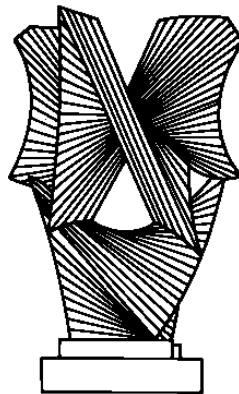


# CHICAGO

JOHN M. OLIN LAW & ECONOMICS WORKING PAPER NO. 131  
(2D SERIES)



## Predictably Incoherent Judgments

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

THE LAW SCHOOL  
THE UNIVERSITY OF CHICAGO

This paper can be downloaded without charge at:  
The Chicago Working Paper Series Index:  
<http://www.law.uchicago.edu/Lawecon/index.html>

The Social Science Research Network Electronic Paper Collection:  
[http://papers.ssrn.com/paper.taf?abstract\\_id=279181](http://papers.ssrn.com/paper.taf?abstract_id=279181)

## Predictably Incoherent Judgments

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

### Abstract

*When people make moral or legal judgments in isolation, they produce a pattern of outcomes that they would themselves reject, if only they could see that pattern as a whole. A major reason is that human thinking is category-bound. When people see a case in isolation, they spontaneously compare it to other cases that are mainly drawn from the same category of harms. When people are required to compare cases that involve different kinds of harms, judgments that appear sensible when the problems are considered separately often appear incoherent and arbitrary in the broader context. Another major source of incoherence is what we call the translation problem: The translation of moral judgments into the relevant metrics of dollars and years is not grounded in either principle or intuition, and produces large differences among people.. The incoherence produced by category-bound thinking is illustrated by an experimental study of punitive damages and contingent valuation. We also show how category-bound thinking and the translation problem combine to produce anomalies in administrative penalties. The underlying phenomena have large implications for many topics in law, including jury behavior, the valuation of public goods, punitive damages, criminal sentencing, and civil fines. We consider institutional reforms that might overcome the problem of predictably incoherent judgments. Connections are also drawn to several issues in legal theory, including valuation of life, incommensurability, and the aspiration to global coherence in adjudication.*

**“Why didn’t the Commission sit down and really go and rationalize this thing . . .? The short answer to that is: We couldn’t. . . . Try listing all the crimes that there are in rank order of punishable merit. . . . Then collect results from your friends and see if they all match. I will tell you they don’t.”<sup>1</sup>**

---

\* 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. For helpful comments, we are grateful to Eric Posner, Richard Posner, and participants in workshops at the John F. Kennedy School of Government at Harvard University, the University of Chicago Law School, and Yale Law School.

<sup>1</sup> Justice Stephen Breyer, as quoted in *The New Republic*, June 6, 1994, at p. 12.

## I. Introduction: Coherence and Incoherence

### A. Basic Claims

Coherence in law is a widely shared ideal. Almost everyone hopes for a legal system in which the similarly situated are treated similarly. But there are many obstacles to the achievement of coherence in the law. This Article is concerned with one particular test of coherence, and with two cognitive limitations that help cause many failures of that test in actual legal systems. We believe that these failures are also failures of justice, and that they suggest a pervasive problem in existing legal systems.

Our test of coherence is straightforward. We ask: When two or more judgments have been made separately, and each seems to make sense on its own, do they still make sense when considered together? When this test fails, and a need is felt to adjust or reverse the judgments that were made separately, we will speak of *judgment reversals*. The test of coherence can be readily applied to decisions by juries and by judges.<sup>2</sup> With suitable modifications, it can also be applied to acts of legislators and regulators. More generally, we ask whether judgments made in isolation fit together in an orderly way when considered as part of the larger whole.

