

April 2003

Can Tort Juries Punish Competently?

Neal R. Feigenson

Follow this and additional works at: <https://scholarship.kentlaw.iit.edu/cklawreview>

 Part of the [Law Commons](#)

Recommended Citation

Neal R. Feigenson, *Can Tort Juries Punish Competently?*, 78 Chi.-Kent L. Rev. 239 (2003).

Available at: <https://scholarship.kentlaw.iit.edu/cklawreview/vol78/iss1/9>

This Book Review is brought to you for free and open access by Scholarly Commons @ IIT Chicago-Kent College of Law. It has been accepted for inclusion in Chicago-Kent Law Review by an authorized editor of Scholarly Commons @ IIT Chicago-Kent College of Law. For more information, please contact dginsberg@kentlaw.iit.edu.

CAN TORT JURIES PUNISH COMPETENTLY?

Punitive Damages: How Juries Decide. By Cass R. Sunstein, Reid Hastie, John W. Payne, David A. Schkade, and W. Kip Viscusi.
University of Chicago Press. 2002. Pp. 285. \$35.00.

NEAL R. FEIGENSON*

INTRODUCTION

Jurors award punitive damages too often. The amounts they award are erratic and unpredictable, even though they start from shared moral intuitions about the reprehensibility of defendants' conduct. And when they deliberate and decide as juries, both the size and variability of the awards increase. These are among the main findings from some twenty mock juror and jury experiments reported in *Punitive Damages*, by Cass Sunstein, Reid Hastie, John Payne, David Schkade, and W. Kip Viscusi. Equally important are the reasons why jurors award punitives as they do. According to the authors' studies, jurors frequently award punitives when experts (*i.e.*, judges) would not because jurors tend to ignore the jury instructions that specify the requirements for punitive liability. Jurors also exhibit systematic biases in their evaluations of the *ex ante* riskiness of defendants' conduct, often blaming defendants for behavior that was cost-justified and thus, under the prominent deterrence theory of tort liability, not even culpable, much less reprehensible. And lacking any clear guidance as to how they should translate their urge to punish into a specific dollar award, jurors resort to various flawed judgmental habits to reach unpredictable and arbitrary outcomes.

* Professor of Law, Quinnipiac University School of Law, and Research Affiliate, Yale University Department of Psychology. I would like to thank Steve Gilles, Steve Latham, Linda Meyer, Tony Sebok, Richard Sherwin, and Christina Spiesel for their helpful comments on an earlier draft of this Article, and Dan Bailis and Jai Park for their guidance on the statistical issue discussed *infra* on pages 253–54. I should disclose that about three years ago Cass Sunstein, the lead author of the book under review, kindly agreed to read a prepublication copy of my own book and to write a jacket blurb for it.

The authors assert that their findings provide support for reforms in the punitive damages process that would increase predictability and avoid outlier judgments: more modestly, having judges determine awards (or exercise greater oversight of jury awards) using awards in similar cases as a standard; more ambitiously, replacing jury and judge decision-making entirely with a quasi-administrative schedule of awards based on guidelines developed by specialists in the relevant subject matter areas. Given the authors' distinction and outstanding achievement in their respective fields of scholarship, their recommendations are likely to be taken seriously by judicial and/or legislative policy makers,¹ so it is worth considering their work very carefully.

Punitive damages, and in particular, allegations that juries' punitive awards are too frequent, too large, and too unpredictable, have prompted much academic and political debate during the last twenty years. The positions of many of the contributors to this debate, including some of the authors of the book being reviewed, are well known to those who follow these matters.² New data, as social psychologists have shown, may only further entrench the views of those firmly on one side of a debate or the other,³ but experimental research of the kind the authors have conducted can help everyone else by pinpointing the effects of specific variables (such as case features or jury instructions) whose impact is difficult to isolate in the complexity of real-world trials and verdicts.

1. For instance, many of the studies reported in the book have been brought to the attention of the Supreme Court in a case on punitive damages being decided this term. Brief of Certain Leading Business Corporations as Amici Curiae in Support of Petitioner, *State Farm Mut. Auto. Ins. Co. v. Campbell*, 122 S. Ct. 2326 (2002) (No. 01-1289). Cf. Neil Vidmar, *Juries Don't Make Legal Decisions! And Other Problems: A Critique of Hastie et al. on Punitive Damages*, 23 LAW & HUM. BEHAV. 705, 713 (1999).

2. For instance, W. Kip Viscusi has published a paper titled *Why There Is No Defense of Punitive Damages*, 87 GEO. L.J. 381 (1998) (cited in PUNITIVE DAMAGES at p. 246 n.7) (this and all subsequent page references to PUNITIVE DAMAGES are in parentheses in the body of the text and footnotes), which was a response to criticisms by Theodore Eisenberg and David Luban of another article by Viscusi in the same issue of that journal. In addition, the empirical studies collected in PUNITIVE DAMAGES have previously been published in some ten articles (pp. x-xi). Based on these publications if not also the partial sponsorship of the authors' research by the ExxonMobil Corporation (p. ix), some readers of the book might anticipate the authors' critical view of juries' punitive damages decision-making. I refuse to prejudice the book in this fashion, and will evaluate the authors' research and arguments on their own merits.

3. The classic study is Charles G. Lord et al., *Biased Assimilation and Attitude Polarization: The Effects of Prior Theories on Subsequently Considered Evidence*, 37 J. PERSONALITY & SOC. PSYCHOL. 2098 (1979). For a more recent discussion, see Robert J. MacCoun, *Biases in the Interpretation and Use of Research Results*, 49 ANN. REV. OF PSYCHOL. 259, 265-67 (1998).

Taken as a whole, the research reported in *Punitive Damages* represents a relatively comprehensive effort to describe with precision not just the outcomes of punitive damages assessments but also, at least as importantly, the cognitive and social psychological processes that lead to those outcomes. The range of phenomena studied is impressive. Some of the authors' findings, such as that jurors are prone to rely on arbitrary numerical anchors in setting dollar figures (chapter 4), are unsurprising, confirming the results of many previous studies. Other findings—for instance, that deliberations increase both the size of punitive awards (the “severity shift”) and their variability (chapter 3); that the less attention juries pay to the judges' instructions, the more likely they are to award punitives (chapter 5); and that jurors are more, not less, likely to punish a defendant that they know performed a competent cost-benefit analysis before acting (chapter 7)—are striking contributions to the literature. Moreover, the concatenation of these findings to form an intelligible account of how jurors arrive at their punitive awards is a valuable addition to our knowledge of the topic. The authors' research in the aggregate involved thousands of jury-eligible participants and, in the group decision-making studies, hundreds of mock juries, enhancing the reliability of the findings. The studies in general offer important results that should be taken into account in subsequent debates about punitive damages.

How the results should be interpreted is a somewhat different matter. The authors⁴ use their research to argue that the arbitrariness and unpredictability of juries' punitive damages awards constitute a “serious problem” for the rule of law (p. 251), warranting corrective measures. There is nothing wrong with using social scientific research data to make an argument; it is certainly preferable to arguing without data. Difficulties arise only to the extent that the argument drives the interpretation of the data rather than the other way around. The authors of *Punitive Damages* state at the outset that “our principal goal is descriptive” (p. vii), and for the most part the

4. It may be perceived as misleading to generalize about “the authors” because different persons are credited as authors of different chapters, including not only the five named authors but also psychologist Daniel Kahneman, who collaborated on the work reported in three of the chapters. For instance, deviations from the ideal of descriptive objectivity may vary by author and/or chapter; in addition to the specific examples mentioned later in this Review, it seems to me that Reid Hastie is the most temperate of the principal authors and Kip Viscusi the least (although I have not conducted a systematic content analysis to test this impression). With this caveat, I will proceed to treat “the authors” (or “the researchers” or “the experimenters”) as a single entity, except where otherwise noted.

ten central chapters of the book in which they report their findings fulfill this goal with the objectivity one expects from leading social scientists. Yet at many points in their presentation, the data do not support their critical view of punitive damages juries as strongly as they would have readers believe.⁵

Any study that sets out to determine whether decision-making processes or outcomes are flawed necessarily employs one or more criteria or norms of sound judgment.⁶ The authors of *Punitive Damages* use four types of criteria to evaluate jurors' and juries' punitive damages judgments. They variously compare participants' reasoning and decisions to (1) what a judge would have decided (what may be called the "judicial" norm); (2) what the relevant legal rules appear to require (the "legal" norm); (3) the outcome dictated by the consequentialist theory of optimal deterrence (the "economic" norm)⁷; and (4) the standards of rational decision-making commonly employed in cognitive and social psychological research (the "rational judgment" norm).⁸ So, for instance, in arguing that jurors and juries vote for punitives more often than they should, the authors use the judicial and legal norms⁹ (chapter 5) as well as the economic norm

5. It may very well be the case, as Phoebe Ellsworth has written, that "[a]lmost every author [of an empirical study] exaggerates the significance of the problem studied, overstates the implications of the results, or both." Phoebe C. Ellsworth, *Sticks and Stones*, 23 LAW & HUM. BEHAV. 719, 720 (1999). Even if this were the only criticism to be made of PUNITIVE DAMAGES, it would still be worth informing readers as precisely as possible of discrepancies between data and interpretation. As I will argue, however, the authors do more than overstate the importance or implications of their findings; in a number of instances they make disputable claims for their results seemingly for the purpose of constructing a slanted picture of jury behavior in punitive damages cases.

6. See MacCoun, *supra* note 3, at 264–65 (referring to "logics" used to determine bias or error).

7. The idea underlying optimal deterrence is that society benefits when it properly balances productive (yet necessarily risky) activity and safety, and benefits most when it spends resources to deter the harmful consequences of risky activity only up to the point at which precautions (including productive activity curtailed) cost as much as the gains in safety that the precautions achieve. Thus, "the task of the legal system [should be] to create penalties that are high enough to produce adequate deterrence [of risky activity], but not so high as to produce overdeterrence" (pp. 109–10).

8. The "rational judgment" norm may be easier to define negatively than positively. It encompasses the ability to make decisions while avoiding framing effects, availability and hindsight biases, the "fundamental attribution error," and other habits of perception and judgment that social scientists would consider errors if present in their own collections, analyses, or reports of data. For a critical discussion of this norm, see David C. Funder, *Errors and Mistakes: Evaluating the Accuracy of Social Judgment*, 101 PSYCHOL. BULL. 75 (1987).

9. In some circumstances the judicial and legal norms may be conflated, namely, if it is assumed that judges follow the relevant law. Nevertheless, the two norms remain analytically distinct: the judicial norm entails a comparison across decision-makers (e.g., jurors' versus judges' verdict preferences), while the legal norm, like the norms of optimal deterrence and rational judgment, entails a comparison of the target decision to an extrinsic standard or

(chapter 10). In characterizing punitive damages awards as arbitrary or irrational, they tend to emphasize the rational judgment norm, although the other norms play important roles here, too.¹⁰

When the authors (or any researchers) evaluate their findings in light of their criteria—for instance, when they identify a discrepancy between juror performance and norm as “bias” or “error”—we should ask at least the following (interrelated) questions: First, have the researchers indeed identified a discrepancy? That is, are the results pointing to the discrepancy significant, reliable, and valid? Second, what should be made of the discrepancy? That is, does it really indicate a shortcoming in the decision-maker’s thinking or decisions, or is there some other explanation for the discrepancy? And if there is bias or error, how important is it as a practical or theoretical matter?¹¹ Third, have the researchers adequately justified their choice and implementation of the particular criterion as the basis for a normative evaluation of the decision-maker’s performance in this context?

Punitive Damages fares very well with regard to the first question. When the authors claim to have identified a discrepancy between juror or jury performance and a clearly specified normative criterion, the discrepancy seems real. What the authors make of these discrepancies, however (the second question), is occasionally problematic. Many of their accounts of what they have found are entirely convincing (as I will observe throughout this Review). But in various ways, generally small in themselves but cumulatively hard to ignore—sometimes in the design of the experiments, more frequently in the language employed to characterize their findings and in the selective use (or, more often, neglect) of other empirical data to provide a context for their findings—the authors’ normative position leads them to present their results tendentiously. I will suggest alternative interpretations of the findings of several of the studies. Among other things, I will indicate where readers’ appreciation of the

criterion. See MacCoun, *supra* note 3, at 264–65. Of course, jurors’ and judges’ decisions may be compared both to each other and to an extrinsic criterion, as the authors do in the research reported in chapter 11 (evaluating the judgments of both jury-eligible participants and real judges in light of the rational judgment norm and other criteria).

10. It is more difficult to pin down the standard on which their crucial description of punitive damages awards as highly variable and unpredictable is based, and there really is no reliable criterion for judging that punitive awards are excessive. I will return to both points later in this Review.

11. See Robert J. MacCoun, *Epistemological Dilemmas in the Assessment of Legal Decision Making*, 23 *LAW & HUM. BEHAV.* 723 (1999).

authors' research could be furthered by additional context in the form of empirical data about real world punitive damages (which would provide a better estimate of the scope of the "problem") and more detailed information about the experimental scenarios used as stimulus materials (which might allow readers to identify reasons for participants' conduct that do not fit as neatly into the authors' argument).

This brings us to the third question. The criteria against which the authors measure juries' punitive damages judgments—norms of good decision-making supplied by judicial outcomes, legal rules, social scientific rationality, and economic theory—are all (with the possible exception of the last) widely employed in jury research. Each can be appropriate; each also has its limitations. For instance, judges can disagree with one another and legal rules can sometimes be reasonably interpreted in conflicting ways, undermining the reliability of the judicial and legal norms, respectively. Moreover, because juries may legitimately bring to the justice system values that judges may not adequately express, and because those different values may sometimes lead juries to decide cases differently than judges would, it would be odd to rely too heavily on the judicial norm (or the legal norm as construed by judges) to evaluate juries' decisions.¹²

Of greater concern is the norm that is most critical to the book. The authors' conception of how big a problem juries' punitive damages awards pose, and thus how badly needed are reforms that shift authority for making the punitive damages decision away from juries, is ultimately grounded in a notion of punitive damages based on the goal of optimal deterrence, and the high levels of consistency and predictability in outcomes that efficient deterrence demands. Of course consistency and predictability are virtues in a legal system, and optimal deterrence is a worthwhile objective, all things being equal.

12. As for the rational judgment norm, it could be argued that it is based on an impoverished notion of how humans think and act. Many narrative theorists and cultural anthropologists, for instance, would contend that human judgment can be properly understood and evaluated only against the background of deep, pervasive cultural patterns and influences that are difficult if not impossible to operationalize in an experimental context. It has also been argued that departures from the rational judgment norm as measured in the laboratory are sometimes not judgmental flaws, and may even be strengths, when evaluated in terms of broader, real world performance standards, or plausible, alternative normative conceptions of social judgment. See Funder, *supra* note 8; Philip E. Tetlock, *Social Functionalist Frameworks for Judgment and Choice: Intuitive Politicians, Theologians, and Prosecutors*, 109 PSYCHOL. REV. 451 (2002). Neither of these observations, however, poses serious difficulties for the authors' use of the rational judgment norm in the research reported in the book.

The authors' focus is troubling, however, because it completely ignores another, undeniably important goal that punitive damages should serve, one more closely allied to corrective justice than to deterrence—retribution for wrongs perceived to be reprehensible. I will conclude the Review by discussing whether the authors have adequately explained and justified the normative vision of punitive damages that underlies much of their research program and provides the standards for their most severe criticisms of jury behavior.

