plaintiffs' lawyers¹³⁸ and defense lawyers¹³⁹ were significant, while the correlation with the awards predicted by judges was only marginally significant.¹⁴⁰ In New York, the correlations between jurors' actual awards and the awards predicted by judges¹⁴¹ and defense lawyers¹⁴² were significant, but the correlation with plaintiffs' lawyers' predicted awards was not.¹⁴³ The range in these correlations (from .21 to .78) is striking, especially considering that it is plaintiffs' lawyers who do both best (in Illinois) and worst (in New York) at anticipating the vertical variation in jurors' reactions to the injuries.

D. *The Stabilizing Power of Aggregation into Groups*

Thus far, we have been speaking as though jurors made decisions individually, as judges do in bench trials, rather than in groups, as juries. The greater observed variability in the awards of jurors, as com-

134. $t(61) = 2.19, p < .05$.

135. $t(61) = 2.45, p < .05$.

136. $t(61) = 2.10, p < .05$.

137. $t(61) = 2.49, p < .05$.

138. $r(60) = .78, p < .001$.

139. $r(60) = .37, p < .01$.

140. $r(60) = .24, p = .06$.

141. $r(60) = .46, p < .001$.

142. $r(60) = .47, p < .001$.

143. $r(60) = .21, p = .10$.

pared to those of judges and lawyers, would decrease if the individuals were combined into juries, as they are in actual trials. That group decisions decrease in variability (and increase in predictability) compared to the decisions of the individuals who compose the groups, is to be expected both from basic statistical¹⁴⁴ and social psychological¹⁴⁵ properties of groups. In addition, studies that have tested this question directly using mock jurors and juries have found that, for the same case, awards of deliberating juries are less variable than awards of jurors deciding damages as individuals.¹⁴⁶ Thus, the data we have presented to this point, on individuals, overstates the variability.¹⁴⁷

Our final set of analyses is essentially (and merely) a demonstration of the phenomenon that when individuals decide as groups their

144. Larger samples produce distributions with smaller standard errors than smaller samples. For example, a set of samples of size 6 or 12 as compared to samples of size 1.

145. People in decisionmaking groups tend to coalesce toward a consensus; their shifts tend to be toward, rather than away from, each other. This is a basic aspect of group behavior, which has long been observed. See, e.g., A. PAUL HARE, *HANDBOOK OF SMALL GROUP RESEARCH* 19 (1976) (groups tend to establish norms which are viewed as legitimate by group members); MUZAFIR SHERIF, *THE PSYCHOLOGY OF SOCIAL NORMS* 104 (1936) (people in social interaction gravitate toward sharing the perceptions and judgments of those with whom they interact).

146. See Shari Seidman Diamond et al., *Juror Judgments About Liability and Damages: Sources of Variability and Ways to Increase Consistency*, 48 *DEPAULL. REV.* 301, 316 tbl. IV (1998) (finding the standard deviation of their individual jurors' general damages awards to be \$4,255,056 compared to \$714,556 for decisions on the same case facts by deliberating juries); Shari S. Diamond & Jonathan D. Casper, *UNDERSTANDING JURIES* (forthcoming) (using data from a completely different set of jurors and a completely different case and finding the standard deviation of jurors' awards to be \$175,988 and of juries' awards to be \$122,749).

147. The only suggestion to the contrary of which we are aware appears in Sunstein et al., *supra* note 51. They allude to the familiar notion that larger groups (such as juries) show less variability than smaller groups or individuals. But they also argue that there exists an "amplification of bias" by which group deliberation would produce greater variance than was found in their study of individual mock jurors. See *id.* at 2101 n.128. This argument confuses several principles upon which it relies. Sunstein et al. are correct that group deliberation produces group polarization, a post-deliberation shift toward positions more extreme than the prediscussion central tendency of the individuals. See, e.g., Martin F. Kaplan, *Discussion Polarization Effects in a Modified Jury Decision Paradigm: Informational Influences*, 40 *SOCIOMETRY* 262 (1977); Serge Moscovici & Marisa Zavalloni, *The Group as a Polarizer of Attitudes*, 12 *J. PERSONALITY & SOC. PSYCHOL.* 125 (1969); David G. Myers & Helmut Lamm, *The Group Polarization Phenomenon*, 83 *PSYCHOL. BULL.* 602 (1976). But they overlook the fact that groups tend to coalesce around that more extreme position. Thus, in a given case, groups will be more consistent than individuals at the same time that they will be more extreme (in whichever direction). When Sunstein et al. jump from the phenomenon of deliberation-induced shifts to the conclusion of "even greater variance," Sunstein et al., *supra* note 51, at 2101 n.128, they have jumped from thinking about the variance for a given case to the variance of the mass of cases. That is, some *types of cases* will shift one way and some the other way, resulting in greater variance across a set of *different cases*. But that phenomenon is irrelevant to the policy issue of variability among *like cases*. The unwanted variability in cases is not that different cases receive different awards (vertical equity requires that they do), but that any given case would receive different awards from different decisionmakers (that would be horizontal inequity). The available evidence supports the conclusion that variance declines when we move from jurors to juries. See *supra* note 146.

decisions are more consistent and predictable than when they decide the same question as individuals. These analyses were conducted on only the first of the two injuries evaluated by each of our respondents; that is, all 62 injuries were included, but only the data of those that were the first of the pair a respondent heard.

Treating each decision by a judge as if it were a decision in a bench trial, we calculated the standard deviation of those raw general damages awards. This is analogous to calculating the standard deviation for a set of awards from a sample of bench trials in a jurisdiction. When considering the same set of injuries, the standard deviation of awards made by Illinois judges was approximately 2.3 million and by New York judges about 1.5 million. Doing the same analysis for individual jurors, as if they decided cases individually as judges do, we find much greater variability in response to the same set of injuries. The standard deviation for Illinois jurors was about 10.0 million, and for New York jurors about 16.6 million.

To estimate the awards that would be made if individual jurors had been grouped into juries and asked to decide awards as a group, we created statistical groups by taking the median¹⁴⁸ of the raw awards of all jurors who evaluated the same injury, for each of the 62 injuries, separately within each state. In this way we made "juries" of 4 or 5 persons each. Although the mix of injuries is as varied as in either of the preceding analyses, the standard deviation shrinks considerably: to about 755,000 for Illinois "juries" and about 543,000 for New York "juries." Thus, the awards of *juries* were less varied than those of judges.¹⁴⁹

These analyses do not purport to give the actual standard deviation that would be obtained by combining the actual individuals into groups and having them reach consensus in group deliberation. These analyses merely illustrate the phenomenon and its general tendencies, namely, that groups produce more stable and more predictable estimates than individuals.¹⁵⁰ Considering that there were few extreme

148. The median of individual judgments provides a very good estimate of the group judgment. See Shari Seidman Diamond & Jonathan D. Casper, *Blindfolding the Jury to Verdict Consequences: Damages, Experts, and the Civil Jury*, 26 L. & SOC'Y REV. 513, 545 tbl. 4 (1992), which found that using medians of individual pre-deliberation judgments to estimate group decisions correlated .62 with the decisions of deliberating juries composed of those same individuals. By contrast, other plausible methods of statistically estimating the group decision had lower, sometimes much lower, correlations.