Our emphasis will be on many categories of harms with which the law is concerned, such as physical injury, commercial fraud, and ecological damage. Our first psychological observation is that in law, as in ordinary life, people's thinking is *category-bound*. People do not easily cross the boundaries of categories<sup>3</sup> of harms in their thinking. When they consider an individual case of physical injury, or commercial fraud, the frame of reference for evaluation is usually a set of instances of the same kind of harm.<sup>4</sup> When setting penalties for a category of cases, such as violations of regulations for occupational safety, regulators will naturally focus on instances that belong to that category. They are much less likely to concern themselves with the consistency of their determinations with punishments for other categories of harmful conduct, such as damage to endangered species. Yet, as we will show, simultaneous

---

<sup>2</sup> Note that this is a test of coherence, not of social consensus, which was explored in Cass R. Sunstein, Daniel Kahneman & David Schkade, *Assessing Punitive Damages*, 107 *Yale LJ* 2071 (1998). But the absence of consensus on dollar amounts, discussed below, does bear on the possibility of achieving coherence.

<sup>3</sup> We say a few words below on the nature of categories.

<sup>4</sup> Daniel Kahneman and Dale Miller, *Norm Theory: Comparing reality to its alternatives*, 93 *Psychological Review* 136 (1986).

consideration of penalties for different kinds of infractions will often reveal that the more severe punishment was assigned to the misconduct which, in context, appears to be the less serious.

A second significant source of incoherence is what we shall call the ***translation problem***. By this term, we refer to the distinctive problem involved in translating a moral judgment of some kind<sup>5</sup> into the terms made relevant by the legal system, such as monetary penalties, civil fines, or criminal punishment.<sup>6</sup> We argue that the act of translation causes serious problems, because it is grounded neither in agreed-upon principle nor in widely shared intuitions. Even when people show coherent and consistent moral intuitions, they may show little consistency and coherence in translating those intuitions into numbers, such as dollars of fines or years in jail. Because of the translation problem, coherence fails: there is no guarantee that the relative severity of punishments administered by the system will still appear sensible, just, or fair when several punishments are considered together. The translation problem helps identify the cognitive foundations of current controversies over criminal sentencing, punitive damages, and contingent valuation. It affects the work not only of juries, but also of legislative and regulatory bodies that determine punishments for different kinds of misconduct within a particular category. The result, we will argue, is that the overall level of penalties set by different regulatory agencies may appear sensible when each set of regulations is considered on its own, but fail our test of coherence when several sets are considered at once.

We consider it self-evident that if it exists, incoherence in punishments is a form of injustice. We shall also assume that when the public would not believe that outcomes fit sensibly together, this is a problem that calls for social response. We will assume that these points are correct, without defending them in any detail.

The coherence and incoherence of punishments, both civil and criminal, will be the focus of our analysis, and in this domain we will attempt to show considerable reason for concern. Juries typically assess cases in isolation; in fact

---

<sup>5</sup> We use this term because as a matter of fact, moral judgments appear to be the foundation of punishments. On the economic theory of punishment, optimal deterrence is the goal, and optimal deterrence is in conflict with ordinary intuitions, as we discuss in detail below. A. Mitchell Polinsky & Steven Shavell, *Punitive Damages: An Economic Analysis*, 111 Harv. L. Rev. 869, 870–76 (1998). For interesting reasons, the translation problem is not a serious one from the standpoint of economic analysis of law. We will refer to this point below.

<sup>6</sup> We initially discussed the translation problem, under the technical term “scaling without a modulus,” in the particular context of punitive damage awards in Sunstein, et al., *supra*. We generalize the problem here.

lawyers are actually barred from referring to awards in other cases. Administrators, and congressional committees setting up penalties for regulatory misconduct, typically deal with one category of misconduct, and do not attend to problems of other types.<sup>7</sup> Criminal sentences are established over time by different legislatures and legislative committees, with little effort to ensure a good fit of penalties to crimes across a broad frame of reference.<sup>8</sup> Because people are not inclined to consider the overall pattern – either because it is too difficult to do so or because it does not occur to them to try – the problem of incoherence does not naturally receive attention.<sup>9</sup> As things now stand, the structure of those institutions charged with making regulatory and legislative decisions reinforces the effects of category-bound thinking. Scandalously large inconsistencies can therefore persist indefinitely, in the absence of a special effort to impose coherence.