I. ARE JURORS TOO INCLINED TO AWARD PUNITIVE DAMAGES?

Two distinct experiments lead the authors to conclude that jurors and juries award punitive damages more frequently than they should. In one (chapter 5), mock juries were presented with stimulus materials based on four actual cases in which courts ultimately ruled that punitives should not be awarded. Across all cases, 58% of the mock juries voted to award punitives; excluding hung verdicts, 67% voted for punitives. In the second (chapter 10, experiment 1), individual participants read one of four versions of an airplane defect story; in all versions the cost of the repair exceeded the expected accident costs, and so the defendant airline should not even have been found negligent (under the standard cost-benefit definition of reasonable care¹³), much less reckless and thus possibly liable for punitives. In this experiment, 84% of participants voted to award punitives. Thus, using two different benchmarks for gauging the appropriateness of punitives—in the first experiment, the judicial norm as indicated by the actual final decisions of trial or appellate courts; in the second, the economic norm as reflected in a correct alignment of perceived culpability and cost-benefit analysis—a majority of participants can be said to have “got it wrong” when they decided to award punitives. The authors of the first study (Hastie, Schkade, and Payne) write that “the number of juries that rendered verdicts discrepant from those reached on appeal is perhaps disturbing” (p. 84).

These authors are careful to disclaim reliance on the courts' resolutions as a test of the correctness of their juries' decisions: “[O]ne

13. Often attributed to Learned Hand's opinion in *United States v. Carroll Towing*, 159 F.2d 169 (2d Cir. 1947), in which a cost-benefit calculus was used to define negligence: where B equals the burden or cost of taking precautions, P equals the probability that an accident will occur if the precautions are not taken, and L equals the severity of the loss should the accident occur, an injurer is negligent if it failed to take those precautions where $B < P \times L$. See also *infra* note 68 (discussing cost side of cost-benefit equation).

could argue that the higher courts were wrong, and punitive damages were warranted in those cases. Our focus, however, is not so much on the right or wrong of the decision as it is on the extent to which juror reasoning considered the necessary conditions for the verdicts rendered” (p. 80). While the latter set of findings may indeed be the “central” ones produced by the study¹⁴ and will be taken up later in this Review, the authors do seem to intend that readers infer from the discrepancies between experimental and actual outcomes that the mock juries that voted for punitives erred. They write that they chose the four cases “because they are frequently cited as precedents, and because the *proper action* on the issue of punitive damages had been decided as a matter of law by trial or appellate court review of the original proceedings” (p. 80) (emphasis added), implying that the actual cases create a reliable benchmark after all. More directly, they write that “discussion of issues that were legally relevant reduced the tendency to decide, *incorrectly*, that punitive damages were warranted” (p. 90) (emphasis added).

Yet the design of this experiment is arguably biased toward producing the impression that jurors tend to award punitives when judges would not.¹⁵ First, while the authors admit that one or more of the appellate courts in the actual cases may have gotten it wrong when they ruled that punitives were inappropriate as a matter of law, they do not make it clear, as they might easily have done, that in two of the four cases the trial judges firmly believed that punitives were appropriate, and the appellate courts went to some length to explain why they thought punitives should not have been allowed. This at least suggests that the ultimate outcomes in the actual cases may not be as reliable a criterion of the incorrectness of awarding punitives as the authors imply.¹⁶

This suggestion may be underscored by considering one of the stimulus cases, which the authors helpfully reproduce (pp. 93–95). (Indeed, only here and in chapter 9 do the authors provide this kind of detail about their experimental materials.) I encourage the reader

14. Reid Hastie et al., *Reply to Vidmar*, 23 LAW & HUM. BEHAV. 715, 715 (1999).

15. The original publication of this study prompted an adversarial forum in LAW AND HUMAN BEHAVIOR in which Neil Vidmar criticized the study, its authors (Hastie, Schkade, and Payne) responded, and other prominent jury researchers (Phoebe Ellsworth and Robert MacCoun) weighed in. Ellsworth, *supra* note 5; Hastie et al., *supra* note 14; MacCoun, *supra* note 11; Vidmar, *supra* note 1. To their credit, the authors refer readers to this forum (p. 223 n.17). Two of the major criticisms in this portion of the Review are taken from Vidmar’s paper.

16. Vidmar, *supra* note 1, at 707–10.

to attempt the task assigned to the participants in the experiment: to read this scenario and then, applying the judge's instructions on punitives (reproduced in Appendix A, pp. 259–60), decide whether punitives should be awarded. When I do so, it strikes me that it would at least be reasonable to conclude that all four elements of the legal definition of “recklessness” are satisfied, in which case, punitives could reasonably be awarded—and the appellate court's determination that punitives were inappropriate as a matter of law would be not merely disputable, but incorrect.¹⁷

In addition, by choosing as their stimulus materials four cases in which the “correct” decision was not to award punitives, the researchers allowed experimental participants to differ from the norm in only one direction—by voting for punitives when they should not have. This is the result that happens to coincide with the depiction of punitive damages jurors as antidefendant (p. 233) and anticorporate (pp. 113–14) presented elsewhere in the book. Had the researchers included scenarios in which the “correct” decision was to award punitives, then discrepancies between participant and actual outcomes might (also) have indicated that juries are generally incompetent to decide punitives,¹⁸ or perhaps that they display biases against as well as in favor of punitives.¹⁹ Since these patterns could not emerge from the experiment as designed, it can be argued that the study was biased toward the outcome indicated by other experiments reported in the book.

The author of the second experiment (Viscusi) purporting to show that jurors are too inclined to vote for punitives interprets his results—84% of participants awarded punitives when, according to the economic criterion of good decision-making, *none* should have—essentially as follows: The defendant airline declined to fix a cargo door on one of its planes at a cost of \$2,000. Because reliable estimates placed the risk of harm at 1 in 10 (in other versions of the

17. Of course, the scenario given to the participants differed in many respects (even if not in the “core facts” represented) from the evidence before the trial court and the body of information contained in the record on appeal, so I cannot confidently say whether I would agree with the actual trial or appellate court's decisions. But this simply reinforces the basic criticism of the experimental design, or at least the significance the authors draw from it.

18. Ellsworth, *supra* note 5, at 721.

19. Vidmar, *supra* note 1, at 710–11. Note that a majority of participants seem to have voted not to award punitives in several of the scenarios used in the experiments reported in chapters 2 (p. 38 tbl. 2.2) and 3 (p. 55 tbl. 3.5), which might temper the view of trigger-happy punitive juries presented in chapter 5. (In the absence of more detailed accounts of those scenarios, however, it is impossible to go beyond this rather superficial speculation.)

scenario, 1 in 1,000, 1 in 100,000, or 1 in 1,000,000, respectively) and the loss in case of accident at \$15,000 in property (or \$1.5 million property loss, \$150 million personal injury loss (29 deaths), \$1.5 billion (290 deaths), respectively), the cost of taking precautions (\$2,000) exceeded the expected benefit (in accident avoidance) those precautions would yield (\$1,500). (In Learned Hand terms, $B > P \times L$.) So the defendant did not act carelessly in not fixing the cargo door. If the defendant was not negligent, then the defendant cannot possibly have been reckless, because recklessness reflects a greater degree of culpability or blameworthiness than mere negligence. Since the defendant must be found reckless for punitives to be awarded, punitives should not have been awarded. So 84% of the mock jurors “got it wrong.” This and other studies reported in the same chapter highlight jurors’ “[f]ailure to think sensibly about risk and risk-cost trade-off issues,” and more broadly show that “jurors particularly fall short of reasonable standards of behavior” (p. 185).

That jurors are not very good risk managers is one of the sturdiest conclusions to be drawn from Viscusi’s (and others’) research (as will be discussed in greater detail later in this Review). If everyone agreed that competent risk management—optimal deterrence—were the sole (or even a necessary) criterion for evaluating jury decision-making, the upshot of this research would be obvious: jurors often do get it wrong (and something should be done about it). The problem is, as Robert MacCoun asks: “If we had an external criterion for judging correct verdicts, why wouldn’t we just use it to resolve the case?”²⁰ The law does not simply adopt the economic (or, for that matter, the judicial) norm of decision-making, and one important reason why it does not is that the power to decide lawsuits is not confined to experts, whether judges or economists; the legal system makes room for lay decision-makers—jurors—and hence whatever substantive values those jurors may properly bring to bear on their judgments.

A discrepancy between jurors’ punitive damage decisions and those dictated by cost-benefit analysis, therefore, may indicate not juror “error” but rather a difference between jurors’ relevant values and those of risk analysis experts. For example, jurors may place greater weight on potential accident victims’ security (as against potential injurers’ liberty to engage in useful but necessarily risky

20. MacCoun, *supra* note 11, at 726.

activity) than the optimal deterrence standard does.²¹ Jurors may plausibly believe that so long as an accident is reasonably foreseeable, the defendant should be liable if it fails to do everything reasonably practicable to avoid it—that is, that the defendant should be considered negligent unless the burden of precaution was *disproportionate* to the risk, not merely slightly greater than the risk.²² In that event, it is not at all clear that participants should not have found (had they been asked) that the defendant in the experimental scenario was negligent, since the precautions (\$2,000) cost about the same as the expected accident costs (\$1,500), and could easily have been undertaken. And if the defendant was sufficiently blameworthy to be considered negligent, it may be less obvious that the defendant should not be judged to have acted recklessly, warranting punitive sanctions.

The data also indicate that the largest increment in the frequency of punitives occurred between the versions of the scenario in which only property was lost (74% to 78% voting for punitives) and those in which 29 or 290 people were killed (95% to 96% voting for punitives), reflecting the high (if not always economically justifiable) value jurors tend to place on risks to life. Moreover, in a group of about 90 judges who responded to the same accident scenarios (chapter 11, experiment 2), about 30% voted to award punitives, including 69% in the case in which 290 people died as a result of the defendant's failure to fix the cargo door. Thus, even though they "exercise[d] more restraint" than jurors did (pp. 196–97), nearly a third of the sample of sitting judges also violated basic economic efficiency principles in deciding to award punitives. This would be really remarkable if the standard for the correct outcomes in these cases were as obvious as the author seems to think it is. An alternative inference to be drawn from the widespread neglect of optimal deterrence by these judges is that something other than optimal deterrence may (sometimes) properly drive punitive damages awards. This is not to say that most or even many of the jury-eligible participants who voted for punitives in this experiment were in fact justified in doing so. My only point is

21. On liberty vs. security in tort law, see Jules Coleman & Arthur Ripstein, *Mischief and Misfortune*, 41 MCGILL L.J. 91 (1995). For an argument that negligence law is properly guided by a norm of reasonableness that is not grounded in economic rationality, see Gregory C. Keating, *Reasonableness and Rationality in Negligence Theory*, 48 STAN. L. REV. 311 (1996).

22. See, e.g., Mark Geistfeld, *Reconciling Cost-Benefit Analysis with the Principle that Safety Matters More than Money*, 76 N.Y.U. L. REV. 114 (2001). According to my colleague Steve Gilles, this is also an important conception of negligence in modern English tort law, although the precedents are far from clear. Stephen G. Gilles, *The Emergence of Cost-Benefit Balancing in English Negligence Law*, 77 CHI.-KENT L. REV. 489 (2002).

that while the authors' findings reliably identify important patterns in juror and jury decision-making, we need not join the authors in interpreting those findings as showing that jurors award punitives too often.

How often do real juries award punitives? The data vary by location and over time.²³ One comprehensive and widely respected study²⁴ of civil verdicts in 74 counties around the country for the period 1988–90 shows that punitives were awarded in 8.3% of verdicts that plaintiffs won and 4.5% of all civil cases tried to a verdict. Punitives were awarded in 8.6% of products liability cases that plaintiffs won and 2.5% of medical malpractice cases that plaintiffs won; in general, punitives were awarded at a much higher rate in cases involving financial harm than in those involving physical harm. Another study, by the RAND Institute for Civil Justice,²⁵ shows that in four geographically diverse counties between 1985 and 1994, punitive damages were awarded in 2% to 7% of civil cases tried to a verdict, with by far the greatest percentages of all punitive awards by case type coming in business cases (47%) and intentional tort cases (36%); product liability and medical malpractice cases each accounted for less than 5% of all punitive awards. These and other empirical data do not tell us, of course, whether real juries' punitive damages awards are too frequent, not frequent enough, or just right; but they do help to put the authors' experimental findings in context.²⁶

The real world data also suggest another limitation of the authors' experimental program, applicable to the variability and size of punitive damages awards as well as their frequency. Throughout the reported research, the scenarios used as stimulus materials are

23. For recent reviews of the data, see THOMAS H. KOENIG & MICHAEL L. RUSTAD, IN DEFENSE OF TORT LAW 141–43 (medical malpractice cases), 181–84 (products liability cases) (2001); Jennifer K. Robbennolt, *Determining Punitive Damages: Empirical Insights and Implications for Reform*, 50 BUFF. L. REV. 103, 161–63 (2002). For data in medical malpractice cases, see NEIL VIDMAR, MEDICAL MALPRACTICE AND THE AMERICAN JURY 254–55 (1995).

24. STEPHEN DANIELS & JOANNE MARTIN, CIVIL JURIES AND THE POLITICS OF REFORM 213–21 (1995).

25. ERIK MOLLER, TRENDS IN CIVIL JURY VERDICTS SINCE 1985 33–35 (1996).

26. In the introduction, George Priest observes with regard to similar data that “[t]he claim that [punitive damages] verdicts . . . are relatively infrequent must surely be true given the hundreds of thousands of civil claims filed and litigated each year in the United States” (p. 2). First, the data mentioned in the text measure punitives as a percentage of cases *tried to verdict*, not cases “filed and litigated,” which makes the single-digit percentages more telling. Second, neither Priest nor any of the book’s authors follow up by trying to relate their experimental findings to the verdict data in any way.

exclusively personal injury and environmental tort cases; the authors did not try to ascertain jurors' punitive responses to business and other intentional torts, which account for the greatest number of actual punitive awards. Obviously the authors cannot study everything about punitive damages, and by confining their research as they did, they not only economized on the creation of their materials but also allowed themselves to compare certain data across experiments. Still, their results leave open the possibility that jurors' punitive damages decisions may reflect *case type effects*: that is, that awards in business tort cases, for instance, may differ systematically in frequency, variability, or size (or in other ways) from those in product liability and/or environmental tort cases. This would be an interesting topic for further research. In addition, by emphasizing atypical punitive damages cases such as those involving product liability, the authors have made the goal of optimal deterrence seem more plausible and compelling than it might be in most actual cases, in which punitives (if any) are grounded in the defendant's intentional wrongdoing and cost-benefit analysis does not come into play.²⁷ The authors also (no doubt unintentionally) may have helped their findings conform to general readers' expectations, based on media-driven prototypes, that personal injury juries are especially inclined to award excessive damages, including punitive damages.²⁸

II. HOW BIG A PROBLEM IS THE VARIABILITY OF PUNITIVE DAMAGES AWARDS?

The unpredictability of punitive damages awards is perhaps the authors' major concern. The current system asks that lay decision-makers translate their (largely shared) moral feelings of outrage at a defendant's conduct into a precise dollar figure on an unbounded scale without any meaningful guidance. "Jurors are told nothing about typical awards in comparable cases, and they are given no help on the complex question of how to infer the punitive impact or deterrent efficacy of a dollar sanction" (p. 213). They are up a creek without a paddle, or, in the authors' formulation, doing "magnitude

27. I thank Tony Sebok for pointing this out to me.

28. See, e.g., 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 (1996); see also *infra* pp. 257-58 (discussing Priest's introduction); William Haltom & Michael W. McCann, *Law and Lore: Urban Legends and the Politics of Tort Reform* (1999) (unpublished paper, on file with author).

scaling without a modulus” (p. 41). The result: awards that are “distinctively unpredictable” (p. 212), “highly erratic” (p. 37), and “arbitrary” (p. 248), displaying “inconsistency and randomness” (p. 242). It is this inconsistency and unpredictability that the authors find most threatening to the rule of law (p. 248), and the need to contain it is ostensibly the strongest motivation behind their policy recommendations.