149. A better test of this, but a far more expensive one, would have jurors deliberate as groups, and would treat those group judgments as the ones to be modeled and compared to those of judges and lawyers.

150. Moreover, larger groups do so to a greater extent than smaller groups. See Michael J. Saks, *The Smaller the Jury, the Greater the Unpredictability*, 79 JUDICATURE 263 (1996); see also Michael J. Saks & Mollie Weighner Marti, *A Meta-Analysis of the Effects of Jury Size*, 21 LAW & HUM. BEHAV. 451, 463-64 (1997).

outliers to begin with,¹⁵¹ it is easy to see how a juror recommending an extreme award could be outnumbered and reined in by the majority of jurors preferring more moderate awards.

V. CONCLUSIONS

This concluding Part will begin by summarizing the study's findings (Sections A through C). Section D will consider the implications of the findings for legal policy concerning the determination of general damages by juries. Section E will raise further empirical research questions that remain to be asked and answered.

A. *Perceptions of Injury Severity*

Several important conclusions can be drawn from the analyses examining the different decisionmakers' judgments of the overall severity of the injuries.

First, the regression models suggest that different decisionmakers — people with different roles in the legal system, different experience with personal injury cases, and different demographic backgrounds — relied on the same injury attributes in similar ways and gave them similar relative weight when evaluating the severity of injuries.¹⁵² For all groups, the extent of perceived disability and mental suffering were the strongest predictors of overall severity judgments, with disfigurement and pain making smaller contributions.¹⁵³ The only notable departure from the pattern was the model for defense lawyers, for whom mental suffering had the most powerful influence (whereas for the other three groups it was disability) and pain did not have a significant unique influence. Nevertheless, all of the regression models accounted for a large, and similar, proportion of variance in the decisionmakers' judgments of overall severity: 69% for judges, 72% for jurors,¹⁵⁴ 75% for defense lawyers, and 78% for plaintiff's lawyers.

Second, all four groups of decisionmakers had generally similar perceptions of what injuries are of relatively greater or lesser severity. Shared vertical equity is evident from the finding that jurors' overall severity assessments were highly correlated with those of the other groups, ranging from .72 to .84 in Illinois and from .81 to .90 in New

151. See the Measures section, *supra* Section III.E.1, for a discussion of how we identified extreme outliers. They constituted 2.32% of the jurors' awards. For comparison, extreme outliers constituted 1.03% of judges' and 1.21% of lawyers' awards.

152. See Table 1.

153. A similar pattern was seen in the two studies in Wissler et al., *supra* note 27.

154. Similarly, in the two studies in Wissler et al., *supra* note 27, at 189, 200, undergraduates' injury perceptions accounted for 75% and 89%, respectively, of the variance in their overall severity assessments.

York. Jurors tended to assess injuries as being more severe than the other groups did, while defense lawyers tended, even more strongly, to see injuries as less severe than the other groups did. Nonetheless, injuries seen as more severe by one group tended strongly to be seen as more severe by the other groups.

The remarkable similarity across the varied decisionmakers' response patterns when judging the severity of injuries has important implications. The findings suggest that jurors approach the task of evaluating injuries no differently than judges and lawyers do. Thus, assertions that jurors cannot intelligently evaluate injuries or have idiosyncratic reactions to injuries are simply incorrect, or at least ill-focused. Indeed, if any one group emerges as being out of step with all of the others, it is defense lawyers. This has somewhat paradoxical implications. When members of the defense bar evaluate the performance of jurors, and gauge them to be off the mark,¹⁵⁵ these lawyers no doubt reach that assessment by comparing the jurors' conclusions to their own. But their own impressions of injuries are the ones that depart most from the pattern shared by the other decisionmaking groups, at least in regard to judgments of injury severity.

Third, for all groups, while the core attributes of injuries accounted for the lion's share of variance in severity judgments, the sociodemographic variables played essentially no role. Because a finding that the background characteristics of the decisionmakers have little or no impact on their decisions may surprise readers unacquainted with the relevant literature, it may be worth mentioning that this merely extends a finding now well established elsewhere in the jury decision-making literature. Though most of that research has been on criminal trials, juror sociodemographic characteristics also have been found to play only a modest role in their civil liability verdicts,¹⁵⁶ certainly compared to the dominant impact of evidence and arguments presented in the cases. The present study, along with other recent research, suggests that what had been learned about liability verdicts can be extended to damages-relevant responses: individual differences make little difference.¹⁵⁷

155. For example, James Griffith, an attorney specializing in medical malpractice defense, wrote, "There's no limit on what jurors can award for pain and suffering, so too often they act like Santa Claus, handing out millions of dollars in cases involving comparatively minor injuries." James Griffith, *What it Will Take to Solve the Malpractice Crisis*, MED. ECON. Sept. 27, 1982, at 195.

156. See Shari Seidman Diamond, *Scientific Jury Selection: What Social Scientists Know and Do Not Know*, 73 JUDICATURE 178, 180 (1990); Steven D. Penrod, *Predictors of Jury Decision-making in Criminal and Civil Cases: A Field Experiment*, 3 FORENSIC REP. 261 (1990); Michael J. Saks, *What Do Jury Experiments Tell Us About How Juries (Should) Make Decisions?*, 6 S. CAL. INTERDISC. L.J. 1, 9-13 (1998).

157. See Diamond et al., *supra* note 146, at 314 (finding, in a study that used a videotaped presentation of a highly realistic simulated trial to 1042 mock jurors from Cook County, Illinois, that "[o]ur attempt to trace the sources of this variation [in damages

B. General Damages Awards

The picture becomes somewhat blurred when the transition is made from perceptions of the overall severity of injuries to placing dollar values on them. First, for awards, the regression models for the different groups of decisionmakers differed more from each other in terms of the factors that account for awards as well as the relative influence of the core injury attributes and the sociodemographic characteristics.¹⁵⁸ Second, the model for each group of decisionmakers accounted for a smaller proportion of the variance in awards than in severity assessments.