The fact that coherence cannot be taken for granted has significant implications for institutional design. It suggests, in some domains, a possible reason to favor judicial decisions over jury decisions, because judges are more likely to have a menu of cases before them. Because judges are human,<sup>10</sup> they too are susceptible to producing incoherent patterns<sup>11</sup>; but especially if the risk of incoherence is brought to judicial attention, they might well likely to do better, on this count, than juries. Our claims also suggest the potential value of “coherence commissions,” assigned the explicit mission to ensure that decisions fit together as an orderly whole, or at least to correct the most serious anomalies. We bring the idea of “coherence commissions” in contact with many areas of the law, including civil and criminal penalties, punitive damages, and valuation of statistical lives. We will also attempt to cast new light on some large topics in legal theory, including the aspiration to similar treatment of the similarly

---

<sup>7</sup> Indeed, an administrator or regulator who shows explicit concern with problems that lie outside his or her jurisdiction risks being criticized as improperly “poaching” on someone else’s turf, or even having his decision held unlawful. See *Pension Benefit Guaranty Corp. v. LTV Corp.*, 496 US 633 (1990) (allowing agency entrusted with pension guarantees not to consider other, relevant areas of law, and suggesting that any such consideration would violate the underlying statute).

<sup>8</sup> The goals of the United States Sentencing Commission are, on this dimension, quite unambitious. See below. In fact a version of the translation problem appears in the international domain, where criminal sentencing also shows a degree of incoherence. See Allison Marston Danner, *Constructing a Hierarchy of Crimes in International Criminal Sentencing*, *Virginia Law Review* (forthcoming 2001).

<sup>9</sup> Indeed we believe that this Article is the first to venture even a tentative assessment of the pattern of administrative penalties. See below.

<sup>10</sup> Chris Guthrie, Jeffrey Rachlinsk, Andrew Wistrich, *Inside the Judicial Mind*, 86 *Corn. L. Rev.* 778 (2001).

<sup>11</sup> This was an explicit rationale for the formulation of Sentencing Guidelines

situated, the twentieth-century movement toward bureaucracy,<sup>12</sup> and the general problem of “incommensurability.” Discussing several different kinds of incoherence, we identify some of the cognitive limits of the aspiration to global coherence in law,<sup>13</sup> while also pointing the way toward institutional reforms that could overcome those limits. We urge that “coherence commissions” could do a great deal to reduce existing injustice, in a way that would provide a twenty-first century analogue to important, but less ambitious, institutional developments in the twentieth century.

## **B. Specific Points**

In this Article, we will be covering many topics, some of them in considerable detail. For purposes of exposition, it will be useful to give an overview of the specific claims that undergird our general arguments about incoherence

- The moral intuitions of the public are firmly retributive in character.<sup>14</sup> The intensity of what we shall call “punitive intention” – the desire to punish wrongdoing – is influenced both by the outrageousness of an action and by the severity of the harm that the action caused.
- It is extremely difficult for people to translate punitive intentions into the terms made relevant by the legal system, such as fines or prison terms. That task is not rooted in shared intuitions, moral or otherwise, and outcomes can be largely arbitrary and unpredictable. Different juries may express the same punitive intention, but come up with quite different dollar awards. The bodies that set administrative punishments for particular categories of misconduct may differ widely in the general range of punishments that they choose, for no principled reason. State legislatures may produce widely varying punishments for the same crime, not because of different moral judgments, but simply because of the translation problem.
- If people are asked to assess cases that fall within a particular category of actions and harms, but are not asked to translate their punitive

---

<sup>12</sup> A helpful overview is Price Fishback & Shawn Kantor, *A Prelude to the Welfare State: The Origins of Workers’ Compensation* (1999); see also Jerry Mashaw, *Due Process in the Administrative State* (1983), for a valuable discussion in the context of social security disability determinations.