Two experiments (chapters 2 and 3) yield much of the data on which the authors rely for their finding that punitive damages awards are too variable for the good of the legal system. In the chapter 2 study, each participant read 10 accident scenarios,²⁹ rated the outrageousness of the defendant’s behavior (on a 0–6 scale), judged how much the defendant should be punished (same scale), and assessed a punitive award in dollars (no limit). The researchers found a remarkable degree of agreement among participants’ levels of outrage at and intent to punish the defendants in the various scenarios. Individuals’ rankings of the scenarios in terms of intent to punish correlated “in the .50 to .60 range” (p. 35). And when individual responses were aggregated by demographic groups, the correlations between the mean responses of different groups’ rankings of the cases were .99.³⁰ So jurors’ moral intuitions, their outrage and intent-to-punish judgments, seem to be widely shared and highly consistent.

29. Each of the 10 scenarios was presented in 2 or more conditions: in all 10, the size of the defendant firm (large vs. small) was manipulated, and in 4 of the 10, the severity of harm (high vs. low) was also manipulated, yielding 28 versions of the scenarios in all. The manipulations are not relevant to my discussion here.

30. As the authors point out, the .99 correlation indicates that “[j]udgments of intent to punish in these scenarios of personal injury cases evidently rest on a bedrock of moral intuitions that are broadly shared in society” (p. 35). But that the mean rankings by *groups* of respondents should correlate at .99, given the .5–.6 correlations among individuals, may not be that surprising. Given that demographic variables predict only 5% to 15% of the variance in verdict preferences generally (*see infra* note 34), grouping respondents by demographic category is likely to be nearly a random grouping for the present purposes as well, and the larger the (randomly chosen) groups, the higher the expected correlation between group means. What may be more striking are the .5–.6 correlations among individuals’ intent-to-punish rankings. Given a set of 10 various fact situations, ranging from air bag deployments to drugs with side effects to workplace toxic harms to inadequate security leading to a parking lot assault, for which punitive damages are arguably appropriate (and thus the defendants in all have behaved with some degree of culpability), these correlations between rankings seem quite high. They are much higher, for instance, than what I have found when I have asked my torts students to rank eight intentional infliction of emotional distress cases from their casebook in order of outrageousness. Of course, we can not determine whether we ought to be surprised at this degree of agreement without knowing more about the stimulus cases; unfortunately, the authors tell us nothing at all about them in this chapter, and provide only one-sentence blurbs about them (and five other cases) in chapter 3 (p. 47 tbl. 3.1). This is another instance in which the authors, by providing more context about their experimental data, could have allowed readers to evaluate the reported findings more thoroughly.

To determine whether dollar awards, by contrast, would be “erratic,” the researchers created “synthetic juries” by randomly sampling groups of 12 individual responses on all 3 dependent measures (outrage, intent to punish, dollar award) for all stimulus cases and computing the medians to create the synthetic jury’s judgments; they did the same for another group of 12 individual responses and then compared the median judgments.³¹ With regard to outrage and intent to punish, the median correlations between pairs of synthetic juries were in the .86–.89 range. In contrast, the median correlation between pairs of dollar awards was .42.³² The authors characterize this level of agreement as “quite weak” (p. 39), producing “severe unpredictability and highly erratic outcomes” (p. 37).

As the authors recognize, a key to the much greater variability of dollar awards is that outrage and intent-to-punish were measured on bounded, labeled scales (e.g., a minimum response of 0 = “not at all outrageous,” and a maximum of 6 = “absolutely outrageous” (p. 34)), whereas dollar awards were measured on an unbounded scale. No particular dollar amount reliably represents “absolutely outrageous” or any other relevant label. This is precisely what the authors mean by “magnitude scaling without a modulus” (p. 41).

But let us look more closely at that .42 correlation between pairs of synthetic juries’ dollar awards. The .42 is presumably a value for r , the usual correlation statistic.³³ Is an r of .42 “quite weak”? It is

31. More specifically, the researchers randomly sampled groups of 12 individual responses to each dependent measure (outrage, intent to punish, and dollar award) and computed the median judgment of that group, a “synthetic jury,” on that measure, for each of the 28 cases. They then randomly selected a different 12-person synthetic jury and computed its median response for each case. Then they correlated the 2 sets of 28 medians to determine how erratic (or consistent) the two juries are in evaluating the 28 cases. To enhance the reliability of their analyses, the researchers created 60 synthetic juries and computed the correlations between every possible pair both within measures (e.g., between two sets of 28 dollar awards) and between measures (e.g., between a set of 28 dollar awards and a set of 28 intent-to-punish ratings).

32. Outrage medians correlated with dollar award medians at .47; intent to punish medians correlated with dollar award medians at .51.

33. What the researchers are actually measuring here is *intercoder reliability*—the level of agreement between independent judges’ decisions (I thank Dan Bailis for pointing this out to me; cf. p. 91, where the authors explicitly use r as a measure of intercoder reliability of research assistants’ independent judgments of whether participants’ responses indicated a correct understanding of jury instructions). Unfortunately the authors of chapter 2 do not identify the statistic they are using (and in general do not report their results in standard APA format, which would facilitate the understanding and critique of their results). Where applicable, a better statistic for reporting intercoder reliability would be kappa, which measures not just the level of agreement between judges but the *difference* between the observed and expected levels of agreement. Jacob Cohen, *A Coefficient of Agreement for Nominal Scales*, 20 EDUC. & PSYCHOL. MEASUREMENT 37 (1960). It is unclear from chapter 2 whether the authors are

certainly much less than one of nearly .9; as best I can tell from the information the authors provide, however, it is significant at $p < .05$, the conventional test for statistical significance. An r of .42 yields an r -squared of .176, that is to say, knowing one synthetic jury's damages awards allows us to predict 17.6% of the variance in another's. This may not sound like much, but it is a greater predictive value than, say, the high end of leading estimates for juror demographic variables as predictors of their verdict preferences;³⁴ it is about the same as the predictive value of LSAT scores on first-year law school grades.³⁵ In all of these cases more than 80% of the variance in outcomes is attributable to factors other than the predictor, including chance. But the fact is that chance (or what looks to us, with our imperfect knowledge, like chance) plays a great role in all kinds of events, indeed a greater role than most people are accustomed to acknowledging. The practical ceiling for the predictability of complex, unbounded judgments like punitive damages awards must be far short of $r = .9$. Because the authors do not clearly and persuasively identify the standard against which the variability of their experimental juries' awards is measured and found wanting, what they make of the degree of variability they found is open to dispute.

Yet the variability is undeniable. Another, very extensive study of both individual and mock jury decision-making (chapter 3; the experiment is discussed in more detail in the next section of this Review) revealed a "strikingly large" degree of variability in different juries' punitive awards in response to identical cases (pp. 54–55). For instance, the mean awards in the 5 cases that generated the highest dollar figures ranged from \$210,000 to over \$83 million; the median was \$10 million, but there was a 10% chance of an award of only

comparing synthetic juries' punitive damages awards (a continuous measure) or their *rankings* of cases in terms of dollar awards; if the latter, this could be converted to a categorical or nominal scale, for which kappa could be used. This technical detail is important in this case because, compared to kappa, a simple correlation exaggerates the difference between participants' near consensus regarding outrage and intent-to-punish and their disagreements regarding dollar awards. It does so precisely because the bounded scale on which outrage and intent to punish were measured would create a much higher *expected* level of interjudge agreement than would the unbounded scale on which damages were measured.

34. See, e.g., Solomon M. Fulero & Steven D. Penrod, *Attorney Jury Selection Folklore: What Do They Think and How Can Psychologists Help?*, 3 FORENSIC REPS. 233 (1990) (reviewing empirical literature and finding overall that demographic and personality variables account for 5% to 15% of the variance in verdict preferences).

35. LAW SCHOOL ADMISSION COUNCIL, LSAT & LSDAS REGISTRATION AND INFORMATION BOOK 121 (2002) (median correlation in 2000 between LSAT scores and first-year law school grades at 183 law schools was .41, yielding r -squared of .168; adding undergraduate GPA to the model yields median r of .50, r -squared = .25).

\$510,000 or less and a 10% chance of an award of \$29 million or more. For the middle 5 cases, awards ranged from \$20,000 to \$60 million, with a median of \$800,000, a 10th percentile of \$140,000 and a 90th percentile of \$11.2 million. Plainly a few outlier awards (of the sort that the trial or appellate court could very well reduce) distend the range; nevertheless, the variation is considerable by any measure.

As the authors explain, the variability of punitive damages awards would hamper a lawyer from being able to make confident predictions about a client's likely exposure (p. 56). Consequences of uncertainty include increases in the rates at which cases go to trial, or "result[] in expensive settlements to avoid uncertain trials, which can stifle innovation and produce other economic inefficiencies" (p. 214). While these statements may be based on slightly inflated expectations about the predictability of legal outcomes in general, as suggested above, and may overstate the economic consequences of uncertainty,³⁶ the authors' basic findings of considerable unpredictability in punitive damage awards are reliable and important.³⁷

Data on the variability of punitive damages awards in real cases would have provided additional context for the authors' research. As the authors point out, the problem with using real world data as a measure of variability is that no two cases are identical, so that differences in awards may be attributable to differences between cases rather than variation in decision-makers, hence the value of the experimental method (p. 214). Nevertheless, an analysis of real verdicts could help suggest the extent to which the degree of unpredictability observed in experiments is likely to be a real world concern as well. In the concluding chapters, the authors address and critique in some detail a leading study by Theodore Eisenberg and others³⁸ of actual punitive awards that found a significant degree of predictability (pp. 214, 245–48).³⁹ While it is not the authors' aim to summarize all

36. One of the authors (Viscusi) has argued vigorously that punitives fail in their deterrent function and lead to economic inefficiencies. W. Kip Viscusi, *The Social Costs of Punitive Damages Against Corporations in Environmental and Safety Torts*, 87 GEO. L.J. 285 (1998). Responses to Viscusi's article by Theodore Eisenberg and David Luban, and Viscusi's response to them, follow in the same issue. See *supra* note 2.

37. See p. 245 (discussion of external validity of experimental findings of punitive award variability).

38. Theodore Eisenberg et al., *The Predictability of Punitive Damages*, 26 J. LEGAL STUD. 623 (1997).

39. The critique (as noted on p. 246 n.7) is largely taken from Viscusi, *supra* note 36. The key arguments, followed by my own brief comments, are as follows: (1) Eisenberg et al.'s use of log awards instead of dollar amounts as the dependent variable, whatever its justification in statistical method, tends to compress the apparent variability of the awards. This point is well

previous research, it does not seem sufficient simply to mention the existence of competing studies on variability and then conclude that “[o]ur reading of this controversial literature is that jury verdicts are highly variable for similar cases, especially across jurisdictions” (p. 214). By presenting some of the data from these studies, and doing so earlier in the book, the authors would have enabled readers to obtain a better perspective on the validity of their experimental work.

III. ARE PUNITIVE AWARDS TOO HIGH? MAKING SENSE OF THE “SEVERITY SHIFT”

Of the three objectionable features of juries’ punitive damages awards—too frequent, too variable, too large—the authors’ research program places the least emphasis on the last. This is understandable, because experts are no more capable than laypeople are of undertaking magnitude scaling without a modulus and producing values reliable enough to be used as benchmarks for “correct” award size.⁴⁰ So the authors would have less of a basis for criticizing participants’ awards as too high than they have when they claim that those awards are too frequent or too variable. The one experiment that directly addresses award size (chapter 3) compares jury awards not to an extrinsic standard (such as the economic norm) or one based on what judges would have decided, but to the predeliberation award prefer-

taken. (2) “Only” 17% of Eisenberg et al.’s sample consists of cases from “more volatile areas of products liability, medical malpractice, and toxic substance liability,” the kinds of cases emphasized in PUNITIVE DAMAGES (p. 246). But these types of cases account for only a very small percentage of all punitive awards (as the research cited *supra* note 28 shows); if anything, the fact that Eisenberg et al.’s data “do not say a great deal about the sorts of cases that we have been emphasizing in this book” (p. 246) reveals the authors’ focus on the atypical, not a shortcoming of Eisenberg et al.’s study. See also *supra* p. 257; *supra* note 38. (3) Most importantly, Eisenberg et al.’s findings that punitive awards are significantly predictable depend on assuming that there will be a nonzero award, which itself is a somewhat unpredictable event. It seems to me that this argument mistakenly shifts the focus away from the predictability of punitives over the relevant long haul. Why could not a defendant simply estimate a likelihood for nonzero awards (based on industry averages, say) and factor that into its *ex ante* determination of expected liability for punitives over a given period of time? If companies can do the kind of cost-benefit analyses that the authors elsewhere promote (and castigate juries for not appreciating, see *infra* pp. 266–76), they ought to be able to come up with good enough projections of P , and hence $P \times L$, for punitive as well as compensatory damages.

40. One of the authors (Viscusi) describes the punitive damages award in the 1999 Chevy Malibu case as “wildly disproportionate to the extent of the harm” (p. 112), so presumably he has some implicit standard of proportionality (or tame disproportionality) in mind; however, this would provide a criterion for judging proper award size only if the extent of the harm were a sufficient basis (perhaps adjusted by some multiplier) for choosing an award. (In chapter 9, Viscusi reports a test of jurors’ compliance with punitive damages instructions that adopts precisely that basis. See discussion *infra* pp. 274–76.)

ences of that jury's members. (This is the "severity shift" experiment, discussed later in this section of the Review.)

But it is the reputed size of punitive damages awards that attracts the most attention in the press,⁴¹ and the book deploys most of what little real world context it offers (prior to Reid Hastie's summary chapter 12) to appeal to this presumed focus of popular interest. Consider in this regard the book's introduction, by George Priest.⁴² Here is how Priest begins:

Over the past two decades, our country has experienced a dramatic increase in the incidence and magnitude of punitive damages verdicts rendered by juries in civil litigation. Perhaps the most extraordinary example is the July 2000 award of \$144.8 billion in the Florida class action brought against cigarette manufacturers. But there are many other examples of huge verdicts. . . . For example, in July 1999 a California jury awarded a punitive damages verdict of \$4.8 billion, and in May of that same year an Alabama jury, in a case in which the economic damages were alleged to equal no more than \$600, awarded a punitive damages verdict of \$580 million. . . .

The[] magnitude [of these verdicts] is remarkable. The award in the tobacco case, for example, equals 2.4 times what our federal government spends each year on education, 52% of the amount spent on national defense, and a full 80% of the amount the federal government collects annually in corporate income taxes (p. 1).

Four things about this passage should leap out at anyone accustomed to reading quantitative research in social science. The first is the argument by anecdote: extreme cases are presented as if they are characteristic of the punitive damages process. Base rate information about mean or median awards, the kind of central tendency data that could at least begin to provide a more accurate picture of punitive awards in general, is completely absent. This is a rhetorical strategy favored by probusiness lobbyists and politicians and uncritical reporters,⁴³ not responsible social scientists.⁴⁴ The second is the use of

41. See, e.g., Marc Galanter, *An Oil Strike in Hell: Contemporary Legends About the Civil Justice System*, 40 ARIZ. L. REV. 717, 726–33 (1998).