While similarities were found, different respondent groups gave differing weight to the various injury attributes when awarding damages. Defense lawyers differed from all other groups in that for them mental suffering weighed more heavily than disability. And plaintiffs' lawyers differed from all other groups in that for them disfigurement did not have a significant impact on awards. Thus, the impact of the injury attributes on jurors' awards more closely resembled their impact on judges' awards, and their impact on the lawyers' awards diverged from both judges and jurors. Nonetheless, there were important similarities among the decisionmakers regarding which injury attributes were predictive of their general damages awards. For all respondent groups, perceived pain was related to awards only indirectly through its association with the other injury attributes, while greater perceived disability and mental suffering were strongly associated with larger awards.¹⁵⁹

Compared to the models of severity assessments, sociodemographic characteristics played a larger role in awards, both directly and through their interaction with other factors. For jurors, gender and income were predictive in both states. Men and wealthier jurors awarded more than women and poorer jurors. One explanation for these findings might be that in a task as undefined as awarding general damages, people rely on the only reference scale with which they are familiar, namely their own. Those who are accustomed to dealing in larger amounts fit the cases onto a mental reference scale that runs into higher numbers. This would explain not only the finding that wealthier individuals give higher awards, but also that on average men give higher awards than women. Another explanation might be more cultural: that something in their value system leads people of higher

awards] to individual differences in the background and attitude of jurors met with little success"); *see also id.* at 307 tbl. I.

158. *See* Table 2.

159. For jurors and lawyers, some of the main effects of the injury attributes are qualified by interactions. *See* the Results Section, *supra* Section IV.B, for the details.

socioeconomic status to place a higher monetary value on personal injuries and their sequella.

In contrast to jurors, the significant demographic variable for judges and lawyers in this study was state. Generally, judges and lawyers in New York gave larger awards than those in Illinois. Several explanations are possible. For the same reasons that wealthier jurors gave larger awards, it may be that wealthier lawyers and judges (i.e., the ones in New York) give larger awards. Another possible explanation is the differences between the two states in the laws regarding damages awards. In New York, loss of enjoyment of life is considered a distinct element of damages, but in Illinois it is subsumed as part of mental suffering. This would be expected to lead to higher awards by New York judges and lawyers. Though all respondents were read the instructions based on those in their state, the same effect might not have arisen for jurors because they are less likely to be attuned to the finer details in the instructions than are lawyers and judges (indeed, most of the lawyers and judges presumably already know their respective state's rules).¹⁶⁰ Yet another explanation might be that the judges and lawyers in each state know that theirs is thought to be a relatively low or high award jurisdiction and act in accord with those expectations, while jurors do not know that they are "supposed" to give higher awards if they are in New York and lower awards if they are in Illinois.

Nonetheless, the paucity of relationships between most of the demographic variables and awards (collectively accounting for only 1% to 4% of the variance in awards) is the more important part of the story.¹⁶¹ Several of these non-differences are especially worthy of note. Illinois jurors did not give awards that differed from those of New York jurors. And the awards of jurors living in rural areas did not differ from those of jurors living in urban areas.¹⁶² These non-differences contrast with findings of archival studies, which suggested that, on a national spectrum, New York jurors give relatively high awards while Illinois jurors give relatively low awards,¹⁶³ and that ur-

160. Instructions are often difficult for jurors to understand. See Michael J. Saks, *Judicial Nullification*, 68 IND. L.J. 1281, 1282-83 (1993).

161. As noted previously, research has found little relationship between jurors' socio-demographic attributes and the verdicts of the jurors or the juries on which they sit. See *supra* notes 156-157 and accompanying text.

162. Although rural or urban locale did not affect the size of awards, it did interact with perceived disfigurement such that disfigurement had a significant impact on the awards of urban jurors but not rural jurors.

163. See DANIELS & MARTIN, *supra* note 12, at 69-90. A possible explanation for the differences observed by Daniels and Martin is that their data were based on total awards, and interstate differences in medical expenses, earnings, and the cost of living could account for differences in awards.

ban jurors award more than rural jurors.¹⁶⁴ Those archival studies could not control for the kinds of cases that jurors in the different locales were deciding because they had no data on the nature of the cases themselves, so they could not determine whether observed differences were due to differences in the tendencies of the jurors or to differences in the cases being presented to those jurors. Because the current study presented an identical set of cases to jurors in the two regions of the two states, our findings suggest that the differences observed in those other studies were due not to differences in the jurors but to differences in the mix of cases that came to trial in the different locales.¹⁶⁵

In addition to differences in the components of the models, awards were less predictable than were severity assessments. The regression model for each group, especially for jurors, accounted for a smaller proportion of the variance in awards: 23% for jurors, 42% for judges, 48% for plaintiffs' lawyers, and 58% for defense lawyers.¹⁶⁶ The lack of information about the plaintiff's gender and pre-injury lifestyle might have contributed to the variability in awards for all groups. The jurors' lack of experience in assigning a dollar value seems an obvious explanation for the greater variability and lower predictability of jurors' awards.

164. See DANZON, *supra* note 33, at 63, 74-75. A possible explanation for the differences observed by Danzon is that her data could reflect differences in medical expenses, earnings, and the cost of living in urban versus rural areas.

165. Another difference is that Danzon's research used *states* as the unit of analysis, rather than individuals or juries. Inferences drawn about relationships at the state level may or may not hold at smaller levels of aggregation. This problem is known as the "ecological fallacy." Thus, while Danzon found that urban areas (actually, relatively more urban states) experienced larger awards than more rural areas (actually, relatively more rural states), there is no reason to expect her findings to apply to urban versus rural areas within states, or to people from urban versus rural areas.

166. It is worth noting that the proportions of variance in awards accounted for by the present data are lower than those of other studies cited in our Introduction. This difference may suggest an important cognitive insight. Our study had respondents rating each injury without comparison to others, or in comparison with one other injury (when rating the second in the pair). Most of the other studies had respondents or coders rating larger numbers of injuries. No doubt if we had people rank order our 62 cases and then used that as a predictor of awards, the correlation would rise considerably. In our own studies of damages, all using similar procedures, when each respondent evaluated 12 or 13 cases, the $R^2 = .74$. See Wissler et al., *supra* note 27, at 201. When each respondent evaluated 5 cases, the $R^2 = .54$. See *id.* at 195. And in the present study, where each juror evaluated 2 cases (so half the time were evaluating the first of the pair), the $R^2 = .23$. If this insight is correct, the higher proportions of variance accounted for in the models of judges and lawyers reflect their having (in their heads, from the numerous other cases they have known) a more complete cognitive reference scale of injuries to provide a context for thinking about the case now under consideration. That, in turn, suggests a relatively simple and yet potentially powerful reform — providing jurors with a frame of reference for determining the appropriate general damages award — that should produce considerable improvement in the predictability and stability of jurors' awards of general damages. The issue of providing jurors with such guidance is discussed below.

The decreased predictability of jurors' awards might also be explained by the difference in education, income, and gender between jurors (as a group) compared to judges and lawyers.¹⁶⁷ When looking only at those subgroups of jurors who most closely approximate judges and lawyers on these dimensions, we find the greatest predictability of awards compared to those jurors who are least similar. For jurors with post-college study or an advanced degree, the full regression model accounted for 29% of the variance in awards, with progressively lesser amounts of variance accounted for as the level of education declined, culminating in a model that accounted for 14% of the variance in awards for jurors who had not graduated from high school. For jurors with household incomes over \$75,000, the model accounted for 32% of the variance in awards, with progressively lesser amounts of variance accounted for as the level of income declined, culminating in a model that accounted for 20% of the variance in awards for jurors with incomes below \$15,000. And for male jurors, the model accounted for 27% of the variance in awards, compared to 15% for women jurors. Thus, to some extent, the differences in the predictability of jurors' awards, compared to those of judges and lawyers, is a product of social differences and not simply differences in the extent of experience in awarding general damages for injuries.