<sup>13</sup> Associated with Ronald Dworkin, *Law’s Empire* (1985).

<sup>14</sup> As recognized by the Supreme Court in *Cooper v Leatherman*, 532 U.S. S. Ct. (2001) at footnotes 5 and 11.

intentions into dollars or years, people’s judgments tend to be both shared and coherent.<sup>15</sup> It follows that if they are evaluating cases within each category, diverse people are likely to agree on how to rank a set of personal injury cases, business fraud cases, sexual harassment cases, or libel cases by their “punishable merit” (Justice Breyer’s term).

- When asked to evaluate a case separately and hence in isolation, people spontaneously proceed by comparing it against others falling within the same category.<sup>16</sup> Thus, for example, people’s responses to a case of business fraud will be generated largely by comparing that case to other cases of business fraud. If the case of business fraud involves extremely egregious misconduct that caused very severe harm in this frame of reference, people will be extremely outraged. The fact that there are other categories of cases, involving actions that people view as more evil and harms that they think more serious, will be effectively neglected. We call this effect “normalization.”<sup>17</sup>
- The requirement directly to compare cases drawn from different categories of harmful actions can cause large shifts in moral evaluations, punitive intentions and actual punishments, relative to the judgments of the same cases in isolation. Such shifts, which we generically label judgment reversals, provide a diagnostic indication of a breakdown of coherence.
- There is less consensus on the ranking of cases in the hierarchy of “punishable merit” when the cases involve different kinds of harms (e.g., personal injury vs. environmental damage) than when they belong to the same category. Harms from different categories may seem “incommensurable,” because they cannot easily be described in terms of the same dimensions: the question of how many animals died, for example, is relevant in one of these situations but not in the other. The difficulty of comparisons across categories of harms may be one of the cognitive sources of the difficulty pointed out by Justice Breyer.

---

<sup>15</sup> The claim is based on results reported by Daniel Kahneman, David Schkade & Cass R. Sunstein, Shared Outrage and Unpredictable Awards, 16 *J. Risk & Uncertainty* 49 (1998) and Sunstein et al., *supra*.

<sup>16</sup> See Kahneman & Miller, *supra*; see also Daniel Kahneman, Ilana Ritov & David Schkade. Economic preferences or attitude expressions? An analysis of dollar responses to public issues, 19 *Journal of Risk & Uncertainty* 220 (1999).

<sup>17</sup> *Id.*

- There is substantial consensus on the ranking of categories, including those created or used by the law. People agree, for example, that murder is worse than rape, than rape is worse than assault, and that assault is worse than libel. We suspect, however, that there is less social agreement on the ranking of categories than the on ranking of cases within categories.
- In spite of the difficulties of comparing categories of cases and cases across categories, there is in many cases sufficient consensus to permit a test of coherence, which examine whether judgments made in isolation are still retained when explicitly compared.

We illustrate these claims by investigating judgments in several kinds of domains – some of them experimental, some involving actual government practice. The first involves punitive damages awards. Here we show that people rank cases within a given category of harms in a consistent and coherent fashion, that their judgments about isolated cases are “normalized” according to the category in which they fall – but that these narrowly-based and easily-made judgments shift when people are compelled to consider the case in a broader context, by forcing a comparison to a case that involves harm of a different kind. The second kind of problem involves contingent valuation – an influential method of valuation by which people assign dollar values to goods that are not ordinarily traded on markets.<sup>18</sup> We observe a similar pattern here. The amount that the public will be willing to pay to prevent or correct some harm to a public good will be very different depending on whether the goods and the harms are viewed in isolation or in explicit comparison with harms to goods and harms from another category.<sup>19</sup> This point fortifies existing doubts about the rationality and reliability of the contingent valuation method.<sup>20</sup>

A third kind of problem involves administrative penalties. Investigating actual statutory practice, we find that such penalties tend to make a great deal of sense within categories, because (for example) the more serious occupational safety and health violations are penalized more severely than the less serious ones. At the same time, administrative practice seems to make little sense as a

---

<sup>18</sup> See, e.g., George Tolley et al., *Valuing Health For Policy* (1995); *Valuing Environmental Preferences* (Ian Bateman & K. G. Willis eds., 1999).

