

University of Chicago Law School

Chicago Unbound

Journal Articles

Faculty Scholarship

2002

Is Incoherence Outrageous?

Cass R. Sunstein

Daniel Kahneman

Ilana Ritov

David Schkade

Follow this and additional works at: https://chicagounbound.uchicago.edu/journal_articles



Part of the [Law Commons](#)

Recommended Citation

Cass R. Sunstein, Daniel Kahneman, Ilana Ritov & David Schkade, "Is Incoherence Outrageous?," 54 Stanford Law Review 1293 (2002).

This Article is brought to you for free and open access by the Faculty Scholarship at Chicago Unbound. It has been accepted for inclusion in Journal Articles by an authorized administrator of Chicago Unbound. For more information, please contact unbound@law.uchicago.edu.

Is Incoherence Outrageous?

Cass R. Sunstein,* Daniel Kahneman,**
David Schkade*** & Ilana Ritov****

For their valuable comments, we are most grateful to Cary Coglianese,¹ Mark Kelman,² and the team of Theodore Eisenberg, Jeffrey Rachlinski, and Martin Wells (hereafter ER&W).³ ER&W offer a great deal of information about the real world of punitive damages, presenting a mixed verdict on our basic claims. Coglianese provides an impressively detailed account of the problem of achieving coherence in the world of administrative penalties. Kelman questions the importance of coherence and also urges that in many cases, people's judgments may be better if they lack a global perspective. We are pleased to say that we have few significant disagreements with the three commentaries. In brief, we think that ER&W's evidence is fully consistent with our basic claims, despite a conservative test; that Coglianese has elaborated helpfully on the problem of achieving coherence in civil penalties; and that Kelman's basic arguments should not be taken to dispute our claims, which are less different from his own than a reader of his commentary might infer. Apart from these points, which we elaborate below, we offer a more general suggestion, one that is perhaps insufficiently emphasized in our main paper: Even if it is demonstrated, incoherence usually is not taken to be outrageous. The persistence of incoherence, in many domains of the law, is closely connected with this psychological fact.

* Karl N. Llewellyn Distinguished Service Professor, University of Chicago.

** Eugene Higgins Professor of Psychology and Professor of Public Affairs, Princeton University

*** Herbert D. Kelleher Professor of Business, University of Texas, Austin.

**** Associate Professor, Hebrew University of Jerusalem.

1. Cary Coglianese, *Bounded Evaluation: Cognition, Incoherence, and Regulatory Policy*, 54 STAN. L. REV. 1217 (2002).

2. Mark Kelman, *Problematic Perhaps, but Not Irrational*, 54 STAN. L. REV. 1273 (2002).

3. Theodore Eisenberg, Jeffrey J. Rachlinski & Martin T. Wells, *Reconciling Experimental Incoherence with Real-World Coherence in Punitive Damages*, 54 STAN. L. REV. 1239 (2002).

I. EXPERIMENTS AND THE REAL WORLD

ER&W provide an instructive look at the real-world of punitive damage awards. The most novel part of their essay offers some intriguing data analysis, suggesting that the incoherence that we describe actually exists, and hence that there is a “degree of support”⁴ for our claims in real-world data. But thinking that few “complaints about a legal system resonate louder than charges of incoherence,”⁵ ER&W object that if our argument is right, we should “be observing a legal system on the verge of collapse,”⁶ because any “society would find a torrent of incoherent legal judgments to be unfair, unjust, unpredictable, and unmanageable.”⁷ And in a discussion of prior findings that appears to be in some tension with their new analysis, ER&W contend that the system of punitive damages is actually quite coherent, showing a high degree of predictability. Of the three comments, ER&W’s raises the most complex issues, and we shall discuss those issues in some detail.