42. Of course, the authors did not write this introduction; it is neither their report of their research nor their assessment of its importance and implications. Nevertheless it does seem reasonable to me to hold the authors somewhat responsible for Priest's remarks: they are the first words the conscientious reader is likely to encounter, thus playing an important role in setting the context for the authors' own words, and it is hard to imagine that the authors did not review and approve what Priest wrote.

43. See, e.g., DANIELS & MARTIN, *supra* note 24, at 29–59, 205–13; Galanter, *supra* note 41, at 729–33; Haltom & McCann, *supra* note 28.

44. Argument by anecdote is no more persuasive when offered to show the variability and unpredictability of punitive damages awards. Priest compares the case of *BMW of North*

irrelevant judgment anchors. The size of the federal education budget quite obviously has nothing to do with whether a given punitive damages award is too high or not. Priest's use of these arbitrary benchmarks is somewhat ironic in light of the authors' study (chapter 4) that critiques jurors' undue reliance on irrelevant anchors when determining their awards.⁴⁵ The third is the absence of any information about the relevant facts of these cases, disabling the reader from even beginning to appreciate what might have provoked the jurors to the level of outrage that presumably inspired the awards. The fourth is the lack of support for the claim of a "dramatic increase in the incidence and magnitude" of punitive awards, other than the observation (without citation to any authority) that "[t]wo decades ago it was unusual to observe a punitive damages verdict greater than \$1 million" (p. 2).⁴⁶ It is somewhat off-putting to find such slanted rhetoric at the very beginning of a book of quantitative social science by outstanding scholars.

Let us turn to the research. Over 3,000 jury-eligible adults watched a videotaped case scenario and read case materials, then rated how severely the defendant should be punished, if at all (similar to the intent-to-punish measure used in the experiment in chapter 2

America, Inc. v. Gore, 517 U.S. 559 (1996), in which the jury awarded \$4 million in punitives, to an identical case brought against BMW by a different customer, in which the jury awarded "roughly similar compensatory damages, but no punitive damages whatsoever, concluding that BMW was innocent of the reprehensible behavior that justifies punitive awards" (pp. 2-3). Without any further information about the case, it is impossible to say whether the defendant's conduct may not have been close to the line with regard to one of the legal prerequisites for awarding punitives, in which case it is conceivable that a jury that found all of the elements to be satisfied would (properly) give a substantial award, while another jury that concluded that one of the elements was not satisfied would (properly) consider itself barred from awarding any punitives at all.

45. The federal budget comparisons are arbitrary with regard to the evaluation of jury verdicts but not rhetorically, of course; Priest means to marshal the readers' attitudes toward particular activities and federal expenditures on them in service of his argument. ("What? You mean one punitive award is more than double what the government spends on *education*? That's outrageous.") In this post-Enron climate, however, one might doubt the efficacy of Priest's invocation of the amount the government collects in corporate income taxes as a benchmark.

46. Priest does acknowledge that "[m]any scholars and policy makers have emphasized the relative *infrequency* of punitive damage verdicts across the wide range of civil litigation" (p. 2), citing to one leading scholar (Michael Rustad) whose research supports this position. But rather than engage in or at least summarize the bases for the controversy on this question, he moves on to the "more general concern, . . . the inability to explain these various punitive damages verdicts on a rational basis" (p. 2). The (lack of) rationality of punitive damage verdicts is indeed an important concern, as chapters 5 and 6 (and the next section of this Review) indicate. My point here is only to show how the opening passages of the book's introduction regarding the size (if not also the frequency) of punitive awards are tendentious and misleading.

discussed earlier, but on a 0–8 scale), or determined a punitive damages award. Then participants met in over 500 mock juries of 6 persons and deliberated until they agreed on a punishment rating or punitive damages award, respectively. The process was then repeated for the other dependent measure (e.g., if asked to rate severity of punishment first, participants then decided damages, first individually and then in groups after deliberation). The experimenters found that deliberation tended to polarize punishment ratings in the original direction, *i.e.*, high-punishment cases yielded higher punishment ratings from juries than the mean of individuals' predeliberation ratings for those cases; deliberation reduced the ratings of low-punishment cases below the mean of the predeliberation ratings. Deliberations shifted damages awards, however, in only one direction—up. Eighty-three percent of nonzero dollar verdicts by juries were above the median individual award by that jury's members; 27% were as high or higher than the *highest* predeliberation award by the individual jurors. This is what the authors call the “*severity shift*.” It appears to be statistically significant,⁴⁷ and it is a striking confirmation and extension of similar results in other studies.⁴⁸

What accounts for the severity shift? The authors offer several explanations, all of which make sense. One is that the results comport with the widely recognized phenomenon of “group polarization” (p. 57), the tendency of groups to adopt a more extreme position in the same direction after deliberating than the group's initial, predeliberation tendency. This is due in part to group members' susceptibility to “social influences” (p. 58): members of small groups are generally reluctant to express views too dissimilar from those expressed by others. (This is an instance of what social psychologists call “social proof.”⁴⁹) In almost all of the experimental juries that eventually awarded punitives, a majority of the individual members favored punitives before deliberating (p. 49); therefore, these juries' deliberations most likely featured a preponderance of propunitive arguments. These arguments, especially if they expressed very strong disapproval of the defendant's conduct, probably encouraged those

47. The authors do not report measures of significance.

48. See, e.g., Shari S. Diamond & Jonathan D. Casper, *Blindfolding the Jury to Verdict Consequences: Damages, Experts, and the Civil Jury*, 26 *LAW & SOC'Y REV.* 513, 555 & n.56 (1992) (cited by the authors at p. 215 & n.4).

49. See, e.g., ROBERT B. CIALDINI, *INFLUENCE: THE PSYCHOLOGY OF PERSUASION* 114–66 (rev. ed. 1993).

jurors (if any) in the antipunitives minority before deliberations to go along, increasing support for a higher award.

A second explanation is artifactual. The authors measured the severity shift by determining whether the jury's award exceeded the *median* of members' predeliberation judgments.⁵⁰ Because dollar awards tend to be skewed right—there is no upper limit—the mean of the individual awards would tend to exceed the median. So even if the juries' awards were equal to the means of their respective members' awards (the reasonable assumption underlying the “synthetic jury” method described earlier), there would still be some upward shift from predeliberation individual to (postdeliberation) jury awards. As the authors correctly point out, however, this alone would not account for the roughly one-quarter of jury verdicts at or above the highest individual awards of that jury's members (pp. 59–60).

The authors' most venturesome explanation for the severity shift is that there is a “rhetorical asymmetry” (p. 58) in deliberations that favors those who support larger awards. “[T]hose who argue that ‘more’ money is necessary to punish a corporation appear to have an upper hand. The unbounded dollar scale affords great latitude in the expression of what ‘more’ means” (p. 58). This seems plausible, although it is difficult to pursue the explanation very far without an account of what the participants actually said during deliberations (*i.e.*, a content analysis of the kind provided for the study in chapter 5, discussed later in this Review). In part this “rhetorical” advantage would seem to recapitulate the purely statistical point just made: all things being equal, arguments associated with the higher dollar figures in a right-skewed distribution would exert more “pull” on the final judgment.

The authors set forth the severity shift without stating or even suggesting that it reflects a flaw in the punitive damages award process. By the same token, they do not offer any explanation that might tend to justify the shift. And yet such an explanation is possible. Of the juries that ultimately voted to award punitives, each juror in the predeliberation majority presumably entered the deliberations with one or more arguments that had convinced that juror that

50. The authors ranked the six individual jurors' predeliberation judgments from lowest to highest and then inserted the jury's award into that series (now seven awards). If the jury's award was equal to the mean of the individuals' awards, then it would be the fourth highest out of the set of seven (six individual awards plus the jury award). The difference between this expected rank of the jury's verdict and the observed rank, as a percentage of the maximum possible shift, is what they labeled the “deliberation-shift measure.”

punitives were warranted. (Those jurors whose individual, predeliberation judgment was against punitives may also have thought of such arguments, but found them to be outweighed by others.) During deliberations, it is fair to assume, all of these propunitives arguments (and perhaps more) were raised. Thus, it is likely that many jurors heard during deliberations arguments for awarding punitives that had not occurred to them beforehand. Some they may have considered not especially weighty, or even redundant. But it seems fair to assume that at least some of those new arguments carried weight and helped persuade those jurors that punitive damages were even more deserved than they had initially thought, and therefore, that the punitive award ought to be higher. This is exactly the kind of reasoned discussion that is supposed to occur during deliberations.⁵¹ By omitting what seems to me to be this rather obvious (if necessarily hypothetical, in the absence of a content analysis of the deliberations) account for the severity shift phenomenon, the authors pass up an opportunity to make at least this aspect of juries' thinking about punitive damages seem reasonable and justifiable.

The same experiment also showed that deliberations exacerbated the variability and unpredictability of awards. The 10th percentile awards by juries were a smaller fraction of the median awards than were the 10th percentile awards by "statistical juries" (based on the means of six individual jurors' awards taken at random); the 90th percentile awards by mock juries were a greater multiple of the median awards than were the 90th percentile awards by statistical juries. Some, but not all, of the increased variability is due to the increased size of the awards given by deliberating juries (p. 57) (*i.e.*, all things being equal, the larger the median award, the more room there is for the distribution to spread out). This finding that deliberations increase the unpredictability of punitive damage awards bolsters the implications the authors want to draw from the results of the experiments discussed previously, because those results were based on the judgments of individual participants, whereas the decisions of

51. Cf. Diamond & Casper, *supra* note 48, at 555–56 (explaining inflation of jury award over mean of individual predeliberation awards by pointing out that jurors' exposure during deliberations to other jurors' views of the blameworthiness of the defendant's behavior may have led group as a whole to consider the defendant's misconduct to be more serious, leading in turn to higher damages). My colleague Steve Gilles points out that thorough, reasoned deliberations should also result in some jurors being persuaded by reasons *against* awarding punitives. This is true, but it seems fair to assume that if the predeliberation majority of jurors are inclined to award punitives, arguments in favor of punitives would predominate. Again, a systematic content analysis of the deliberations could test these and other hypotheses.

deliberating mock juries are likely (again, all things being equal) to have greater external validity.⁵²

IV. DO JURIES DECIDE PUNITIVE DAMAGES IRRATIONALLY?

Throughout the present research, we have been impressed by the serious and energetic manner in which citizens performed the difficult legal judgment tasks that are demanded by the punitive damages decision. The many systematic patterns of behavior we have observed are convincing evidence of the jurors' conscientiousness. Nonetheless, the legally required decision tasks often seemed to exceed their individual and social capacities. The decision task is not well defined by the jury instructions[, and] jurors are not provided with the necessary background information or experiences to make reliable judgments. . . . (p. 241).⁵³

So how do jurors respond in these trying circumstances? Studying the processes by which jurors and juries reach punitive damages judgments is at least as important to the authors as examining the pattern of judgments themselves (*see, e.g.*, pp. 2, 80). More than half of the experiments reported in the book lend support in one form or another to the general proposition that juries' punitive damages awards are flawed in relation to the norm of rational judgment (as well as the other norms they employ). Jurors tend to downplay or ignore factors that ought to influence their decisions (such as the jury instructions and whether the defendant's failure to take precautions was cost-justified), to take into account factors they ought to ignore (such as the plaintiff's lawyer's requested dollar figure), and to fall prey to systematic cognitive biases (such as the hindsight effect and the overestimation of low-probability risks). All of these decision-making habits tend to lead to larger, more frequent, and more unpredictable awards than more rational judgment processes would yield.

Some of the reported findings are relatively unsurprising but still useful confirmations of phenomena frequently noted in the cognitive and social psychological literature. I will discuss these briefly first. Then I will turn to the several chapters (7-11) the book devotes to establishing that jurors perform poorly as risk managers. I will conclude by returning to the study (chapter 5) showing that jurors tend to ignore instructions in deciding whether punitives are proper.

52. Note, however, that by concluding that "shared moral judgments do not produce predictable dollar awards" (p. 61), the authors imply that predictability is either present or absent, as opposed to being a matter of degree—a slightly leading interpretation of the data that, once again, seeks to reinforce the authors' overall picture of juries' punitive awards.

53. *See also* pp. 26, 79, 212 (where Hastie makes similar assertions).

V. DO JURORS DISPLAY COGNITIVE BIASES IN THINKING ABOUT PUNITIVES?

The study reported in chapter 4 offers convincing evidence that jurors' punitive damages awards are prone to being influenced by a salient but legally irrelevant anchor, the dollar figure for which the plaintiff's lawyer asks (the punitive ad damnum). Jurors saw a videotaped narration and read a written summary of an environmental accident case, and were told that the defendant had already paid \$24.5 million in compensatory damages; they were also told that the defendant's actions had already been found to be reckless, so that some punitives were permitted. In some conditions, the plaintiff's lawyer requested a punitive award of "between \$15 million and about half a year's profit, \$50 million" (low anchor); in others, the lawyer requested "between \$50 million and about a year's profit, \$150 million" (high anchor) (pp. 65–66). The results were that, in the low-anchor condition, jurors' median award was \$15 million (mean \$19.5 million); in the high-anchor condition, it was \$50 million (mean \$51.9 million) (pp. 68–69).⁵⁴

Anchoring and adjustment is a typical if not universal judgment strategy for making quantitative estimates (pp. 216–17). If the anchor is relevant and the adjustment is reasonably sufficient, the strategy yields good enough estimates. The problem, however, is that people tend to be unduly influenced by irrelevant or even absurd anchors, and having selected an anchor (relevant or not), tend to make insufficient adjustments. All of this is well established in the cognitive psychology literature going back to the classic work of Tversky and Kahneman in the 1970s.⁵⁵ So it is no surprise to find jurors faced with an especially difficult estimation task—"magnitude scaling without a modulus," as they put it (p. 41)—being particularly likely to rely unduly on arbitrary anchors.

54. To emphasize that the difference in median awards between the high-anchor and low-anchor conditions was "\$35 million for an identical fact situation" (p. 73; see also p. 219) may characterize jurors' cognitive pliability in a somewhat unduly pejorative way, to the extent that readers treat the \$35 million figure as independently meaningful ("Wow, that's a lot!") and not as the experimental manipulation itself (the high-anchor was \$50 million and the low-anchor was \$15 million). And both damages requests were plausible given the \$24.5 million in compensatory damages already awarded.

55. The authors (p. 73 n.9) cite the classic article, Amos Tversky & Daniel Kahneman, *Judgment Under Uncertainty: Heuristics and Biases*, 185 SCI. 1124 (1974). Kahneman is coauthor of three of the book's chapters, although not chapter 4.

Yet the authors' experiment showed that the plaintiff's award request was actually the *second* most popular anchor according to participants' own explanations for their judgments; the modal anchor was the compensatory award (stipulated in the stimulus materials) (pp. 68–69). The Supreme Court, in *BMW of North America, Inc. v. Gore*,⁵⁶ looked to the relationship between the punitive and compensatory awards as one factor in determining whether the punitive award is unconstitutionally disproportionate.⁵⁷ This implies that the participants most often chose to anchor their judgments in an arguably justifiable number.⁵⁸ The difficulty of adjusting from the anchor remains, of course, and neither the Supreme Court nor any other legal authority has indicated the multiple (or fraction) of compensatory damages jurors ought to employ to arrive at their punitive awards (p. 248).⁵⁹

The study reported in chapter 6 purports to show that, in deciding whether to award punitives, jurors are also prone to the hindsight bias, just as they are in every other legal judgment task with regard to which possible hindsight effects have been examined. Participants watched a videotaped narration and read a written summary of a case in which the defendant railroad wanted to carry toxic herbicides along a stretch of mountainous track that the National Transportation Safety Board (“NTSB”) had declared hazardous. In the foresight condition, jurors were to decide whether the NTSB should lift its order stopping the railroad's operations along the track before any

56. 517 U.S. 559, 580–83 (1996).