Given the small variability in assessments of injury severity within each respondent group and the strong relationship between overall severity assessments and awards, the variability in awards would seem to be due largely to differences in the metrics people use to convert perceptions of injury severity into awards.¹⁶⁸ Why didn't we find more variability in the severity ratings instead of only in the dollar awards? Perhaps the five-point severity rating scale limits the expression of one's evaluation of injuries or consideration of subtle distinctions among them, compared to the more fine-grained and unbounded nature of dollar scales. Thus, the lack of variability in severity ratings might be due to the insufficiency of the rating scale producing an apparent consistency in judgments rather than to a truly better quality of decisionmaking.

Despite the lesser predictability of awards, the correlations between severity ratings and awards (using the injury scenarios as the unit of analysis) showed a strong association between the degree of perceived injury severity and the size of the awards given. These correlations were .83 for jurors, .88 for judges, .77 for plaintiffs' lawyers, and .88 for defense lawyers. This finding replicates the high degree of vertical equity in awards found in other studies.¹⁶⁹ In addition, shared

167. See *supra* Section III.C.

168. See related discussions in Sunstein et al., *supra* note 51, at 2106-07, and Wissler et al., *supra* note 27, at 193.

169. See studies cited *supra* notes 26-27.

vertical equity is evident from the finding that jurors' damages awards were well correlated with those of the other groups, ranging from .63 to .73 in Illinois and from .60 to .69 in New York.¹⁷⁰ These findings indicate that all groups tended to agree on which injuries deserved higher general damages awards and which deserved smaller ones.

Even an examination of the absolute level of awards, however, does not find jurors to be out of line. Averaged over all cases, jurors and plaintiffs' lawyers tended to give the highest awards, defense lawyers the lowest, and judges gave an intermediate amount.¹⁷¹ What might account for the higher awards given by jurors? One obvious "statistical" explanation is that the greater variability of individual jurors' awards allows the average to shift upward more readily than downward. A second, "vertical equity," explanation is that it is appropriate that the jurors' awards would be larger than those of other groups because jurors rated the injuries as more severe. A third possible explanation is an artifact of the media's reporting of jury awards, which tends to cover higher award amounts.¹⁷² While lawyers and judges would be aware that there are many small awards that are not reported for every large verdict that is reported, jurors are less likely to know that, and their frames of reference would be affected accordingly. A final possible explanation is an artifact of our interview procedures. By asking judges and lawyers, but not jurors, in both states to make what they regarded as appropriate awards, regardless of any caps in effect in their states, we were unavoidably reminding them of caps, which might have artificially pushed down their awards.

In sum, the way in which the damages decisionmaking models of jurors differ most from those of judges and lawyers is in their variability, not in the patterns of the factors that affect their awards. And this variability among *jurors* is almost certain to decline when the decisionmakers consist of *juries*.¹⁷³ This study was not able to do more than demonstrate that general effect: the awards of (statistically created) *juries* were less varied than those of judges, while the awards of *jurors* were far more varied than those of judges.

170. The findings of Vidmar & Rice, *supra* note 18 — that jurors and arbitrators have similar patterns of correlations of perceptions of disfigurement, etc., with awards — parallel our own. In both states, our findings showed the lowest correlation was between jurors and plaintiffs' lawyers, suggesting an interesting area of further study.

171. See Table 4.

172. See Daniel S. Bailis & Robert J. MacCoun, *Estimating Liability Risks with the Media as Your Guide: A Content Analysis of Media Coverage of Tort Litigation*, 20 LAW & HUM. BEHAV. 419, 421, 426 (1996) ("It is a small step to assert that reading about million-dollar verdicts might lead people to overestimate the frequency of large jury awards."); see also Michael J. Saks, *Public Opinion About the Civil Jury: Can Reality be Found in the Illusions?*, 48 DEPAUL L. REV. 221, 231-34 (1998).

173. See discussion *supra* notes 144-151 and accompanying text.

C. *Judges' and Lawyers' Predictions of the Awards of the "Average Juror"*

When asked to predict the awards that would be made by the "average juror," judges in New York, defense lawyers in both states, and the jurors themselves, predicted that the average juror would give awards that were larger than they themselves would give in the same case. Apparently, the members of most groups, with plaintiffs' lawyers being the unsurprising exception, think that the awards they personally would give are lower than those given by jurors. Despite these differences in the predicted awards, judges and lawyers expected jurors to share their relative assessments of damages — giving larger awards in the cases where they themselves gave larger awards, and smaller awards in the cases where they themselves gave smaller awards.¹⁷⁴ Thus, all groups expected the average juror's awards to have vertical equity that strongly paralleled their own, but to be shifted upward.

The awards our jurors actually gave tended to be smaller than the awards the judges and lawyers predicted they would give. One might expect the expertise of judges and lawyers to include accurate predictions of what jurors will do. We would be hesitant to make too much of this systematic bias (overprediction), except that it is consistent with findings where lawyers have been asked to estimate the size of jury awards in their states.¹⁷⁵ Systematic errors call for explanations; we can only speculate on one. A generations-old legal folklore about the excesses of civil jurors, compounded by the public rhetoric of more recent years, may prevent lawyers from recognizing what jurors actually do. Even without such folklore, or perhaps at the root of it, the "availability heuristic" might be operating¹⁷⁶ — a cognitive bias which suggests that judges and lawyers will have a tendency to more easily recall surprisingly large awards, and then to estimate the overall behavior of juries from those more "available," but less representative, instances.¹⁷⁷

174. The correlations between the lawyers' and judges' own awards and the awards they predicted the "average juror" would give ranged from .49 to .93 in Illinois and from .89 to .96 in New York.

175. See, e.g., Donald R. Songer, *Tort Reform in South Carolina: The Effect of Empirical Research on Elite Perceptions Concerning Jury Verdicts*, 39 S.C. L. REV. 585 (1988).

176. See Amos Tversky & Daniel Kahneman, *Availability: A Heuristic for Judging Frequency and Probability*, 5 COGNITIVE PSYCHOL. 207 (1973); Amos Tversky & Daniel Kahneman, *Judgment under Uncertainty: Heuristics and Biases*, 185 SCIENCE 1124 (1974).

177. The correlations between what judges and lawyers predicted the average juror would award and what our jurors in fact awarded present an odd pattern, with some of the most wide-ranging correlations found in the entire study. Plaintiffs' lawyers do both the best (in Illinois, $r = .78$) and the worst (in New York, $r = .21$) job of predicting our jurors' awards in the cases. Perhaps the lack of a real pattern is the most one can say about these correlations.