A. *A New Analysis*

ER&W’s innovative empirical analysis is based on their (welcome) claim that our essay can be used to generate testable hypotheses. They offer three such hypotheses. The first involves juries. In their view, we would predict that juries, which make decisions in isolation, would produce, across diverse categories of cases, the same statistical relationship between punitive awards and compensatory awards. The second hypothesis involves judges, who observe cases from many categories. ER&W think that we would predict that judges will be engaging in cross-category comparisons and hence that the effects of case categories “should be greater in judge-trying cases than in jury-trying cases.”⁸ Perhaps most interestingly, ER&W urge that if we are right, judges will offer higher awards than juries in cases involving the more prominent category of harm. The reason is that in cross-category comparisons, the awards in the more prominent category typically increase; and judges, in ER&W’s view, are engaged in cross-category comparisons. ER&W rightly say that the third hypothesis seems counterintuitive, even a bit odd. Could it be right to suggest that when the harm is prominent, judges will produce higher damage awards than juries?

Comparing actual awards in cases involving bodily injury to actual awards in cases not involving bodily injury, ER&W find support for all three hypotheses. For juries, the ratio of compensatory to punitive awards is about

4. *Id.* at 1243.

5. *Id.* at 1239.

6. *Id.* at 1240.

7. *Id.*

8. *Id.* at 1249.

the same between cases involving bodily injury and cases not involving bodily injury. For judges, punitive awards have a far higher ratio to compensatory awards in cases involving bodily injury. Most strikingly, the counterintuitive prediction is vindicated. The punitive award is likely to be much higher, compared to the compensatory award, when judges are making the decision than when juries are doing so.

ER&W have done a very creative job of adapting our account of punitive damage assessment to produce testable hypotheses for their extensive data set. We are of course delighted that our model passed the tests that they set for it. Still, we should point out that these tests are overly severe and that our model might have failed them even if it is basically correct. Three points deserve mention. First, ER&W's analyses use compensatory damages as a proxy for the "outrage \times harm" component of our model. But in our model, as in the law, compensatory damages are linked only to the harm, not the outrageousness of the defendant's behavior. The outrageousness (often called egregiousness) of the defendant's behavior is not represented in the ER&W analyses, though it is central to our model of how punitive damages are set. Second, ER&W's analysis takes for granted that judges consider each case in a broader context, one that includes cases from other categories. What is certainly true, and what we emphasized, is that judges *could* consider cases through a wider viewscreen because, unlike juries, they have encountered a varied set of cases. However, we are not aware of any direct evidence that judges *do* spontaneously consider particular cases in a broad context. Indeed, our original paper drew on work by Guthrie, Rachlinski, and Wistrich in raising doubts about precisely that point.¹⁰

The Supreme Court apparently has similar doubts about the width of the screen judges use. In *Leatherman* the Court found it necessary to *require* even appellate judges to engage in an explicit comparison (i.e., to use a wide screen) in evaluating the constitutionality of damage awards, rather than trusting their experience.¹¹ Finally, ER&W's analyses use the contrast between cases with

9. Cass R. Sunstein, Daniel Kahneman, David Schkade & Ilana Ritov, Predictably Incoherent Judgments, 54 STAN. L. REV. 1153, 1166 (2002).

10. See Chris Guthrie, Jeffrey Rachlinski, & Andrew Wistrich, *Inside the Judicial Mind*, 86 CORNELL L. REV. 778 (2001).

11. See *Cooper Indus., Inc. v. Leatherman Tool Group*, 532 U.S. 424, 440 (2001). Differences in the institutional competence of trial judges and appellate judges are consistent with our conclusion. In *[BMW of North Am. v.] Gore* [517 U.S. 559 (1996)], we instructed courts evaluating a punitive damages award's consistency with due process to consider three criteria: (1) the degree or reprehensibility of the defendant's misconduct, (2) the disparity between the harm (or potential harm) suffered by the plaintiff and the punitive damages award, and (3) the difference between the punitive damages awarded by the jury and the civil penalties authorized or imposed in comparable cases. Only with respect to the first *Gore* inquiry do the district courts have a somewhat superior vantage over courts of appeals, and even then the advantage exists primarily with respect to issues turning on witness credibility and demeanor. Trial courts and appellate courts seem equally capable of analyzing the second factor. And the third *Gore* criterion, which calls for a broad legal comparison, seems more suited to the expertise of appellate courts. Considerations of institutional competence therefore fail to tip the balance in favor of deferential appellate review.