57. Viscusi (p. 112) implicitly uses “the extent of the harm” as an anchor for judging the inappropriateness of a punitive award. *See supra* note 40.

58. Given this, the rhetoric of the concluding paragraph of the chapter (p. 74) seems especially tendentious. The authors begin: “In conclusion, the present results reinforce the description of punitive damages awards as highly variable and arbitrary,” and, after remarking on the range of awards, continue: “However, the two experiments reported in this chapter do point to the fact that punitive damages awards can be influenced by predictable factors”—namely the *inappropriate* factors of the plaintiff's lawyer's request and the geographical location of the lead plaintiff. By omitting the predictability generated by the compensatory award, the authors overstate the gap between participants' decision-making and the norm of rational judgment.

59. On the other hand, one might even question whether the jurors who anchored their judgments in the plaintiff's lawyer's request were really behaving completely irrationally under the circumstances: without any other reliable guidance, they may well have deferred to the figure chosen by an “expert”—a person they should have recognized as biased, but one they might properly assume was as knowledgeable as anyone else about the value to be matched to the defendant's reprehensible conduct. Given the law's complete lack of guidance as to how to translate outrage at a defendant's conduct into a dollar amount, no anchor currently available to juries can entirely avoid the label “arbitrary.” In the book's concluding chapter, Sunstein recommends providing the punitives decision-maker with anchors in the form of punitive awards in comparable cases; I will discuss this promising suggestion later in the Review.

accident occurred. In the hindsight condition, jurors were told that the accident had occurred and that full compensatory damages had been paid; jurors were asked to decide whether punitives should also be awarded and, if so, in what amount (pp. 100–02).

Knowing that the accident that the defendant's behavior risked had actually occurred led participants to estimate a higher *ex ante* probability that it *would* occur. Two-thirds of participants in the hindsight condition voted to award punitive damages; by contrast, only one-third of those in the foresight condition decided that the NTSB should not lift its order (*i.e.*, that the railroad should not be allowed to continue its operations). And hindsight made participants significantly more likely to judge that each of the component elements of punitive damages liability (*e.g.*, that the defendant was reckless, that the defendant disregarded the risk of a grave danger, etc.) was met than participants in the foresight condition were to judge that each of the equivalent conditions for not lifting the NTSB order was met (pp. 103–05).

Although not apparent from the authors' description of the experiment (pp. 100-01), their earlier report of what appears to be the same study and their stimulus materials indicate important differences in the information provided to participants in the respective foresight and hindsight conditions (apart from the intended manipulation of whether the accident occurred) which arguably confound their results.⁶⁰ Nevertheless, the hindsight effect has been so well established by other research, if not also by the study reported in chapter 6, that it is reasonable to believe that jurors' punitive damages decisions may be prone to this effect. Such an effect would be normatively unjustifiable. The hindsight effect constitutes a bias in relation to the social scientific norm of rational judgment, as well as in the sense that it makes jurors more likely to find defendants liable (for compensatory or punitive damages). In addition, there is no known reliable cure for the bias (pp. 107–08). So to the extent that the hindsight bias affects punitive damages decisions, the authors' case for the irrationality of the jurors' decision-making regarding punitives would be strengthened.⁶¹

60. Richard Lempert, *Juries, Hindsight, and Punitive Damages Awards: Failures of a Social Science Case for Change*, 48 DePaul L. Rev. 867 (1999) (reviewing Reid Hastie & W. Kip Viscusi, *What Juries Can't Do Well: The Jury's Performance as a Risk Manager*, 40 ARIZ. L. REV. 901 (1998)).

61. Nor does there appear to be any normative justification for jurors to grant larger punitive damages awards to local as opposed to out-of-state plaintiffs, as the authors found in

VI. ARE JURORS POOR RISK MANAGERS?

The authors (W. Kip Viscusi is the sole author of four of the five chapters in this section) marshal several studies to show that when making punitive damages awards, jurors are oblivious to the most basic principles of optimal deterrence—the idea that in regulating risky behavior, legal decision-makers ought to encourage actors to avoid only those accidents worth avoiding (p. 109). Jurors incorrectly estimate the magnitude of mortality risks. They seem incapable of following instructions to take deterrence into account as a goal of their punitive awards. And they are often unwilling to place a reasonable dollar value on safety, as is required to determine whether their award would promote optimal deterrence or not. Perhaps least justifiably, they actually judge a defendant that competently performed a cost-benefit analysis and decided that taking safety precautions would cost more than the expected benefit in accident costs *more* harshly than they do a defendant that simply declined to take precautions without first weighing costs against benefits. Most of these findings are sound, and cumulatively they paint a convincing portrait of jurors as poor risk managers. That jurors should punish defendants for weighing costs and benefits seems especially troublesome, even if this finding may not surprise observers of tort juries.⁶² In two of the studies, however, the authors' data does not entirely justify the pejorative inferences they draw regarding jurors' unwillingness and inability to implement the goal of deterrence.

the study reported in chapter 4. On the other hand, the jurors in the study reported in chapter 2 who assessed larger punitive awards against wealthier defendants (a "large" company with profits of \$100 million to \$200 million per year as opposed to a "medium-sized" company with profits one-tenth that large) may well have been legally justified in doing so. While this defendant's wealth effect might suggest to some readers (as it apparently does to the authors, *cf.* pp. 40, 113–14) an anticorporate defendant bias, it is perfectly legitimate for jurors to take the defendant's wealth into account in deciding how much to punish the defendant, because all things being equal, a greater award may be necessary to inflict equal punishment on a wealthier defendant. *See* *Pac. Mut. Life Ins. Co. v. Haslip*, 499 U.S. 1, 21–22 (1991) (defendant's wealth one of several factors that may properly be taken into account in setting punitive award); W. PAGE KEETON ET AL., *PROSSER & KEETON ON THE LAW OF TORTS* 15 (W. Page Keeton ed., 5th ed. 1984). The authors conceal their own preference for the optimal deterrence goal of punitives behind their assertion that "[w]ithin the academic community, opinion is sharply divided on the question whether the amount of punitive awards should depend on the size of the defendant firm" (p. 40); in fact, opinion is "sharply divided" only between those economic theorists who think it is irrelevant and everyone else. *See also* Eisenberg et al., *supra* note 38, at 629 & nn.15–19 (noting that even some economic theorists allow that the defendant's wealth may be relevant to punitives).

62. *See, e.g.*, Gary T. Schwartz, *The Myth of the Ford Pinto Case*, 43 *RUTGERS L. REV.* 1013, 1035–36 (1991); Michael Wells, *Scientific Policymaking and the Torts Revolution: The Revenge of the Ordinary Observer*, 26 *G.A. L. REV.* 725, 736 (1992).

A. *Jurors' Risk Perceptions Are Biased*

The first requirement of a good risk manager would seem to be the ability to appraise risks in a relatively objective, unbiased, and accurate fashion. Laypeople's inability to do these things is well documented,⁶³ and the experiments in chapter 10 confirm several of these flaws. When participants were asked to estimate annual mortality risks from a variety of causes, they tended to overestimate the frequency of unlikely causes of death (*e.g.*, fireworks, botulism) and underestimate the frequency of common killers (*e.g.*, diabetes, stroke, cancer). This kind of overestimation of low-probability risks would incline jurors to hold the defendants who create those risks liable for not taking steps to avoid them (pp. 180–83). A second cognitive bias compounds the first. When presented with a range of scenarios in which the expected accident loss was held constant but its component values—the probability of an accident and the severity of the loss should an accident occur—were manipulated, participants were insufficiently sensitive to reductions in the probability of the accident. Consequently, their thinking was dominated by the corresponding increase in the magnitude of the potential loss, making defendants who are responsible for large-loss but low-probability accidents (a commonly litigated products liability scenario) especially vulnerable to punitive damages (pp. 175–77). Then, when participants were implicitly asked to place a value on life that could be plugged into a cost-benefit calculus (specifically, they were asked how much they would pay to cut their annual motor vehicle fatality risk in half, from 1 in 10,000 to 1 in 20,000), nearly 10% of participants responded essentially that no amount was too much to spend on safety. This “zero-risk mentality” is plainly not feasible as a matter of policy, and would also lead participants to hold defendants liable for failing to avoid risks even if they could have done so only by means of precautions that were clearly not cost-effective (pp. 183–84).⁶⁴ Finally, when accident losses were described as uncertain (*e.g.*, the defendant actually caused \$10 million losses but was fortunate in that there was a 90% chance that losses could have been much worse, as high as

63. See, *e.g.*, Paul Slovic et al., *Facts Versus Fears: Understanding Perceived Risk, in JUDGMENT UNDER UNCERTAINTY: HEURISTICS AND BIASES* 463–89 (Daniel Kahneman et al. eds., 1982).

64. The fewer than 10% of respondents who placed an infinite value on life may not have much practical effect on jury decision-making, because there would rarely be more than one such person on a six-person jury (p. 184).

\$100 million), many participants tended to award damages based on the worst-case scenario—which, if done in all cases, would lead to greatly excessive deterrence (pp. 177–80). When all of these robust little irrationalities are cumulated, the result (experiment 1) is that jurors tend to hold defendants liable for punitive (and not merely compensatory) damages for failing to avoid a low-probability risk even though the cost of avoiding that risk was greater than the expected benefit in accident losses—so that, according to optimal deterrence theory, the defendant should not be considered negligent, much less reckless and thus potentially liable for punitives.⁶⁵

B. Jurors Punish Defendants for Conducting Cost-Benefit Analyses

To my mind some of the most interesting and telling results in the book concern jurors' responses to a defendant's performance of a cost-benefit analysis (chapter 7). In all scenarios, the defendant automobile company knew that it had designed a line of cars with a defective electrical system; the defect had led to ten burn deaths per year. The researcher (Viscusi) manipulated whether the defendant conducted a cost-benefit analysis before deciding not to fix the design; whether, in doing the analysis, the defendant used a low value-of-life figure (based on average compensatory damages awards) or a high one (based on National Highway Traffic Safety Administration data); and whether the analysis was based on a mistaken overestimate of the cost of precautions per life saved (pp. 115–21).

Far from being more lenient toward a defendant that could justify its actions on the ground that a competent cost-benefit analysis showed that the cost of precautions exceeded expected accident losses, participants punished the defendant more severely when it had done such an analysis: the geometric mean of punitive damages awards was one and a half times higher, and the median punitive award was ten times higher, when the defendant had done the analysis than when it had not.⁶⁶ And when the defendant, in conducting its

65. I have already discussed this study in an earlier section of this Review. I will only repeat here that the tendency of actual judges as well as jury-eligible adults to vote for punitives where the accident was very unlikely (1 in 100,000 or 1,000,000) but the magnitude of the loss was very great (29 or 290 deaths, respectively) may indicate not so much an insensitivity to decrements in accident probability as it does a sensitivity to the value of life—not necessarily an undue or even infinite sensitivity, just a judgment that when many lives are at stake, a company ought to be judged reckless for not “overspending” on safety by \$500 to fix a cargo door. (pp. 176–77).

66. The authors overstate their results in one small respect: in table 7.2 (p. 118), summarizing the result of the comparison of jurors' awards in scenario 1 (no cost-benefit analysis, \$4

analysis, placed a higher value on life (using the government figure of \$4 million per life rather than the \$800,000 figure based on compensatory damages), participants again punished it for doing so, awarding punitive damages that were 20% higher (geometric mean) or nearly three times higher (median) (pp. 122–24).

Whether jurors' lack of receptivity to corporate cost-benefit analyses should be "linked to anticorporate bias" (p. 113), as Viscusi implies, is debatable, given at least equally compelling evidence of jurors' generally probusiness attitudes.⁶⁷ The experimental results are better explained by other psychological phenomena. Viscusi, for instance, suggests that jurors may punish corporations for risk-benefit balancing because jurors may regard trading-off money and lives as taboo,⁶⁸ or because, when they judge in hindsight, the trade-off seems to be not between costs and the low probability of an accident, but between a small expenditure by a (big) company and an identifiable death (p. 130).⁶⁹ He also points out that the defendant's use of a

million per life) vs. scenarios 3-5 (cost-benefit analysis), the authors describe the finding as "[s]ignificant effect with large influence on award level." The "large influence" is a euphemism for a result that approaches but does not meet the conventional .05 standard for statistical significance (here, $t = 1.85$, $p < .10$).

67. See VALERIE HANS, *BUSINESS ON TRIAL* 138–77 (2000). Viscusi's support for his claim on anticorporate bias is a 1998 survey published in the *NATIONAL LAW JOURNAL* (p. 114 n.7). Such poll responses, however, are unlikely to be reliable predictors of respondents' behavior should they actually become jurors, mainly because (1) people's responses to the very general kinds of questions asked in this and most other polls (to use one of Viscusi's examples, whether executives of big companies often try to cover up the harm they do) may very well differ from their judgments as jurors with regard to the specific, vastly more detailed facts presented at trial (given all of the evidence, did *this* defendant try to cover up harm?), and (2) when people respond to surveys, their thinking and answers are not constrained by oath, adversarial presentation and argument, and deliberation with fellow jurors. Moreover, whether any "general suspicion of corporate motives" (pp. 113–14) on jurors' part would be unwarranted, as Viscusi also implies, is also debatable, given the well-known incentives that corporations and their officers have to value profit and professional advancement over customer safety. David Luban, *A Flawed Case Against Punitive Damages*, 87 *GEO. L.J.* 359, 371–74 (1998).

68. Cost-benefit analysis need not be and often is not confined to a trade-off between safety and "costs" defined as decrements in the company's profits. Rather, costs are typically understood to include as well the *social* costs of precautions, including the unavailability of other desired products and decreases in safety due to the company's inability to take other, possibly more efficacious precautions. The stimulus materials for at least one scenario used in the experiment reported in chapter 7 appear to have presented participants with the broader sense of costs (p. 120), but others did not (pp. 117–18). To the extent that participants were allowed or encouraged to identify the cost side of the equation solely with lower company profits, their anger at companies that used these cost-benefit analyses to justify their failure to take precautions becomes more understandable; such participants might have been more lenient had they believed that the defendants were not simply trading lives for money. (I thank my colleague Linda Meyer for pointing out this argument to me.)