D. Policy Implications

The dominant theme of these findings is one of considerable similarity across the various groups of decisionmakers in the structure of thinking about injury severity and awards. Most importantly, an impressive similarity exists in the injury attributes that drive their decisions, the weight given to those attributes, and the shared sense of vertical equity held by jurors, judges, plaintiffs' lawyers, and defense lawyers alike. These findings suggest that commonly voiced speculations about the inability or irrationality of jurors in evaluating injuries are misconceived, because on that task jurors were nearly indistinguishable from judges and lawyers. These data also reveal, contrary to some discussions in the literature, no differences between jurors in two different states and in different regions of the same state when they are asked to respond to the same set of injuries.

If there is any basis for concern, it is the translation from perceptions of injuries into dollar valuations. But that is a problem not with jurors alone, but with the task — judges and lawyers similarly display decreased predictability, albeit a smaller decrease, as they move from severity assessment to damages valuation, and reveal more between-group differences in the models that predict their decisions. Thus, improvements, if justified as desirable policy, should be concerned less with who is to make the decisions about general damages and more with making the task more achievable.

1. *Whether the Role of the Jury Should Be Reduced*

Should the role of the jury in making general damages awards be reduced or eliminated? Should the determination of those awards, as others have suggested in the context of punitive damages, be transferred from jurors to judges?¹⁷⁸ The main argument for doing this is that judges would produce more predictable awards than juries.¹⁷⁹

178. See Sunstein et al., *supra* note 51, at 2112-13 ("It would be reasonable to react to our study by suggesting a simple reform: Juries should decide questions of civil liability, just as they do questions of criminal liability. But judges should decide on the appropriate level of punitive damages, just as they do criminal punishment, subject, in both cases, to guidelines laid down in advance."). Sunstein et al. presented 28 variations of 10 case vignettes to 899 randomly selected registered voters in Travis County, Texas, and asked them to (a) evaluate the cases, and (b) recommend punitive damages awards to assess against the defendants in those cases. *Id.* at 2095. Sunstein et al. found that jurors were breathtakingly consistent in their judgments of the defendant's outrageousness ($r = .99$), but that variability increased (reliability and predictability decreased) in the second step, namely translating those judgments into punitive awards ($r = .42$). *Id.* at 2098, 2103. Sunstein et al. conclude that "individual differences in dollar awards produce severe unpredictability and highly erratic outcomes," *id.* at 2103, and implicitly assume, without any data, that judges would make very different decisions than jurors do, *id.* at 2113.

179. Perhaps it goes without saying that predictability and consistency might be purchased at the cost of more important values. For example, the rule of thumb some lawyers use to come up with a figure for general damages for purposes of settlement negotiation —

First of all, it is not clear from the empirical evidence that transferring the task from juries to judges would increase predictability at all, much less to the satisfaction of those calling for such a reform. Our findings show a fairly high correlation between awards given by judges and those given by individual jurors (r 's = .73 and .69 in the two states respectively).¹⁸⁰ Based on the decision models we developed, the predictability of jurors' general damages awards, expressed as a correlation, is $r = .48$, while that of judges is .65.¹⁸¹ That suggests that judges also are far from perfect. And jurors, as well as judges, perform this task at a level of reliability that compares favorably to that of other important societal decisionmaking. As Table 5 indicates, juror evaluations of injury severity and damages are of comparable consistency and predictability to the decisions of a considerable array of other decisionmakers who are thought to be far more expert and who are performing tasks thought to be much better defined than that of assessing general damages.

multiplying medical specials by three — would, if adopted as a legal rule for assigning damages, produce highly predictable outcomes, but at the expense of injury- and victim-specific considerations. Sometimes the amount awarded would be too great and sometimes it would be too small.

It may be worth noting that some level of variability is beneficial to the legal process. The proper level of uncertainty helps to promote settlements, while not promoting so many that courts do not receive the cases necessary to monitor society's disputes, and thereby to continually refine and announce the law. For more detailed discussion of this point, see Michael J. Saks, *Enhancing and Restraining Accuracy in Adjudication*, 51 LAW & CONTEMP. PROBS. 243 (1988). A rational goal would be to aim to calibrate the level of uncertainty to seek its optimal level, rather than to aim to eliminate uncertainty from the litigation system.

180. Thus, we found higher reliability in awards (in terms of agreement between judges and jurors) than Sunstein et al. did (in terms of agreement among jurors). Perhaps this reflects differences in the challenge of assigning dollar values for general damages versus for punitive damages. If so, it implies that Sunstein et al.'s speculation that their findings would generalize to other juror decision tasks about damages is unwarranted. Sunstein et al., *supra* note 51, at 2131-38.

181. Taking the square root of the R-squareds in Tables 2A and 2B, for jurors and judges, respectively.

TABLE 5: RELIABILITY OF VARIOUS DECISIONMAKERS MAKING VARIOUS DECISIONS

<i>Decisionmakers and Tasks</i>	<i>Strength of Association</i>
Jurors evaluating outrageousness of tortious conduct (reliability) ^a	.99
Jurors and Judges evaluating injury severity (reliability) (in IL & NY, respectively) ^b	.84, .90
Jurors evaluating injury severity (predictive validity) ^b	.85
Judges evaluating injury severity (predictive validity) ^b	.83
Jurors and Judges deciding amount of general damages (reliability) (in IL & NY, respectively) ^b	.73, .69
Judges deciding general damages (predictive validity) ^b	.65
Juries and Judges deciding on personal injury liability (reliability) ^c	.58
Federal judges deciding criminal sentences (reliability) ^d	.58
Four different kinds of business experts rating strategic attributes associated with competitive moves in 15 different companies (reliability) ^e	.55, .55, .50, .47
Scientists deciding merits of research proposals (reliability) ^f	.50
Jurors deciding general damages (predictive validity) ^b	.48
Doctors diagnosing skin diseases (reliability) (three different diseases) ^g	.36, .48, .48
Jurors deciding punitive damages (reliability) ^a	.42
Employment interviewers deciding on qualifications for hiring (reliability) ^h	.40
Navy pilots rating technical effectiveness of aircraft (reliability) (three different sets of measures) ⁱ	.28, .12, .16
Psychiatric diagnoses (reliability) ⁱ	.13, .22

Note: Pearson correlations or multiple correlations are in bold; Cohen kappas are in regular font. In a few instances, the tabled correlation values had to be calculated from rates of agreement given in the original sources.