and without physical injury as a proxy for the prominence of the case category. While certainly a reasonable operationalization, this classification is imperfect since the “non-physical” class includes many highly prominent categories (as possible examples, consider sexual harassment and racial discrimination¹²), again limiting our hypotheses’ chances to succeed. In sum, because our account of punitive damage assessment succeeded on this somewhat unfavorable terrain, what ER&W call “modest” support should probably be interpreted as much stronger.¹³

One further point. ER&W contend that we predict that case categories will play *no* role in jury cases; but the claim strikes us as too strong, and it is their speculation, not ours. Indeed, compensatory awards typically differ among categories, and hence categories will matter insofar as compensatory awards operate as anchors. In our study compensatory damages were held constant. But in an earlier study, mentioned briefly in our article, compensatories were not constant in the isolation condition, and we found a *large* difference between categories in the isolation condition.¹⁴

B. *Predictability and Coherence*

Notwithstanding the evidence just discussed, ER&W urge that “research on quantitative judgments demonstrates substantial consistency in judgments.”¹⁵ Drawing largely on an impressive and influential study by Eisenberg, Wells, and three co-authors,¹⁶ ER&W claim that the legal system has proved its “ability to reduce and manage the limitations of human judgment.”¹⁷ In our view, the real-world data do not support this optimistic conclusion, at least if the optimism is applied to jury determinations.¹⁸ We have noted that in *Leatherman*, the Supreme Court found the risks of arbitrary jury outcomes sufficiently troubling to require *de novo* review of the constitutionality of jury awards by appellate courts. Further, over half of punitive damage awards are

Id. (emphasis added, citations and footnote omitted).

12. We describe these as possible examples because we lack evidence on what categories of “nonphysical” injury count as prominent.

13. See Deborah Prentice & Dale Miller, *When Small Effects Are Impressive*, 118 PSYCHOL. BULL. 112, 160-164 (1992) (giving an instance in which small effects are impressive).

14. See Cass R. Sunstein, Daniel Kahneman, David Schkade, & Ilana Ritov, *Legal Coherence and Incoherence* (2001) (unpublished manuscript, on file with author).

15. Eisenberg, et al., *supra* note 3, at 1241.

16. See Theodore Eisenberg, John Goerdt, Brian Ostrom, David Rottman & Martin T. Wells, *The Predictability of Punitive Damages*, 26 J. LEGAL STUD. 623, 633-37 (1997).

17. Eisenberg, et al., *supra* note 3, at 1242.

18. See W. Kip Viscusi, *Why There Is No Defense of Punitive Damages*, 87 GEO. L.J. 381, 384-87 (1998).

reduced or overturned.¹⁹ The process of appellate review can impose on all parties involved substantial transaction costs and lead to situations in which justice delayed becomes justice denied.

Before explaining the problem, we emphasize our agreement with many of ER&W's conclusions. Punitive awards are indeed relatively rare. They are especially uncommon in cases lacking intentional misconduct. Significantly, the compensatory award plays a substantial role in determining the punitive award. For this reason punitive awards are far from entirely unpredictable. But ER&W downplay the existence of a great deal of noise in the system. The studies that purport to show predictability find that the compensatory award explains 46-47% of the variation in punitive damage awards, an impressive figure to be sure, but one that leaves the majority of the variation unexplained. The claim of predictability is further weakened by the fact that in the relevant work, the focus of analysis is not on actual punitive damage awards, but on the *logarithm* of such awards, which is the dependent variable that is being explained. Any transformation of real awards into log awards will naturally shrink much of the variation in the size of punitive damages awards. But defendants pay their penalties in dollars, not in log dollars, and the uncertainty they face when a punitive award could range from \$10,000 to \$1,000,000 is rather more impressive than is conveyed when the same range is expressed in log units (7 vs 9). With respect to predictability, judicial oversight and appellate review can and do help, but we respectfully disagree with ER&W's suggestion that the legal system has overcome the cognitive problems that we have described. Consider the fact that another empirical study, based on a larger sample than that used by ER&W, finds that punitive awards "are highly variable and unpredictable."²⁰