69. To put this point another way: when jurors hold defendants responsible for failing to take (suboptimal) precautions to avoid low-probability events, they are not necessarily giving vent to an anticorporate or antidefendant bias (see p. 233); they may be reflecting the difficulty

higher dollar figure for the value of life, while it ought to have signaled the defendant's greater concern for safety and thus led participants to perceive the defendant as less deserving of punishment, probably offered participants a higher anchor for their awards (pp. 125–26).⁷⁰

Still another possibility is that the results reflect the operation of prototype effects. Some or many of the participants who punished the defendant more severely for conducting a cost-benefit analysis may have reasoned that “car company, burn deaths, cost-benefit analysis” equaled “Ford Pinto,” which equaled “defendant is bad and should be punished.”⁷¹ To be more precise, in the real Ford Pinto case, jury outrage was likely based on certain perceptions regarding Ford's behavior, such as what the jurors perceived to be Ford's manipulation of the terms of the cost-benefit analysis it offered to justify its refusal to redesign the car to avoid hundreds of burn deaths and its vigorous lobbying against a government safety standard that would have required it to redesign the gas tank.⁷² Participants in the experiment may have been aware (even if vaguely) of these aspects of the Ford Pinto case (and any other cases they believed to be similar) and incorporated them into their prototype of the profits-over-safety car company. In the absence of information that distinguished the defendant from the prototype, participants *assimilated* the defendant to the prototype, and reacted with corresponding outrage. Had the stimulus materials more pointedly differentiated the defendant from the prototypical corporate wrongdoer, however, participants' thinking may well have displayed a *contrast* effect instead, leading to less

they have in using their ordinary, intuitive blaming schema to evaluate corporate behavior. Jurors' blaming judgments typically array the blameworthiness of behavior along a continuum, with intent-to-harm being judged the most reprehensible, followed by bringing about harm knowingly or recklessly, and so on down to unforeseeable, accidental injury as the least blameworthy. The shared outrage and intent-to-punish rankings reported in chapter 2 reflect this common schema (pp. 24–25). Corporate behavior that causes physical injury, however, as in the typical products liability setting, seems to be *both* accidental (which ought to indicate low blame) and *knowing* (which ought to indicate high blame): if a company knows that one out of every million widgets will cause harm, then if it makes enough widgets, it knows to a statistical certainty that harms will occur. The result of this tension or dissonance may be judgments of blame that are discrepant from the ones the schema routinely prescribes.

70. This resembles another instance in which jurors' use of anchors not only seems odd but is precisely contrary to the aims of the entity that unwittingly provided the anchor—jurors' use of statutory damage caps as an anchor for damage awards, which actually increase average awards, contrary to legislative intent (see p. 218, where Hastie, citing the leading studies, makes precisely this point).

71. See Schwartz, *supra* note 62 (mythic status of Ford Pinto case).

72. Cf. Mark Dowie, *Pinto Madness*, MOTHER JONES Sept.-Oct. 1977, at 18; Schwartz, *supra* note 62, at 1017–20.

punitive behavior.⁷³ Thus, the reported findings may not generalize to situations in which the defense attorney presents evidence that persuasively distinguishes the defendant's conduct from that of jurors' prototype.

Viewing the findings in their best light, though, the participants' judgments remain hard to justify. "The direction of the effect [of the defendant's performance of a cost-benefit analysis on punitive awards] is disturbing because jurors are doing the opposite of what juries should be doing if corporations are to be encouraged to think systematically about risk and cost trade-offs" (p. 124).⁷⁴

C. *Jurors Avoid Thinking in Terms of Economic Efficiency*

Two other studies offer results that are internally valid but do not quite support the broadly pejorative characterizations of punitive damages juries that the authors draw from those results. In chapter 8, the authors ask: "Do people want optimal deterrence?" (p. 132). The answer, as one might expect by this point, is no. The argument is as follows: if all accidents for which defendants were legally responsible resulted in tort suits and (appropriate) compensatory damage awards, then defendants would have to internalize all expected accident costs. Out of self-interest they would risk incurring those costs only when it was cheaper to do so than to spend on precautions to avoid them. Thus, tort law would regulate accident-causing behavior in an economically efficient fashion. Some accident-causing behavior goes

73. See Norbert Schwarz & Herbert Bless, *Constructing Reality and Its Alternatives: An Inclusion/Exclusion Model of Assimilation and Contrast Effects in Social Judgment*, in *THE CONSTRUCTION OF SOCIAL JUDGMENTS* 217 (Leonard L. Martin & Abraham Tesser eds., 1992).

74. The same experiment also showed that jurors were not responsive to factors that ought to have affected their judgments if they were thinking in terms of optimal deterrence: they did not punish more harshly a defendant who could have avoided the accident at a lower cost (\$1 million versus \$4 million per life saved) (pp. 119–20), nor did they punish more harshly a defendant who made a mistake in its cost-benefit analysis, calculating a cost of \$4 million per life saved when the actual amount was \$2 million) (p. 126). Perhaps the manipulation of cost-per-life-saved did not affect jurors' decisions because they considered even the more expensive precaution cheap enough to be well worth taking, so that the perceived reprehensibility of not taking worthwhile precautions to save life drowned out any differences in how much those precautions would have cost. Note that even at \$4 million per life saved, the precaution here is pretty close to the NHTSA value of life figure (\$3 million), so jurors might well have been outraged at the defendant for taking such a chance with the lives of others. Perhaps a stronger manipulation, including an experimental condition in which the precaution was obviously not cost-effective (e.g., \$100 million per life saved), would have produced a significant effect. And perhaps similar observations apply to the lack of effect of the defendant's making a mistake in the cost-benefit analysis (jurors may have thought: "Two million, four million, what's the difference? The defendant should have avoided the harm in any case.").

undetected, however, because injured persons cannot identify the injury or the culpable party, or choose not to sue, or sue and lose when they deserve to win. To make up for this shortfall in deterrence, in cases in which the defendant is sued and compensatory damages are awarded, punitive damages should also be assessed, and increased to the extent that the probability of detection is low.⁷⁵

The authors' first experiment shows that participants' punitive awards were unaffected by manipulations in the probability that the defendant's tortious conduct would be detected, as optimal deterrence theory requires. (There is actually an insignificant trend in the opposite direction: awards *increased* as the probability of detection increased.) (pp. 137–38). In a second experiment, a majority of University of Chicago law students disagreed with the decision of the trial judge in the stimulus case to set aside the jury's punitive award based on the judge's (uncontested) finding that punitives were not needed to deter the defendant's conduct because the plaintiff would receive full compensation. From these studies the authors conclude that "[p]eople do not spontaneously think in terms of optimal deterrence, and their proposed punishments do not vary with the probability of detection" (p. 141). People do not want optimal deterrence.

It is possible, as the authors suggest, that the pattern of punitive awards indicates that people care about deterrence but just do not understand how to implement it—specifically, that they do not understand the connection that the authors posit between the probability of detection and optimal deterrence (pp. 138–39). I would go further. It seems to me that these results support at most the proposition that people are unwilling to use deterrence to *trump* a punitive award that they deem justifiable in terms of the other functions that punitives serve (including retribution and the public declaration of

75. For a critique of the theory that punitive damages should be awarded if and only if necessary to make up for a shortfall in deterrence created by the failure to detect acts of negligence, see Mark F. Grady, *Efficient Negligence*, 87 GEO. L.J. 397, 397–98 (1998). In any event, it is certainly inconsistent with current law. For instance, in *BMW of North America, Inc. v. Gore*, 517 U.S. 559, 575–83 (1996), the Supreme Court used the size of the plaintiff's compensatory damages and the size of the civil fines for misconduct comparable to the defendant's as benchmarks for the proportionality and hence constitutionality of punitive awards, thus indicating that punitives should be positively and not negatively correlated with the extent of detection (*i.e.*, the *greater* the liability to which the defendant is or would otherwise be subject, the *larger* the permissible punitive damages award). Moreover, the Court explicitly identified the reprehensibility of the defendant's conduct as another "guidepost" for determining the reasonableness of the punitive award, quite apart from questions of deterrence.

wrongdoing⁷⁶). In the second experiment, for instance, the defendant company's injury-causing conduct was described as "grotesquely reckless"; the facts stated that the defendant "did not even try to take the most minimal and obvious precautions to protect workers against serious risks to life and health" (p. 140). The majority of participants who declined to endorse the trial judge's decision to set aside the punitive award on the ground that compensatory damages would provide adequate deterrence may very well have been happy to further deterrence in other contexts,⁷⁷ but here, they perceived it to be so obvious that the defendant deserved punishment that they preferred to uphold the punitive award rather than eliminate it entirely (which was the only choice presented by the experiment).

Or perhaps, as the authors suggest (p. 138), probability of detection is just too subtle a cue to trigger jurors' awareness of deterrence concerns. In chapter 9, the authors offer strong evidence that jurors do not follow punitive damages instructions that explicitly call their attention to, explain, and even provide a formula for implementing optimal deterrence. These instructions were proposed by A. Mitchell Polinsky and Steven Shavell in a *Harvard Law Review* article,⁷⁸ and are referred to as the "Polinsky-Shavell instructions" (pp. 142, 165–66). The instructions on the "deterrence objective" consist of four paragraphs, about 250–300 words, which no doubt are clear enough to attentive, educated readers. For instance:

2. To achieve the deterrence objective, your principal task is to estimate the likelihood that the defendant might have escaped having to pay for the harm for which it should be responsible. Thus, for example, if the harm was noticeable and likely to lead to a lawsuit, your estimate of the likelihood of escaping liability would be relatively low. But if the harm might not have been attributed to the defendant, or if the defendant tried to conceal its harmful conduct, your estimate of the likelihood of escaping liability would be relatively high (p. 165).

The deterrence instructions also include a table correlating the probability of escaping liability (in percentage terms) with a "punitive damages multiplier" (in decimal terms); jurors are supposed to

76. Marc Galanter & David Luban, *Poetic Justice: Punitive Damages and Legal Pluralism*, 42 AM. U. L. REV. 1393, 1428–38 (1993); see *infra* p. 286 & n.7.

77. For example, data from the study reported in chapter 5 indicates that the occurrence during deliberations of the proposition that the purpose of punitive damages is to deter others was the *strongest* predictor that the jury would award punitives (p. 87).

78. A. Mitchell Polinsky & Steven Shavell, *Punitive Damages: An Economic Analysis*, 111 HARV. L. REV. 869 (1998).

multiply the compensatory damages (specified in the scenario) to yield the “base punitive damages amount” (p. 165), elsewhere called the “deterrence value.”⁷⁹ The instructions continue with about 200 words on the “punishment objective” of punitives and how to determine an appropriate “punishment amount.” No table is offered here, but rather guidelines such as:

3. In considering how well the imposition of punitive damages will fulfill the punishment objective, you should also bear the following in mind: . . . b) the extent to which you believe that innocent parties will suffer as a result of the imposition of punitive damages on the defendant; such parties might include shareholders as well as customers, who may have to pay higher prices for the defendant’s products or services. The more likely it is that innocent parties will be punished, the lower should be the level of punitive damages (p. 166).

The instructions conclude with about 100 words explaining that the final punitive award should be between the deterrence and punishment amounts.

The major finding is that “people did not carry out the Polinsky-Shavell instructions in setting punitive damages” (p. 163). Across all scenarios (in which the researchers manipulated, among other things, the probability that the defendant’s risky conduct would be detected and whether an anchor for the punitive award was presented), 7% of the participants simply did not provide a deterrence value (*i.e.*, they did not complete the seemingly straightforward computation, required by the deterrence instructions, of multiplying the compensatory damages by the appropriate “punitive damages multiplier” derived from the table); another 78% provided an incorrect value. So most participants failed to correctly do the math they were instructed to do. In scenarios in which the probability of detection was very low (1%), so that the correct computation according to optimal deter-

79. So, for instance, if the probability of escaping liability is .5, the multiplier is 1.0—punitives are equal to compensatory damages, so that the total tort liability (compensatory plus punitives) is twice the accident losses; because the probability of detection is .5 (*i.e.*, there was a 50% chance that actual losses would go uncompensated), doubling the damages is needed to make liability equal risks created ($.5 \times 2 = 1$), as is required for optimal deterrence to work. If the probability of escaping liability is 90%, the multiplier is 9 (*i.e.*, if the defendant had only a 10% chance of having to pay out any damages), setting punitives at 9 times compensatory damages yields total liability of 10 times compensatory damages, so $.1 \times 10 = 1$; again, liability equals risks created. If the probability of escaping liability is 99%, the multiplier is 99. (Why the left column of the instruction table is set in terms of “probability of escaping liability” instead of the probability of *detection*, the complementary concept the authors choose to explain optimal deterrence elsewhere (chapter 8), is unclear to me, though it would seem to add a needless complication to the jurors’ task.)

rence (as indicated in the table in the instructions) would have been to multiply compensatory damages by 99, participants' deterrence values reflect their use of a multiple of less than 2, indicating that they did not use the deterrence table as instructed.⁸⁰ Participants did, however, raise their deterrence values (and overall punitive awards) in response to anchors, whether purportedly irrelevant (plaintiff's lawyer's request) or arguably more relevant (media report of punitive award in a similar case). Finally, despite the Polinsky-Shavell instructions' clear emphasis on the deterrence objective, the punishment value, not the deterrence value, played a larger role in participants' final punitive damages awards (pp. 151–58).

Does the failure to follow the Polinsky-Shavell instructions show that jurors simply cannot or will not grasp the deterrence function of punitives? Perhaps a more modest conclusion would be appropriate. The syntactical and conceptual complexity of the instructions, taken as a whole, strongly suggests that even well intentioned and otherwise capable participants may have decided to abandon their efforts to comply and fallen back on their common sense notions of what punitive damages are for.⁸¹ Viscusi, the chapter's author, writes that “[m]any respondents are simply reluctant or unable to carry out even the most basic mathematical calculations” (p. 164). Maybe—but those calculations were set in a dauntingly complex set of verbal instructions, so that many participants probably just gave up before they bothered to do the math.⁸² “Moreover, [many respondents] appear quite willing to abandon the jury instructions when they have other rationales for setting punitive damages that they find to be either more convenient or more compelling,” and this “is consistent

80. The resulting awards were, however, much *lower* than they should have been, contrary to the tenor of the results of other studies in the book.

81. Cf. p. 93, where the authors offer a similar explanation for mock juries' failure to follow punitive damages instructions modeled on actual ones rather than on the Polinsky-Shavell proposal.

82. Viscusi observes: “Perhaps most troubling is that these difficulties [in accurately calculating the deterrence or ‘base’ punitive damages amount, taking into account the probability of detection] are not random, but are highly concentrated among particular demographic groups, specifically minorities and the less well educated” (p. 163). This is based on correlations showing that significantly larger percentages of women, Hispanics, and those with some college education or less are to be found among those who incorrectly determined the deterrence amount than among those who determined it correctly (p. 161 & tbl. 9.5). Viscusi speculates that the discrepancy “may be indicative of a greater reluctance by female respondents to surrender their punitive damages judgment to a mathematical formula” (p. 161); he does not, however, offer this rationalization for the other groups, whose disproportionate failure to get the deterrence amount right is chalked up to “[f]inding] the basic multiplication tasks . . . too difficult” (p. 163).

with the performance of juries more generally” (p. 164). Maybe—but it seems rather hasty to draw such broad inferences from participant behavior that may very well reflect a situation-specific response to these particular instructions. Rather than, or in addition to, criticizing participants for failing to respond positively to the “detailed rationale and mathematical formula for setting punitive damages” (p. 162), Viscusi might have contemplated the possibility that the instructions were *too detailed* (and yet still insufficiently clear) for the typical jury-eligible adult to understand, and that an alternative, simpler set of instructions might have yielded different results.⁸³

83. My argument in the text is based on the sheer complexity of the instructions, but the instructions may also have failed to elicit compliance because they are steeped in a style of thinking—consequentialist reasoning—with which participants were unfamiliar. The participants in this experiment were surely more accustomed to thinking in terms of punishment than deterrence; as the authors of the preceding chapter write, “intuitive punishment judgments are not tailored to consequentialist goals” (p. 135) and “[p]eople do not spontaneously think in terms of optimal deterrence” (p. 141). Yet rather than seek to explain this unfamiliar way of thinking in terms a typical adult might understand, perhaps by trying to reconcile the idea of deterrence with that of punishment-as-retribution or at least address jurors’ likely preference for the punishment objective of punitives, *see infra* p. 285 n.2, the Polinsky-Shavell instructions plow ahead and define *both* deterrence and punishment in consequentialist terms. Consider, for instance, the respective opening sentences of the explanations of the deterrence and punishment objectives. Deterrence:

Punitive damages fulfill the deterrence objective to the extent that they deliver a message and warning to the defendant and to other similarly situated firms to take appropriate steps to prevent harm in the future. But punitive damages will not fulfill the deterrence objective if they cause firms to take wasteful steps to prevent harm, if they cause the prices of products and services to rise excessively, or if they cause firms to withdraw socially valuable products or services from the market. (p. 165).