The data in the table are taken from the following studies:

- a. Cass R. Sunstein et al., *Assessing Punitive Damages (with Notes on Cognition and Valuation in Law)*, 107 YALE L.J. 2071 (1998).
- b. The study reported in this Article.
- c. Henry Kalven, Jr., *The Dignity of the Civil Jury*, 50 VA. L. REV. 1055 (1964).
- d. Shari S. Diamond & Hans Zeisel, *Sentencing Councils: A Study of Sentence Disparity and Its Reduction*, 43 U. CHI. L. REV. 109 (1975).
- e. Ming-Jer Chen et al., *An Exploration of the Expertness of Outside Informants*, 36 ACADEMY OF MANAGEMENT J. 1614 (1993).

- f. S. Cole et al., *Chance and Consensus in Peer Review*, 214 SCIENCE 881 (1981).
- g. John D. Whited et al., *Primary Care Clinicians' Performance for Detecting Actinic Keratoses and Skin Cancer*, 157 ARCHIVES OF INTERNAL MEDICINE 985 (1997).
- h. R.S. Uhrbrock, *The Personnel Interview*, 1 PERSONNEL PSYCHOLOGY 273 (1948).
- i. Leonard Adelman & Michael L. Donnell, *An Empirical Study Comparing Pilots' Interrater Reliability Ratings for Workload and Effectiveness*, 18 IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS 978 (1988).
- j. Pierre Leichner et al., *A Study of the Reliability of the Clinical Oral Examination in Psychiatry*, 29 CANADIAN J. PSYCHIATRY 394 (1984).

Moreover, there is good reason to believe that juries (groups) produce considerably more stable and predictable outcomes compared to jurors (individuals).¹⁸² Consequently, decisions of *juries* may approach, or even exceed, the stability and predictability of decisions of individual judges. Without persuasive data to the contrary, entertaining a reform that would reduce or eliminate the role of juries in making damages awards seems premature.

But assuming that, at the end of the day, it is concluded that a marginal advantage in the predictability of general damages awards favoring judges does exist, is the transfer of responsibility from juries to judges wise policy? The answer lies in weighing this small gain in predictability against what would be lost.

A core benefit of juries is the sampling function they provide. The law apparently has discovered that the best substitute for a market to generate the values of the sorts of losses subsumed under general damages is to sample the judgment of the community. Moreover, if it is to tailor damages to the actual losses incurred, it must do its sampling on a case-by-case basis. To accomplish these functions, juries not only are quite an efficient method, they are the only available method. Without juries, the ability to estimate the community's sense of the value of the ability to see or to hear or to walk or to have children or to be a whole human being is forfeited. That juries consist of groups, drawn more or less at random from the community, is essential to the sampling function. Imperfect sampling though it may be, it is less imperfect than having no sampling at all.

In addition, juries offer several collateral benefits. One is that they reduce bias by dividing decisionmaking functions between judge and jury. The judge is free to know everything, and can act to screen out biasing, speculative, or otherwise inadmissible evidence and arguments about what the damages should be. The jury then is able to consider only proper evidence and arguments in reaching its decisions. Even when this is imperfect, it is a considerable improvement on re-

182. See *supra* Section IV.D.

quiring judges to know what a factfinder should not know, and then to try to decide in spite of that knowledge.¹⁸³

Another collateral benefit is the “lightning rod” function of juries. That is, by being the primary decisionmakers, juries become the targets of criticism, thereby drawing the fire away from judges. Moreover, juries can dissipate the attacks better than judges could. A jury that attracts the ire of the public fades back into the community and is replaced by another jury for the next trial. In contrast, as judges render one verdict after another, they risk accumulating criticism for successive unpopular decisions. Further, the unpopular decisions of judges, unlike those of juries, might be viewed as the edicts of an unelected elite rather than as the sentiments of the general community. The very criticism directed at some juries may illustrate their performance of a lightning rod function.¹⁸⁴ To say the least, it is not obvious that the marginal benefit of increased predictability of judges’ decisions (if such a benefit were to materialize) outweighs the various benefits provided by the use of juries — especially considering that most of the time juries reach decisions that judges regard as reasonable.¹⁸⁵

2. Reforms Consistent with the Available Data

Sensible and moderate reforms to reduce unpredictability, focusing on procedural adjustments, could improve on an essentially sound system.

One approach would be to increase the size of juries (from the six or eight civil jurors now found in many federal and state jurisdictions to the traditional twelve of the common law) in order to reduce unwanted variability and increase stability and predictability.¹⁸⁶

A second method, which emphasizes the diversity among injuries, would be to provide jurors with descriptions of a range of injuries as a basis for forming a reference scale with which to compare the case at bar.¹⁸⁷ By seeing the injury under consideration in the context of other

183. Of course, many judges believe that they can perform this cognitive feat, but what little hard data exist on the subject indicates that judges are no better than jurors at disregarding evidence they are not supposed to know. See Landsman & Rakos, *supra* note 18.

184. Especially in cases where jurors have made unpopular decisions with which judges have not manifested disagreement.

185. See Valerie P. Hans, *Attitudes Toward the Civil Jury: A Crisis of Confidence?*, in VERDICT: ASSESSING THE CIVIL JURY SYSTEM 248, 262-65 (Robert E. Litan ed., 1993). It is also worth noting that the law has mechanisms for reining in the rare runaway jury, namely additur/remittitur review and appeals.

186. See Diamond et al., *supra* note 146; Saks, *supra* note 150; Saks & Marti, *supra* note 150.

187. See *supra* note 166 and accompanying text; see also Sunstein et al., *supra* note 51, at 2104.

injuries, jurors would gain the advantage possessed by judges and lawyers and insurance claims representatives — an appreciation of the diverse universe of injuries.

Another powerful yet modest reform would be to pool jury awards made for similar injuries, and to present these cases and their award distributions to juries for guidance in reaching their general damages awards¹⁸⁸ and to judges for conducting their additur/remittitur reviews.¹⁸⁹ New decisions would continuously be added to update the data pool.¹⁹⁰ In either use, they could be offered as true “guidelines” — suggestions which could be departed from to the extent that the instant case differed from those guideline cases.¹⁹¹ Pooling verdicts in similar cases, and funneling those back to juries (or judges), magnifies the benefit of sampling community judgments. Thus, rather than reducing the role of the jury, we can reduce a possible disadvantage (unpredictability) by making more use of one of its principal advantages (sampling). As the data of this study suggest, the problem with general damages is not who makes the decision; rather, the problem is the inherent difficulty of the decision being made.

E. *Postscript*

The present study leaves open several important questions. Although there is a strong underlying similarity of judgments among groups, especially concerning the nature of injuries, how might the models of award decisions be improved? Relatedly, what accounts for any differences that do exist among the models of the groups studied? Why, exactly, do defense lawyers’ decision models depart from those of the other groups? Why, exactly, are jurors more variable than the others, even though their decision models are so similar?¹⁹²

The research described in this Article presented jurors, judges, and lawyers with bare-bones cases, consisting almost exclusively of facts about the injuries and their sequella. Understanding how decision-

188. See Saks et al., *supra* note 27 (reporting a jury simulation experiment showing that giving mock jurors guidance tended to reduce the variability of general damages awards for medium- and high-severity injuries).