But coherence, not predictability, is our focus here. To the extent that ER&W are able to establish predictability, they have not established coherence. To see this point, begin with the fact that because of the difficulty of translating moral judgments into dollars, anchors will loom very large. The compensatory award is the most readily available anchor, and in light of the translation problem, it should be no surprise that that award helps to explain the punitive award. And when people are making judgments in isolation, the anchor will have an especially big effect, indeed a bigger effect than it has when people look at two or more cases at once. Here, then, is the problem. Because

19. See, e.g., Marc Galanter, *Real World Torts: An Antidote to Anecdote*, 55 MD. L. REV. 1093, 1115-16 (1996) (citing studies finding that jury punitive damage awards are reduced in fifty percent or more of product liability cases).

20. See Jonathan Karpoff & John Lott, *On the Determinants and Importance of Punitive Damage Awards*, 42 J.L. & ECON. 527, 571 (1999). The authors do find that compensatory awards, alongside other identifiable factors (such as lawsuit type, firm characteristics, and location of suit), can explain as much as 50% of the variation in punitive awards. But without prior information that there will be a punitive award, "only 1-2 percent of the variation in punitive awards can be explained." *Id.*

compensatory awards in certain categories of cases (for example, commercial fraud cases) are typically larger than compensatory awards in other categories of cases (for example, personal injury cases), punitive awards will be larger in the former set of cases as well, regardless of the relative outrageousness of cases from the two categories. The existing research shows that this difference is entirely predictable. *But that very difference is a form of predictable incoherence in the sense that we discuss.* The reason is that people are most unlikely to approve of the pattern of outcomes that emerges from their own judgments in isolation. And one reason for the disapproval is that anchors have a larger effect on judgments in isolation than on comparative judgments. We can now see that if, as ER&W urge, compensatory awards are acting as the anchor, people's judgments will appear to make sense on their own, but will not make sense, by people's own lights, when taken together.

In this light, it is possible to dissolve the tension between the apparently conflicting empirical claims reported by ER&W. First, punitive awards do indeed have some degree of predictability (although that predictability is more impressive for scholars who think in terms of logarithms than for plaintiffs and defendants, who think about actual dollars). Second, punitive awards can be shown to be incoherent, in the sense that people's decisions in isolation do not match their decisions if they are engaging in cross-category comparisons. The two claims are not in conflict.

C. *Public Indifference*

Now let us turn to ER&W's broadest claim. They contend that if incoherence is pervasive, the legal system would be "on the verge of collapse,"²¹ and people would be up in arms. The "paradox" lies in the fact that "the system seems to work well, despite the well-founded concerns"²² of our article.

We do not think that there is any paradox. Because people do not naturally seek coherence, it seems to us wrong to suggest that incoherence, in the domain of punitive damages, would produce any kind of public outcry. Suppose, for example, that a presidential candidate of either party, armed with sensational data, wanted to make an energetic and dramatic campaign speech, in which he demonstrated, beyond a shadow of a doubt, that punitive awards in some categories of cases were far too low when taken in comparison with awards in other categories of cases. If you were the candidate's campaign manager, would you urge him to deliver that speech? Would the speech be taken to have uncovered an outrage? Would the candidate make a public splash? We doubt it. For the same reason that people do not spontaneously worry over coherence, incoherence, when brought to public attention, is a reason for embarrassment

21. Eisenberg, et al., *supra* note 3, at 1240.

22. *Id.* at 1241.

and possibly reform—but hardly an outrage. Because of the relative indifference to incoherence, the system may *appear* “to work well”²³ even when it in fact does not.

II. ACHIEVING COHERENCE

Like ER&W, Coglianese is concerned to connect our arguments to real-world evidence about law and policy.²⁴ But while ER&W concentrate on punitive damage awards, Coglianese examines administrative penalties. His overall goal is to show that incoherence cannot be established simply by finding disparities in congressionally-specified penalty amounts. It is necessary to ask many more questions.