Punishment:

Punitive damages fulfill the punishment objective to the extent that they cause defendants to penalize their *blameworthy employees* who engaged in reprehensible behavior. In considering punishment, you should keep in mind that the defendant’s payment of compensatory damages already may lead to the punishment of blameworthy employees to some extent. (p. 166).

(The next paragraphs of the punishment instructions, one of which is quoted in the text, are also entirely consequentialist in nature.) Faced with this consequentialist onslaught and no place to put their retribution-based intuitions, some participants may very well have given up the attempt to follow the instructions to the letter. (Some may not have gotten started.)

Or perhaps some participants reacted against the instructions’ evident slant, not just toward consequentialism but against awarding punitives at all, by jettisoning the instructions altogether. (Of course, without any data in the form of participants’ think-aloud protocols or other appropriate measures of what they thought about the instructions, this hypothesis, like the one above, is entirely speculative.) A reasonable person could well believe that the thrust of the instructions was to discourage the awarding of punitives—both in blatant ways such as those quoted above, and in more subtle ways as well. For instance, concerning paragraph 3.b of the “punishment” instructions quoted in the text: since the imposition of tort damages on a company may always, in theory, be reflected in charges on “innocent parties” like shareholders (when company stock goes down) or customers (if prices are raised to cover liability costs), when *would not* jurors be expected to lower their awards? And note that what paragraph 3.b asks jurors to lower is not the “punishment amount,” which ought to be the topic in this part of the instructions, but the “punitive damages”—the ultimate award.

VII. DO JUDGES DO BETTER?

The preceding sections of this Review explain the extent to which the experiments reported in the book show that jurors' and juries' punitive damages awards and component judgments are biased in relation to the rational judgment and economic norms of good decision-making, respectively. With regard to each criterion it may fairly be asked whether the likeliest alternative decision-maker—the trial judge—would do any better. Research reported in chapter 11 indicates that, by and large, judges' punitive damages decisions are less prone to cognitive bias and more capable of accommodating the goal of optimal deterrence. For instance, judges, like jurors, displayed hindsight bias when evaluating the *ex ante* magnitude of a risk, but, unlike jurors, their liability judgments were not affected by knowing that the accident had occurred (pp. 189–90). Judges' decisions matched cost-benefit principles better than jurors' did, so that judges began to favor punitives only when the benefits of accident avoidance (not taken) far exceeded the costs (again unlike jurors, who seemed to favor punitives regardless of cost-benefit judgments) (pp. 191–92). And while judges, like jurors, overestimate low-probability mortality risks and underestimate high-probability ones, judges' estimates form a better overall fit with actual data (pp. 200–04). Unfortunately, the researchers do not seem to have asked judges to determine dollar awards,⁸⁴ so there is no way to compare the variability of their awards to that of jurors—a critical omission, since the excessive variability of awards is the authors' most serious charge against juries' punitive damages decisions.

There is a considerable amount of research elsewhere tending to show similarities rather than differences between jurors' and judges' decision-making processes and damages awards. In a recent, careful, and extensive review of this work (some of which she has herself conducted), Jennifer Robbennolt writes:

The existing empirical research investigating decision-making about other legal issues has shown that . . . judges and jurors have

84. The authors did so in experiment 3 in chapter 11, testing judges' ability when setting damages (in response to a very brief scenario) to screen out what might have occurred and focus only on actual losses (pp. 198–200), but the authors do not report any variability data. Interestingly, such data is reported in the article on which the relevant portion of chapter 11 is based. W. Kip Viscusi, *Jurors, Judges, and the Mistreatment of Risk by the Courts*, 30 J. LEGAL STUD. 107 (2001). There Viscusi reported that jurors' damage awards were more variable than judges', which would support the book's overall argument. See Robbennolt, *supra* note 23, at 150–57, for a discussion of this and other comparative studies.

similar responses to statistical information, and that judges, not unlike ordinary citizens, are unable to ignore inadmissible evidence and are vulnerable to cognitive illusions such as hindsight bias, anchoring, egocentric bias, framing, and the representativeness heuristic.

....

[T]he empirical results provide a somewhat mixed picture of possible differences between judges and jurors in the decisions they make about damages . . . [but a]t this early stage of the research there is no clear evidence to support the notion that judges will make qualitatively different decisions than juries across cases.⁸⁵

To their credit, the authors of the concluding chapters acknowledge some of this research, twice referring specifically to Robbennolt's paper (pp. 234 & n.37; 257 & n.28). Their assertions in chapters 11 and 12 regarding "[t]he superior performance of judges" (pp. 207, 235), on the other hand, contrast a bit more dramatically with what may reasonably be inferred from the bulk of the other research on this topic.

VIII. DO JURY DECISIONS TO AWARD PUNITIVES REFLECT IGNORANCE OF THE LAW?

Perhaps the strongest criticism that might be made of how juries decide to award punitive damages is that they do not conform to the legal norm—that juries do not follow the law. It is a serious criticism because it is generally (and reasonably) assumed that the likeliest alternative decision-maker (the judge) would do much better on this score, and because it is the criticism most likely to be appreciated by the public at large. An experiment already discussed (chapter 5) offers compelling evidence that juries' punitive damages decision-making does not fare well when measured by this standard.

In this study, over 700 jury-eligible adults deliberated in over 120 mock juries to decide whether to award punitive damages in one of four scenarios, based on actual cases in which the courts ultimately rejected punitives as a matter of law. As noted before, the majority of these juries voted to award punitives. Content analyses showed that the juries' attention to, use of, and understanding of the legal prerequisites for awarding punitives, as set forth in the judge's instructions,⁸⁶ were incomplete at best. On average the juries dis-

85. Robbennolt, *supra* note 23, at 146–47, 152, 154.

86. The elements counted, based on actual and typical punitive damages instructions (app. A at 259–60) included: (1) whether the defendant's conduct was malicious, or whether it was

cussed only 3.13 of the 5 possible elements, and drew conclusions about only 2.78 (p. 86). More importantly, juries that voted to award punitive damages were significantly less likely than those that voted not to award them to discuss or state conclusions about each of the 5 elements (p. 88). And the fewer the number of legal elements considered, the more likely a jury was to vote for punitives (pp. 88–89). Finally, 30% of the jurors did not get a single item right on a postverdict recall-comprehension test on the instructions, and the median score was only 5% correct (pp. 90–91).

A number of considerations may temper somewhat the apparent force of these data. Juries may have thought that some legal elements were too obviously satisfied to be worth discussing; thus, whether the dangerous outcome occurred was the most frequently ignored element (by over 50% of propunitives juries and over 30% of antipunitives juries). In this light, the average number of elements considered (3.13) or concluded (2.78) should perhaps be contrasted not to the ceiling of 5 but to something lower. The recall-comprehension test scores are so low, even by the usual fair-to-poor results obtained in most studies of juror comprehension of instructions, that perhaps the test instrument was unusually difficult and not a reliable measure of jurors' working understanding of the relevant concepts.⁸⁷ And the inverse correlation between thoroughness of discussion and likelihood of awarding punitives does not necessarily reflect a flaw or bias unique to the punitive damages decision: in any situation in which liability is premised on the satisfaction of all of the elements of a conjunctive rule, the failure to pay attention to any one or more elements will, all things being equal, increase the likelihood of a conclusion of liability.⁸⁸

reckless, where "recklessness" requires findings that (2) the defendant was subjectively conscious of a particular grave danger or risk of harm, which was a foreseeable or probable effect of the conduct, (3) the danger or risk in fact occurred, (4) the defendant disregarded the risk in deciding how to act, and (5) the disregard of the risk was a gross deviation from ordinarily reasonable behavior.

87. As was pointed out by Ellsworth, *supra* note 5, at 720–21. The authors address this external validity issue by contending that their findings are consistent with other studies of real and mock jury decision-making, and that, if anything, such features of their experimental design as making written instructions available during deliberations would have been predicted to yield higher levels of recall and comprehension than in the real world (pp. 91–92). Note also that recall-comprehension significantly correlated with predeliberation verdict preference for only one of the four cases (the better a juror's comprehension and recall, the more likely that juror was to vote against awarding punitives) (p. 91).

88. By the same token, the structure of the participants' decision-making task—a conjunctive rule, the proper application of which should have led to the conclusion that fewer than all elements were satisfied and therefore that the defendant was not liable for punitives, *see*

Finally, the significant negative relationship between the thoroughness of juries' attention to the legal requirements for a punitive award and their preference for punitives does not mean that ignorance of the law is the only explanation for their awards. Some of the juries that awarded punitives, and thus decided differently from the highest court to rule on the actual case, may have followed the law (at least, as well as did the juries whose judgments agreed with the actual decisions) but reasonably interpreted and applied it differently than the judge(s) did. To that extent, discrepancies between judges' and juries' punitive damage decisions may reflect legitimate, justifiable disagreements about the policies or values reflected in the law and its application to the case rather than, say, jury nullification or unwitting deviation from the legal norm.⁸⁹

These points notwithstanding, the results offer strong evidence that the processes by which juries reason their way to punitive damages decisions sometimes do not conform to the legal norm. The authors properly attribute this incomplete reliance on legal rules not to jurors' lack of motivation, but to their recourse, in the face of a difficult and complex judgmental task, to their everyday habits of thinking about responsibility and blame (pp. 92–93).⁹⁰ It is possible that the outcomes to which these common sense blaming habits lead jurors would largely coincide with those of judges or other legal experts; punitive damages juries could be “right for the wrong reasons.”⁹¹ But divergence from legally prescribed decision-making

discussion above,—did not allow the researchers to determine whether, in other situations, participants might make a different sort of error in failing to follow the law—*adding* an element to those specified in the instructions and thereby finding a defendant not liable who satisfied all of the legal elements of the rule. Cf. Vicki L. Smith, *Prototypes in the Courtroom: Lay Representations of Legal Concepts*, 61 J. PERSONALITY & SOC. PSYCHOL. 857 (1991) (scenarios in which a criminal defendant's conduct lacked certain elements of participants' prototypical conception of the crime charged yielded fewer guilty verdicts than scenarios containing all elements of the prototype, even though both scenarios met the legal definition of the crime). Of course, while such outcomes could offset the authors' implication that jurors are biased against defendants (at least when they decide punitive damages), they would compound the problem of juror inaccuracy relative to the legal norm. (I thank my colleague Steve Gilles for emphasizing this last point.)

89. See MacCoun, *supra* note 11, at 727; see also *supra* pp. 244, 249.

90. The authors mention, almost offhandedly, that jurors' “rough-and-ready everyday reasoning habits [were] probably influenced by their sympathies for one party or the other” (p. 93; see also p. 223). Unfortunately they do not appear to have sought the data that would have allowed them to test this, or any other, speculation about how jurors' emotions may have affected their decisions. Considering that in the authors' own model of punitive damages decision-making, an emotion—outrage—drives judgments (p. 32 fig. 2.1), it would have been worthwhile to study in greater depth the role of emotions in those judgments.

91. NEAL FEIGENSON, LEGAL BLAME: HOW JURORS THINK AND TALK ABOUT ACCIDENTS 104–11 (2000) (summarizing jurors' efforts to achieve “total justice” by using habits

processes could very well increase discrepancies between juries' punitive damages decisions and those of judges (p. 92) and/or other expert observers, lending support to calls for reform. I turn next to the book's concluding chapter, in which remedies are suggested for what the authors perceive to be the ailments of punitive damages juries.

IX. WHAT SHOULD BE DONE?

Cass Sunstein, the author of the book's last chapter, recommends two sorts of measures to address the arbitrariness and variability of juries' punitive damages awards, as described by the studies in the preceding chapters. The more modest is to encourage judges to exercise greater oversight of juries' punitive awards. In particular, the judge should make sure that the jury's award is in line with awards in similar cases (pp. 248–51). The more ambitious proposal is to replace juries as punitive decision-makers altogether with a schedule of punitive damages determined by “specialists in the subject matter at hand” (p. 258), analogous to workers' compensation and other administrative payment schemes (pp. 252–55). Sunstein also contemplates various intermediate “mixed approaches” (pp. 255–57).

The recommendation that a given punitive damages award be constrained by the range of awards in similar cases seems very sensible. While defining the set of similar cases may be a matter of dispute, almost any such set would provide at least a more relevant and meaningful anchor for the punitive dollar judgment than any currently used, and at most a determinate range for each award, thereby addressing major sources of current variability and arbitrariness (the unbounded scale, the absence of relevant anchors, and the presence of irrelevant ones). As long as counsel are permitted to help identify the comparison cases⁹² and to argue in exceptional circumstances for an award outside the range, it is hard to see how this

of thinking and feeling that often do not match the decision-making processes prescribed by law but usually yield results that are consistent with those of legal experts).

92. Sunstein does not explicitly state but implies that the judge would identify the comparison cases as part of his or her review of the jury's verdict. Without the lawyers' input, the selection of comparison cases would itself shift a great deal of punitive decision-making authority from juries to judges; certain groups of awards (those approved by appellate courts would presumably predominate) would likely become widely used, creating a kind of judicial punitive damages schedule (consisting of ranges rather than single numbers). Sunstein recognizes and applauds this; after all, even in this “modest” proposal he would be happy for judges to play an “exclusive” (p. 248) role in determining punitive dollar awards.

proposal would be less fair than the current system. But once the comparison cases have been chosen, there would not seem to be any particular reason for the judge rather than the jury to choose the appropriate dollar figure within that range, as Sunstein seems to prefer (pp. 248–49).⁹³ After all, the authors' own research (chapter 11) shows that judges as well as jurors are prone to some of the cognitive biases that affect punitive awards, albeit in some cases to a lesser degree,⁹⁴ and that above all, judges would share with jurors the "difficulties in mapping normative judgments onto dollar amounts" (p. 257).

The "more dramatic and radical shift" would be to move away from jury determinations of punitives entirely toward a damages schedule on the administrative model (p. 242):

As compared with the current approach, such a system would have many advantages. It would simplify matters, thus reducing the costs of litigation and decision. It would dramatically increase predictability. It should increase fairness, by ensuring that similarly situated people will be treated similarly. To the extent that officials seek to pursue optimal deterrence, a system of civil fines could be designed with that goal in mind (p. 253).

Sunstein mentions half a dozen other administrative schemes for determining compensation and/or penalties, including workers' compensation schedules, criminal sentencing guidelines, and social security disability determination grids, as precedents for his proposed "schedule of fines and penalties" (p. 252). Beyond stating that the punitives schedule "would be developed after extensive discussion

93. Sunstein acknowledges the possibility of judicial review of jury determinations as one of his "mixed approaches" (pp. 255–57).