<sup>19</sup> See Julie Irwin, Paul Slovic, Sarah Lichtenstein & Gary McClelland, Preference reversals and the measurement of environmental values, 6 *Journal of Risk & Uncertainty* 5 (1993)

<sup>20</sup> See Peter Diamond & Jerry Hausman, Contingent Valuation: Is Some Number Better Than No Number, 8 *J. Econ. Persp.* 45, 49-52 (1994); Note, Ask A Silly Question, 105 *Harv L Rev* 1981 (1992).

whole: Once the practices of diverse agencies are put together, the area appears pervaded by cross-category anomalies. What requires explanation here is both how such anomalies arise and why they persist. We argue that the anomalies arise from the combined effects of category-bound thinking and the arbitrary nature of the translation of moral judgments into punishments. The anomalies persist because of the absence of a strong need, at the individual and social levels, to ensure coherence. Apparently no constituency is seeking to ensure that the system of penalties fit together as a whole. In fact we believe that ours is the first effort even to explore the question of coherence in civil penalties – a point that attests to the fact that people can live (perhaps in ignorant bliss) with patterns that make little general sense. Here as well, underlying sources of the difficulty are category-bound judgments and the translation problem. The result is injustice and arbitrariness.

While emphasizing the problem of incoherence, we do not suggest that coherence is sufficient to produce good outcomes. Systems that are internally coherent can and should be criticized on independent grounds.<sup>21</sup> It is also possible to insist that in some domains, the price of coherence is too steep, perhaps because of the administrative costs of achieving it, perhaps because coherence can be achieved only by altering some institutions, outcomes, and judgments that, on independent normative grounds, turn out to be good. But incoherence of the sort documented here is at the very least a serious problem, because it ensures a set of results that would widely be seen as indefensible and arbitrary. One of our largest goals is to uncover the mechanism that helps to produce this state of affairs, and to see what might be done about it.

### **C. Plans**

This Article is organized as follows. Part II provides some general background on the concept of coherence. Part III explores the psychology underlying the forms of incoherence that is our focus here. Part IV turns to the basic cases of punitive damages and contingent valuation, offering experimental evidence of judgment reversals. In order to provide a focus for normative work on incoherence in law, Part V discusses the implications of our findings for punitive damages awards. Part VI explores more general implications by showing patterns of administrative penalties that most reasonable people would

---

<sup>21</sup> See Joseph Raz, *The Relevance of Coherence*, in *Ethics in the Public Domain* 261 (1994). Raz emphasizes that “Coherence conveys a specific good, the value of which is undeniable. What is incoherence is unintelligible, because it is self-contradictory, fragmented, disjointed.” *Id.* at 264. Raz raises serious doubts about coherence theories in *id.* at 265-70; we do not investigate the resulting complexities here.

reject as absurdly inconsistent, and that nevertheless persist indefinitely, reflecting the pervasive indifference to issues of global coherence that we have attributed to human cognition.

Our basic purpose is descriptive, but we also contend that at a minimum, we have uncovered a serious problem, one that infects judgments and penalties in many areas of the law. While emphasizing the difficulty of achieving agreement on the requirements of full coherence, we suggest that many steps can be taken to correct the worst anomalies. In Part VII, we urge that as an ideal, the legal system should attempt to create institutions that would create more in the way of systemic rationality, and that where results do not fit, there is likely to be a problem of injustice. With respect to the budget, the Office of Management and Budget was originally created on just this ground, and the Sentencing Guidelines had similar aspirations (as yet unrealized, partly because of the absence of explicit cross-category comparisons). With respect to regulation, there are some related problems: the existence of large and apparently inexplicable disparities in expenditures per life saved is, in part, testimony to the absence of sustained cross-category comparisons. Of course the jury system raises special problems and concerns, and there are large questions about the extent to which reforms, even dramatic ones, might overcome the problems stressed here.