Coglianese says, rightly, that our discussion of administrative penalties is focused on what he calls the problem of “comparative coherence.”²⁵ He insists that seemingly inexplicable variations in dollar amounts are only partly informative. Suppose, for example, that those who injure birds are subject to a federal penalty of \$20,000, but that those who subject employees to dangerous working conditions are subject to a federal penalty of \$10,000. Is this proof of incoherence? Not necessarily. For example, employers may be subject to stiff, additional penalties under state law while those who injure birds may not be; the lower federal penalties may make a great deal of sense once federal and state penalties are taken together. It is also possible that the government regularly enforces the latter statute, and infrequently enforces the former—and hence the disparity in penalties is sensibly responsive to varying enforcement levels.

Coglianese offers an impressive account of the many variables that must be examined in order to reach firm conclusions about coherence. He is correct to insist that judgments of incoherence might themselves reflect a form of bounded rationality—a failure to take account of considerations (such as additional penalties and the level of enforcement activity) that actually matter a great deal. But we are sure that Coglianese would agree that people, including policymakers, rarely attempt to make anything like the necessary assessment. It would be heroic, and in our view implausible, to think that the existing array of administrative penalties reflects considered judgments about the factors that Coglianese rightly identifies. So too with existing variations in cost-per-life-saved. As Coglianese urges, no sensible person could urge that every program for risk reduction should have the same cost-per-life-saved. Contextual factors, including distributional issues, matter a great deal. But incorporating contextual factors is most unlikely to rescue the current system from the charge of incoherence.

23. *Id.*

24. Coglianese, *supra* note 1.

25. *Id.* at 1219.

Coglianesi raises the possibility that after all relevant sanctions have been imposed, there will be less incoherence than it would seem if we focus on only one type of penalty. But we think that Coglianesi would agree that incoherence is likely to be substantial, all the more in light of the fact that the greater the number of different proceedings that must be initiated to achieve a coherent penalty, the less likely it is that a sensible overall outcome will be achieved. The execution of a complex strategy is harder than a simpler one, and a complex task will inevitably have higher transaction costs. Indeed, we greatly doubt that administrative penalties are in fact set through deliberate consideration of all the variables to which Coglianesi draws attention. The relative coherence of penalties within categories itself suggests that these variables do not play a major role, simply because there is no reason to assume that the variables in question are kept constant within categories. For example, if we compare a violation of the African Elephant Conservation Act²⁶ and a violation of the law protecting Bald and Golden Eagles (carrying the same penalty),²⁷ it is not obvious why the existence of other liability, or the availability of other remedies to ensure compliance, or the probability of detection, and so forth, would be the same in both cases.

One of the most striking features of Coglianesi's discussion is the support that it gives to our claim that full coherence is exceedingly difficult or perhaps impossible to achieve in practice (assuming we could agree on what it means in principle). Nonetheless, we think that much can be done to reduce the most serious anomalies. This point brings us directly to Kelman's paper.

III. PRESCRIPTIVE, NOT NORMATIVE

Our major concern with Kelman's interesting discussion²⁸ is that he has taken issue with claims that we do not make. Contrary to Kelman's suggestion, we do not believe that category-bound judgments are always inferior to the judgments made in cross-category comparisons. And in analyzing punitive damage reform, our approach is prescriptive rather than normative. We seek to show how certain social goals might be obtained, without saying anything about the appropriate goals of punitive awards. We attempt to dispel four misunderstandings here.

First, we do not say (and do not believe) that those who make category-bound judgments are "irrational."²⁹ Indeed we do not mean to make any controversial claims about the nature of rationality. Our more cautious suggestions are that category-bound thinking does give rise to incoherence, in the sense that we describe it, and that incoherence is indeed a problem. We

26. 16 U.S.C. §§ 4201-4246 (2002)

27. 16 U.S.C. § 668 (2000)

28. Kelman, *supra* note 2.

29. *Id.* at 1275.

plead innocent to Kelman's charge that we have offered an "unduly casual invocation of the 'rationality' norm."³⁰ We have no view about Kelman's claim that law students are rational to be intensely disappointed when they do not receive job offers from their preferred law firm; and nothing we say speaks to his suggestion that intense emotional reactions are generally rational if they promote people's subjective ends.