189. See Baldus et al., *supra* note 26.

190. For details of a methodology for gathering and using such data in courts, see Baldus et al., *supra* note 26. See also Bovbjerg et al., *Valuing Life and Limb*, *supra* note 26.

191. Others have discussed these and additional techniques for dealing with this problem. See ABA REPORT OF THE ACTION COMMISSION TO IMPROVE THE TORT LIABILITY SYSTEM 10-15 (1987); ALI REPORTERS’ STUDY ON ENTERPRISE LIABILITY FOR PERSONAL INJURY, 199-230 (1991); Bovbjerg et al., *Valuing Life and Limb*, *supra* note 26; Diamond et al., *supra* note 146; Mark Geistfeld, *Placing a Price on Pain and Suffering: A Method for Helping Juries Determine Tort Damages for Nonmonetary Injuries*, 83 CAL. L. REV. 775 (1995); Sunstein et al., *supra* note 51.

192. The obvious hypothesis that the latter finding is due to a lack of experience with the task is only that, a hypothesis.

makers respond to the core attributes of injuries is only a first step. The models we have obtained can be built upon, layer by layer, systematically adding legally permissible contextual factors that should influence perceptions of the impact of injuries and the consequent damages. Then extra-legal contextual factors, including characteristics of cases, plaintiffs, and defendants, can be added. Although these ought not to affect damages awards, some of them are likely to, and identifying them can add to the predictive accuracy of the model. Finally, we can experiment with procedural and evidentiary variations that may affect the uses jurors (and judges and lawyers) make of the information they receive, and which may increase the predictability and decrease the variability of awards. Thus, the models developed in the present study, which already embody a considerable amount of predictive power, seem likely to be amenable to improvement through additional stages of model development.

APPENDIX A:
EXAMPLES OF INJURY DESCRIPTIONS

The two bones in the plaintiff's lower right leg were badly broken in several places. The 33-year-old plaintiff's leg was so disrupted that repair of the blood vessels was impossible and surgical amputation of the leg just below the knee was necessary. Over parts of three years, the plaintiff was hospitalized on many occasions and underwent several painful operations. The plaintiff was fitted with an artificial leg. Because the artificial leg causes the leg stump to become raw and inflamed, the plaintiff can walk for only brief periods.

The 33-year-old plaintiff suffered numerous cuts on the left forearm. The cuts required 40 stitches. The injury left several small scars that are permanent.

The 37-year-old plaintiff suffered scarring of the lungs, which affects their ability to absorb oxygen. Also, the lining of the lungs has hardened, making it difficult for them to move air in and out. These conditions make the plaintiff short of breath, more susceptible to flu and pneumonia, and more likely to develop lung cancer. The plaintiff cannot walk very far without getting winded and becoming exhausted. The plaintiff will continue to get worse and has a 50-50 chance of dying within five years.

The 33-year-old plaintiff suffered burns that caused blistering and destroyed tissue on the face and over 80% of the body. The plaintiff's wounds and their treatment were extremely painful. During the first two months in the hospital, the plaintiff underwent massive infusions of blood plasma and a procedure of salves, bandaging, chlorine baths, and cutting away the dead tissue from the burned areas. Over several years, the plaintiff had extensive skin graft surgery and received further treatment. The plaintiff suffered from painful tightening of the scar tissue, which impaired the mobility of joints and required constant surgical repairs. All of this required 70 operations, and at least ten more operations will be necessary over the next ten years. The plaintiff's skin resembles a patchwork quilt, and extremes of heat and cold will always be painful.

The 33-year-old plaintiff suffered a broken right wrist. The plaintiff, who is right-handed, had to wear a cast for one month. For a few months, the plaintiff's wrist was somewhat weak and less flexible. Thereafter, the wrist healed completely.

The 33-year-old plaintiff suffered a broken pelvis, such that the bone was splintered into fragments. In addition, the fragments on the left side were out of position and surgery was required to set them. The plaintiff spent one month in the hospital, followed by another month of bed rest at home. Over the next several months, the plaintiff gradually was able to walk and to resume normal activities.

APPENDIX B:
INTERVIEW PROTOCOL FOR JURY-ELIGIBLE CITIZENS¹⁹³

Introduction to the interview:

We are studying people's decisions about how much monetary compensation an injured person should receive in a lawsuit. We want to talk to people like you who could end up serving as jurors and making these kinds of decisions.

The interview will last about ten minutes. I will read you brief summaries of two real cases and then ask you for your reactions to them. The cases may contain some graphic details about the injuries.

These cases involve an injured person, called the plaintiff, and the party responsible for the injury, called the defendant. According to the law, when the defendant is found to be responsible, plaintiffs are entitled to receive compensation for their medical expenses, lost wages, and the pain and suffering they have experienced.

In each case that I'm going to read you, a jury in an earlier trial found the defendant to be legally responsible for the plaintiff's injury. Also at that trial, the jury compensated the plaintiff for medical expenses and lost wages. So all you have to focus on here is how much pain and suffering this injury has caused the plaintiff and how much money would provide fair compensation for that pain and suffering.

So for each case, I am going to ask you to come up with a dollar amount for the plaintiff's pain and suffering and to give me your impressions of the injury. For example, to rate how much physical pain the injury has caused. If I ask any question that you do not wish to answer, let me know and we'll move on to the next question. All your answers will be kept confidential and will be seen only by the researchers directly involved with this project.

The first case was then read, followed by these instructions and questions:

To remind you, a jury in an earlier trial found the defendant in this case to be legally responsible for the plaintiff's injury or illness. Also in that trial, the jury compensated the plaintiff for medical expenses and lost wages. So leave those out of your calculation.

193. The introduction to the interview for lawyers and judges varied somewhat from the introduction for jurors. The jurors' introduction had more information about the study (since they had not received a letter about the study in advance of the phone call) and more explanation of personal injury cases. The same instructions for determining an award were read to all respondent groups within each state. The questions about the awards and the injuries were asked of all participants in the same order. The particular socio-demographic items varied with the respondent group; the questions asked of each group are included here.

New York:

You should focus only on compensating the plaintiff for the injury and for any conscious pain and suffering that the plaintiff experienced in the past, or will experience in the future with reasonable certainty. In deciding the amount to be awarded for the injury and for conscious pain and suffering, you may take into consideration any physical pain, mental suffering, disability, disfigurement, or loss of the ability to enjoy life. There is no exact standard for fixing the compensation to be awarded. Any award that you make should be fair and reasonable in light of the facts you have heard. Disregard any caps on pain and suffering in your state and give the award you think the injury should receive without regard to any cap.

Illinois:

You should focus only on compensating the plaintiff for any physical pain, mental suffering, disability, or disfigurement that the plaintiff experienced in the past, or will experience in the future with reasonable certainty, as a result of the injury. In deciding the amount to be awarded, you may consider the nature, extent, and duration of the injury. There is no exact standard for fixing the compensation to be awarded. Any award that you make should be fair and reasonable in light of the facts you have heard. Disregard any caps on pain and suffering in your state and give the award you think the injury should receive without regard to any cap.¹⁹⁴

With all this in mind, how much money would you award to the plaintiff?