94. Also, one of these studies (chapter 11, experiment 5) indicates that judges placed a much lower mean dollar value on life (\$3.6 million) (p. 206) than the \$5 million figure derived from labor market data (based on the implicit valuations workers make when they accept extra pay for higher-risk jobs) (p. 183); the mean valuation by jury-eligible participants responding to the same question was \$5.1 million (p. 184). Thus, juries' punitive awards would be more likely than judges' to reflect community sentiment (*cf.* pp. 40–41, where the authors point out that the supposed ability of a jury's punitive award to reflect community sentiment is seen as one of its major functions), at least to the extent that the level of outrage that drives the award depends on the value the community places on the well-being that the defendant's conduct threatened. I might add that in characterizing judges' and jurors' value-of-life figures, the author of chapters 10 and 11 (Viscusi) seems to engage in less than even-handed treatment. The jurors' mean value, which is almost exactly at the midpoint of economists' values, is described as "quite reasonable and in line with estimates in the literature" (p. 184), while the judges' values—which, as noted, are much further from that midpoint—"seem to be somewhat low, but [are] by no means outside of the range of estimated value-of-life statistics in the literature. . . . [T]he general order of magnitude of the [judges'] responses seems appropriate" (p. 205). This description reduces the disparity between the two sets of judgments and minimizes jurors' superior performance relative to the scale the author himself has posited.

within the democratic process, discussion informed by specialists in the topics at hand” (pp. 252–53) and would be overseen by administrative officials, he does not offer any more detailed description of the program.

This proposal raises three concerns. First, without a better idea of the factors that would be considered in setting awards and the dollar ranges that might be viewed as appropriate penalties for different kinds and degrees of reckless or malicious misconduct, it is difficult to say how well the proposal would accomplish its express purpose of correcting for the perceived arbitrariness and unpredictability of the current system. Even if the relevant factors are specified, there may be a considerable range of disagreement and hence unpredictability in judges’ or administrators’ applications of those factors.⁹⁵ Second, in the attempt to define penalties that are both appropriate and predictable, the scheme may become so complex and rigid that it disserves rather than promotes fairness—as has arguably been the case with the sentencing guidelines Sunstein offers as one of his models.⁹⁶ Third, the shift of authority in deciding punitives from juries to judges or administrators may be opposed on political grounds. Sunstein recognizes this and argues: “To those who think that this [proposal] is an unacceptably undemocratic measure, the best response might involve the workers’ compensation system[, which] . . . commanded an extraordinary consensus from those frustrated by the unreliable use of jury awards” (p. 255). The problem with this analogy is that workers’ compensation laws promised (and delivered) something to each camp: employers and their insurers got caps on damages, while injured workers and their families got more certain recovery.⁹⁷ By contrast, the proposed administrative scheme may very well give defendants their caps, but it does not seem to offer the injured any increase in the likelihood of receiving a punitive damages benefit (at least, one cannot discern from the information provided that such an effect is likely).

95. For this same reason, it is not clear whether the proposed system would indeed increase fairness, because the specified factors and/or the ways in which they are applied might subject people who are similarly situated (according to important definitions of “similarity” extrinsic to the schedule) to dissimilar treatment.

96. See, e.g., Kate Stith & José A. Cabranes, *Judging Under the Federal Sentencing Guidelines*, 91 NW. U. L. REV. 1247 (1997).

97. See Lawrence M. Friedman & Jack Ladinsky, *Social Change and the Law of Industrial Accidents*, 67 COLUM. L. REV. 50, 65–72 (1967).

One set of recommendations that Sunstein does not make might help reduce the perceived arbitrariness if not also the unpredictability of awards with a minimum of dislocation to current practices. This would involve amending jury instructions on punitive damages to make them clearer, and requiring special verdicts to encourage juries to pay attention to each and every legal prerequisite for making a punitive award. Current instructions are far from ideal, and the study reported in chapter 9, as discussed above, does not in any way support the conclusion that jurors would not be able to follow truly clear and simple instructions. Perhaps most importantly, current instructions do not even attempt to address jurors' likely preconceptions regarding punitives. Research has shown that while jurors' adherence to their everyday notions of blame and responsibility may impede them from complying with typical instructions, revised instructions that acknowledge those common sense conceptions and explain how the law differs elicit greater compliance.⁹⁸ Requiring juries to complete a special verdict form would not prevent them from deliberating as they like and then retrofitting their conclusions into the form, but it would be predicted to increase those measures of juror attention to the law that the authors' own research found most lacking (chapter 5).⁹⁹ Such improvements would be entirely compatible with the use of punitive damages awards in comparable cases as a constraint on, or at least anchor for, juries' decisions and/or judicial review thereof, as Sunstein recommends.

CONCLUSION: WHAT ARE PUNITIVE DAMAGES FOR?

Implicit in much of the research program and results reported in *Punitive Damages* and explicit in the concluding chapter is the authors' normative view of punitive damages. In this view, predictability and rationality are the primary virtues of punitive

98. Diamond & Casper, *supra* note 48 at 534, 558 (participants in antitrust case who were told not only that the law would treble their compensation awards but also the reason for the trebling and why they should not reduce their award to adjust for it gave larger (and thus more appropriate) awards than did those who were told only that their awards would be trebled or that the awards would be trebled but that they should disregard the trebling); Vicki L. Smith, *When Prior Knowledge and Law Collide*, 17 LAW & HUM. BEHAV. 507 (1993) (instructions designed to address and revise participants' pre-existing prototypical conceptions of crimes led to more accurate decisions than instructions that ignored those prototypes or told participants to disregard them).

99. See Elizabeth C. Wiggins & Steven J. Breckler, *Special Verdicts as Guides to Jury Decision Making*, 14 LAW & PSYCHOL. REV. 1 (1990) (special verdicts somewhat increased comprehension of jury instructions, but participants' impressions of the parties rather than verdict form were strongest predictors of their decisions).

ability and rationality are the primary virtues of punitive awards, and optimal deterrence is the most important goal that they should strive to achieve.¹⁰⁰ These virtues and this goal are linked in the list of advantages, quoted earlier, that would accrue when erratic punitive damages juries, antipathetic to cost-benefit thinking, are replaced by administrators applying a predetermined schedule of civil penalties with the goal of optimal deterrence built in.

It is hard to imagine a liability system providing optimal deterrence without predictability. But one could easily have predictability without optimal deterrence—for instance, punitive damages could simply be prohibited, or, at the other extreme, every defendant found liable for punitives could be assessed a \$10 billion penalty. Yet what is obviously wrong with these schemes is not that they would be inefficient, though they probably would. It is that they would be unfair.¹⁰¹ Under neither regime would most malicious or reckless tortfeasors receive their just deserts. If tort law is going to be in the punishment business at all (as it is if *punitive* damages are allowed, or even a system of civil *penalties* as the authors recommend), then the punishments should be in some sense proportionate to the blameworthiness of the conduct, or the punishment scheme will lose its basis in the community's common sense of what justice requires.¹⁰²

100. I do not believe that I am overstating this point. It is true that Sunstein, having introduced the five chapters of studies on jurors and judges as risk managers with an explanation of optimal deterrence, goes on to say:

These ideas are controversial, and we do not, as a group, intend to take a position on them here. Some people insist that the purpose of punishment is to ensure retribution rather than optimal deterrence. We suggest only that many policy analysts believe that the task of the legal system is to create penalties that are high enough to produce adequate deterrence, but not so high as to produce overdeterrence" (pp. 109-110).

The authors' research program itself endorses the importance of the optimal deterrence goal—more than half of the descriptive portion of the book is devoted to studies designed to determine whether jurors are up to the task of optimal deterrence, and concludes that they are not (as compared to the other norms). Later, Sunstein notes that one of the advantages of his proposal to replace jury decision-making with an administrative damages schedule is that "[t]o the extent that officials seek to pursue optimal deterrence, a system of civil fines could be designed with that goal in mind" (p. 253); no such accommodation of the goal of retribution (for instance) is promoted, or sought.

101. Cf. California's "three strikes" law, mandating a minimum sentence of twenty-five years to life for a felony conviction if the defendant has previously been convicted of two or more violent or serious crimes. CAL. PENAL CODE § 667 (West 2002). The law has been widely criticized as unfair, and has led prosecutors in some counties to exercise their discretion in a way that restores some proportionality to the punishment scheme (but at the same time violates the uniformity and predictability sought by the "rule of law"). See Joshua E. Bowers, "The Integrity of the Game is Everything": *The Problem of Geographic Disparity in Three Strikes*, 76 N.Y.U. L. REV. 1164 (2001).

102. Experimental research confirms that lay people's punishment decisions are driven by just deserts rather than deterrence. Kevin M. Carlsmith et al., *Why Do We Punish? Deterrence*

This is precisely what the retribution goal of punitive damages both permits and requires. Retribution has to do with the wrongness of a defendant's conduct rather than (solely) the magnitude of the harm the defendant has inflicted (which, considered prospectively, is the concern of deterrence theory).¹⁰³ Punishment informed by the goal of retribution allows society to inflict "a publicly visible defeat on the wrongdoer," thereby reaffirming society's moral standards.¹⁰⁴ Retribution has traditionally been recognized as a purpose of punitives,¹⁰⁵ and the Supreme Court has approved of it.¹⁰⁶ The typical reference in punitive damages jury instructions to the reprehensibility of the defendant's conduct (*e.g.*, p. 13) makes sense only in terms of the retribution goal. In short, while retribution is not the only legitimate aim of punitive damages, it is certainly one such aim.¹⁰⁷

At the heart of any fair retributive scheme, as noted, is proportionality or just deserts.¹⁰⁸ And the best way to keep punishments attuned to the community's sense of just deserts is to maintain the community's role in administering those punishments. The constitutional recognition of jury decision-making attempts to ensure that the legal decisions to which it applies will reflect not only judicial and legal norms (embodied in the substantive and procedural law the judge administers) but also the community's values and common sense, as each jury reflects them. The authors would exclude the community's traditional role in self-governance as expressed in the

and Just Deserts as Motives for Punishment, 83 J. PERSONALITY & SOC. PSYCHOL. 284 (2002); see also NORMAN J. FINKEL, COMMONSENSE JUSTICE 334-37 (1995) (proportionality as hallmark of common sense justice).

103. Galanter & Luban, *supra* note 76, at 1432.

104. *Id.*

105. See Michael Rustad & Thomas Koenig, *The Historical Continuity of Punitive Damages Awards: Reforming the Tort Reformers*, 42 AM. U. L. REV. 1269 (1993).

106. *Pac. Mut. Life Ins. Co. v. Haslip*, 499 U.S. 1, 21 (1991).

107. Note that Sunstein, in the passage quoted in footnote 100 above, writes as if *either* retribution *or* optimal deterrence may be accepted as *the* goal of punitive damages. I see no reason why the legal system must opt for one or the other. Indeed, it has been argued that liability and punishments based on just deserts (as determined by the community's shared principles of justice) would create the best deterrence. Paul H. Robinson & John M. Darley, *The Utility of Desert*, 91 NW. U. L. REV. 453 (1997). Nor are the purposes that jurors impute to punitive damages limited to deterrence and retribution. For instance, jurors may also use punitives as compensation for injured plaintiffs (especially if they believe that the compensatory damages will not adequately serve this function, as some participants in the authors' research, who were not allowed to decide compensatories themselves, might well have believed) and/or as a way of achieving restorative justice. Michelle Chernikoff Anderson & Robert J. MacCoun, *Goal Conflict in Juror Assessments of Compensatory and Punitive Damages*, 23 LAW & HUM. BEHAV. 313 (1999).

108. Galanter & Luban, *supra* note 76, at 1432; Rustad & Koenig, *supra* note 105, at 1320.

determination of punitive damages awards and replace it with managerial expertise in the form of risk analysis. They facilitate this move by simply omitting the retribution goal from the normative vision supporting their appraisals of juries' punitive damages decisions. The authors' construction of a research paradigm based on the norm of optimal deterrence has both interest and value, as I hope I have recognized in previous sections of this Review. Their failure, however, to explain and justify their reliance on optimal deterrence *to the exclusion of retribution* in their evaluations of jury behavior—how big a “problem” juries' punitive damages awards are and what, if anything, ought to be done in response—is a major drawback of their work.¹⁰⁹

Once the authors take out of the picture what juries may be most capable of doing well (making particularized judgments of just deserts) and confine their evaluative standards to what juries are least equipped to do well (managing risks the way risk analysts would), it is not surprising that they find juries' punitive damages awards to be so flawed, especially in their variability. Reinstating the retributive goal puts the authors' focus on the predictability of awards in a different light. To the extent that one focuses, as *Punitive Damages* does, on optimal deterrence as the goal of punitives, considerable predictability may be both demanded and reasonably expected.¹¹⁰ In contrast, there are severe limits to how much predictability we can reasonably expect of decisions that are finely tuned to the particular facts of each case, and especially to nuances in the blameworthiness of the defendant's conduct and its consequences, as determinations of just deserts ought to be. Furthermore, the authors' characterization of jurors' punitive damages awards as arbitrary (relative to the norms they

109. The authors might also have confronted some of the theoretical, *see, e.g.*, Neal K. Katyal, *Deterrence's Difficulty*, 95 MICH. L. REV. 2385 (1997), and practical, *see, e.g.*, DON DEWEES, DAVID DUFF, & MICHAEL TREBILCOCK, *EXPLORING THE DOMAIN OF ACCIDENT LAW* (1996) (discussing equivocal evidence for deterrent effects of punitive damages and other tort liability rules in products liability, environmental, and other types of cases), difficulties in identifying what deterrence is optimal and whether punitives (or civil fines) might plausibly achieve it.

110. More generally, however, the authors may be demanding more predictability than a complex or even chaotic system like that of civil litigation, across case types and jurisdictions and over time, can possibly offer. Priest's statement in the introduction to the book—“it is an appropriate subject of concern to observe any important legal outcome that appears unpredictable and that cannot clearly be explained by principle” (p. 2)—is surely excessive as applied to individual punitive damages awards; as applied to the pattern of decisions as a whole, the degree of predictability that might reasonably be expected rises (that's just the law of large numbers) but the power of any one or more principles to account coherently for all outcomes radically shrinks (as Critical Legal Studies scholars and the Legal Realists before them established).

employ) should be augmented by an examination of how well those awards may serve the goal of just retribution.¹¹¹ The broader point, once again, is that judgments that may appear defective in terms of the economic norm may be quite rational and appropriate when viewed in terms of a different, entirely justifiable norm of jury decision-making.¹¹²

How well do those awards serve the retributive goal? I do not know, although I suppose that experimental research could help address the question.¹¹³ The authors of *Punitive Damages* are not to be faulted for not adding such studies to the exceptionally broad and fruitful research program they have already completed. However, had they sought to explain at greater length their emphasis on optimal deterrence to the exclusion of retribution and to justify the standards of predictability against which jurors' and juries' awards were measured and found wanting, they might have allowed readers to draw from their extensive and valuable data somewhat different conclusions about the state of punitive damages practice and the need to reform it.

111. Thus, Sunstein's assertion that "when there is a wide range of possible awards, and when differences among them cannot be explained by reference to factors in the case, the legal system is not complying with the aspirations of the rule of law" (p. 247) may be insufficiently sensitive to case factors that (properly) affect *retributive* responses, and may set too high a standard for the consistency required by the "rule of law."

112. See Tetlock, *supra* note 12, at 464.

113. The authors might argue that the extent of variability of juror or jury responses to *identical* scenarios that they found in their research indicates that those awards cannot be doing a very good job of implementing a rule or principle of proportionate justice. Such an inference would be premature. It could be that different jurors or juries could reasonably assign different weights to the arguably culpable aspects of a defendant's conduct, resulting in (justifiable) variability in *inputs* to the proportionality judgment, and/or that they could reasonably assign ranges of appropriate dollar sanctions for any given level of perceived reprehensibility, resulting in (justifiable) variability in *outputs* reflecting a proportionality judgment.