Second, we do not think, as Kelman suggests, that "the global perspective is always optimal."³¹ We meant to be very tentative here, suggesting only "that most people, in most conditions, prefer a wide viewscreen to a narrow one, and will in principle (though not always in actual practice) have greater confidence in the judgments they make in a broad context rather than in a narrow context."³² We do not disagree with Kelman's suggestion that in some circumstances, people's judgments may be better when and because they are category-bound. (Certainly we do not mean to say anything about "the appropriate response to misbehavior by one's child"³³)

Third, Kelman misunderstands our argument in objecting that we emphasize coherence—what he calls "horizontal equity"³⁴—at the expense of other social goals, and in insisting that it might be sensible to sacrifice coherence for the sake of ensuring that punishments are not too severe or too lenient. We agree with Kelman's suggestion that horizontal "equity is at best but one value."³⁵ We said that "coherence is not a trumping value, and a system displaying incoherence may well be better than one that is coherent but pervasively unjust."³⁶ We urged only that at "the very least, efforts should be made to correct the most conspicuous anomalies—a goal that can be obtained without thinking that it is easy or even possible for people to agree on what full coherence actually requires."³⁷

Fourth, Kelman is mistaken to say that according to our argument, "we should assess punitive damages in accord with a sense of outrage because we typically do."³⁸ He urges (far more modestly than Hume) that "it is never clear why we should derive an 'ought' from an 'is';"³⁹ he adds that perhaps punitive damages should not be understood in terms of outrage at all. But we did not intend to accept any particular normative theory of punitive damages. We certainly did not mean to express any view on the retributive view of

30. *Id.* at 1278.

31. *Id.* at 1275.

32. Sunstein et al., *supra* note 14, at 1179.

33. Kelman, *supra* note 2, at 1276.

34. *Id.* at 1280.

35. *Id.*

36. Sunstein et al., *supra* note 14, at 1203.

37. *Id.* at 1203.

38. Kelman, *supra* note 2, at 1289.

39. *Id.*

punishment. Our goals, here as in other work, are descriptive and prescriptive rather than normative: Why are awards what they are? Assuming that the legal system has as a certain goal, how does it best get there? If optimal deterrence is the goal, it does not make much sense to rely on juries. If the goal is to reduce incoherence while retaining a large role for retributive thinking, the legal system might attempt to use additur and remittur to increase anomalously low awards and to decrease anomalously high ones. But as we also say, a society “should be willing to tolerate a degree of predictable incoherence if this is the price of avoiding a (coherent) set of biased and unjust awards.”⁴⁰

IV. OUTRAGE

Much of our work on punitive damages has been concerned with the topic of outrage. In our view, outrage is associated with an urge to punish, which is the foundation for such awards.⁴¹ Jurors’ reactions to misconduct are rooted in judgments about whether the defendant has done something that qualifies as genuinely outrageous. We have also shown that outrage itself is category-bound. What appears outrageous in isolation may appear barely objectionable when placed in the context of far more egregious wrongdoing.

Our investigation has brought new clarity to a not unrelated point: Within-category incoherence is easily noticed and arouses a strong sense of injustice. Incoherence across categories is harder to see and less impressive when it is seen. People do not spontaneously question the coherence of their judgments across category boundaries, and are not greatly exercised by incoherence when it is drawn explicitly to their attention. Nonetheless, incoherence can be a serious problem, indeed a source of real injustice. We are not exactly outraged. But we think that much more can, and should, be done to reduce the problem.

40. Sunstein et al., *supra* note 14, at 1183.

41. Indeed, the legal criteria for awarding punitive damages seem almost a recipe for outrage, for example, “reckless disregard for the welfare of others” or “gross deviation from the normal standard of care.” CASS R. SUNSTEIN, REID HASTIE, JOHN W. PAYNE, DAVID A. SCHKADE & W. KIP VISCUSI, PUNITIVE DAMAGES: HOW JURIES DECIDE 259 (2002) (quoting punitive damage instructions).