[If the respondent answered zero]

Why do you think the plaintiff should receive no compensation?

If the respondent said, in effect, the defendant was "not guilty" or "not liable," the interviewer replied: Let me remind you that a jury in an earlier trial found the defendant in this case to be legally responsible for the plaintiff's injury or illness.

With this in mind, how much money would you award to the plaintiff?

How much money do you think the average juror would award this plaintiff?

Now I'm going to ask you five questions about your impressions of the injuries the plaintiff suffered. For each question, I'd like you to indicate your judgment on a five-point scale.

194. The last sentence in the instructions in both states was included for only the judges and lawyers, who would tend to be aware of existing caps while the jurors would not.

1. Considering both the amount of physical pain and how long it will last, how much physical pain will this injury have caused the plaintiff? From one, no pain at all, to five extreme pain?
2. Considering both the amount of mental suffering (that is, worry, fear, depression, and so forth) and how long it will last, how much mental suffering will this injury have caused the plaintiff? From one, no mental suffering at all, to five, extreme mental suffering?
3. Considering both the amount of disability and how long it will last, how much disability will this injury have caused the plaintiff? From one, no disability at all, to five, extreme disability?
4. Considering both the amount of disfigurement and how long it will last, how much disfigurement will this injury have caused the plaintiff? From one, no disfigurement at all, to five, extreme disfigurement?
5. Overall, considering all aspects, how severe is the injury to the plaintiff? From one, not at all severe, to five, extremely severe?

The second case was then read, followed again by the award instructions. The same award and injury questions as above were asked for the second case.

After the awards and injury ratings were made for the second case, the following socio-demographic information was obtained.

JURORS

In order to be sure that we have a fair representation of all kinds of people, we'd like to ask you the following questions.

Do you personally know anyone who has had either of the conditions I have read to you about? IF YES: which one?

How would you describe where you live? Would you say it is: a large city; the suburbs of a large city; a medium-sized city; or a small town or rural area?

What is your age?

What is the highest level of education you have completed?

0 - 8th grade

9 - 11th grade

high school graduate

some college

college graduate

post-college study or degree

Please tell me which of the following ranges best describes your total annual household income.

- less than \$15,000
- \$15,000 - \$30,000
- \$31,000 - \$50,000
- \$51,000 - \$75,000
- over \$75,000

Thank you for your participation in this interview. The answers you have provided will assist the researchers in their study of legal decision making. Thank you very much for your time.

The following information also was recorded but was not asked of the participants:

gender: male, female

state: Illinois, New York

from urban or rural county group: urban, rural

JUDGES

About what percentage of the cases that come before you are personal injury cases?

Do the people who serve as jurors in your court come mostly: from a large city; from the suburbs of a large city; from a medium-sized city; or from a small town or rural area?

Do you personally know anyone who has had either of the conditions I have read to you about, or have you heard a case involving either of these conditions? IF YES: which one?

For approximately how many years have you been hearing personal injury cases as a judge?

That concludes the interview. Would you like to receive a summary of the results of the study when it is completed?

Thank you for your participation in this interview. The answers you have provided will assist the researchers in their study of damages decision-making. Thank you very much for your time.

The following information also was recorded but was not asked of the participants:

gender: male, female

state: Illinois, New York

from urban or rural county group: urban, rural

LAWYERS

What percentage of your work involves representing clients in personal injury cases?

In personal injury cases, do you represent: mostly plaintiffs, mostly defendants, or both about equally?

In the personal injury cases that you have worked on, do the plaintiffs come mostly: from a large city; from the suburbs of a large city; from a medium-sized city; or from a small town or rural area?

Do you personally know anyone who has had either of the conditions I have read to you about, or have you been involved in a case in which the plaintiff had either of these conditions? IF YES: which one?

Have you served as an arbitrator in a personal injury case?

IF YES: Have you done that: rarely, occasionally, or often?

For approximately how many years have you been handling personal injury cases?

That concludes the interview. Would you like to receive a summary of the results of the study when it is completed?

Thank you for your participation in this interview. The answers you have provided will assist the researchers in their study of damages decision-making. Thank you very much for your time.

The following information also was recorded but was not asked of the participants:

gender: male, female

state: Illinois, New York

from urban or rural county group: urban, rural

from plaintiff or defense group: plaintiff, defense

APPENDIX C:
SUMMARY OF REGRESSION ANALYSIS CONCEPTS

Percentage of variance. The percentage of variance accounted for tells us how well a group of variables included in a predictive model accounts for what one seeks to predict, such as awards. Zero percent of the variance in awards accounted for by injury ratings and demographics would reflect a complete absence of predictive power; 100% of the variance accounted for would reflect perfect predictive power.

Statistical significance of the predictor variables in the model. Only those variables that have a statistically significant relationship to the variable being predicted make a contribution to the predictive model. Thus, variables that are not statistically significant are omitted from the model.

Interactions. Sometimes the relationship of a predictor variable to the variable being predicted is affected by another predictor variable. In such circumstances one can add to the predictive power of a model by including these statistically significant "interactions" between predictor variables in the model. For example, we found that disability interacted with state, such that disability had a greater impact on awards for New York judges than for Illinois judges.

Direction of effect. The main effects (variables acting individually) and interaction effects (variables acting jointly) can predict either increases or decreases in the variable being predicted.

Magnitude of effect. Some predictors are more powerful than others. The magnitude of an effect is typically reflected in a "regression coefficient" (such as a Beta weight), and these are presented in a table of the regression results. One of the most intuitively meaningful ways to understand the magnitude of effect of a predictor is simply to compare it to other predictors.

Shared variance and unique variance. Sometimes different predictors are correlated with each other (that is, they "share variance"), so that there is redundancy in their predictive power. The amount of variance attributable to a given predictor and that predictor alone ("unique variance") is determined statistically by regression analysis by removing the redundancy. In order to compare the extent to which a given predictor's predictive power is a result of its shared versus unique influence, we compare the magnitude of its uncontrolled effect on the variable being predicted (given by a simple zero-order correlation, bivariate r , in which no other variables are controlled) with the magnitude of its effect on the variable being predicted in the regression model (given by Beta, in which the effect of other variables has been removed).

The reader who wishes to understand these statistical techniques even more deeply might consult one of the many textbooks on the

topic, or, for a treatment designed for judges and lawyers, read Daniel L. Rubinfeld, *Guide to Multiple Regression*, § 4, in MODERN SCIENTIFIC EVIDENCE: THE LAW AND SCIENCE OF EXPERT TESTIMONY (David L. Faigman, David H. Kaye, Michael J. Saks & Joseph Sanders eds., 1997).