In Part VIII, we show that our analysis bears on some larger issues in legal theory, including the debate over the value and possibility of coherence in law and the nature and existence of “incommensurability.” A main theme is that any effort to proceed “one case at a time” will produce serious problems, because of identifiable features of human cognition. We also attempt to show the cognitive basis for the experience of incommensurability. Our closing plea is for institutional changes designed to overcome the problems we identify, replacing predictably incoherent judgments with reforms whose goal is to produce systemic rationality.

## **II. Coherence in General**

Our emphasis here will be on the particular types of incoherence that stem from the human tendency to make category-bound judgments, and from the arbitrariness of the translation of punitive intent into actual punishments. But this type of incoherence should be understood against a more general background, formed by the broader interest in coherence as a goal and by a continuing debate within economics and other social sciences.

Ethical and political philosophers have often viewed inconsistencies with concern,<sup>22</sup> treating them as local warps in a web of beliefs that must be repaired by adequate reflection, in which specific beliefs are brought in line with each other and with broader principles.<sup>23</sup> The search for “reflective equilibrium” is designed to ensure that one’s beliefs, at multiple levels of abstraction, fit together as a sensible whole.<sup>24</sup> There is a similar aspiration to coherence within law,<sup>25</sup> though a decentralized system, with numerous judges, may well have special difficulty in achieving that goal.<sup>26</sup> Our central findings here will show some new difficulties with efforts to achieve anything corresponding to reflective equilibrium within the legal system, or even in moral judgments.

“Rational agent” theories in social sciences, and much important theorizing in the domain of law, rest on the assumption that human agents are endowed with coherent systems of beliefs and preferences, and define coherence as the principal criterion of rationality.<sup>27</sup> An influential definition of rationality avoids any normative evaluation of the specific contents of beliefs and choices, and refers neither to the truth of beliefs nor to the consequences of choices.<sup>28</sup> Only internal consistency matters.<sup>29</sup> In modern economic thinking, and in the economic analysis of law in particular, coherence is considered a touchstone of rationality.<sup>30</sup> The preferences of the idealized rational agent provide a coherent ordering of possible states of affairs, and the beliefs of that agent permit an ordering of events by their probabilities.<sup>31</sup> Furthermore, the dispositions to form new beliefs in the light of evidence, or to make choices when new options are offered, are also assumed to belong to the same coherent structure.<sup>32</sup>

Doubts about the possibility of achieving incoherence lie at the core of a continuing debate about the “rationality” of human agents, involving

---

<sup>22</sup> See Joseph Raz, *The Relevance of Coherence*, in *Ethics in the Public Domain* 261-278 (1994).

<sup>23</sup> See Gilbert Harmon, *Change in View* (1986); David Brink, *Moral Realism and the Foundations of Ethics* (1989); Germain Grisez et al., *Practical Principles, Moral Truth, and Ultimate Ends*, 32 *Am J. Jurisprudence* 231 (1987).

<sup>24</sup> See John Rawls, *A Theory of Justice* 19-22, 46-51 (1971).

<sup>25</sup> See Dworkin, *supra* note.

<sup>26</sup> See Cass R. Sunstein, *Legal Reasoning and Political Conflict* 48-52 (1996).

<sup>27</sup> For overviews, see *Rational Choice* (Jon Elster ed. 1994); Gary Becker, *Accounting for Tastes* (1996); Richard A. Posner, *Economic Analysis of Law* ch. 1 (5<sup>th</sup> ed 1999).

<sup>28</sup> See John Von Neumann & Oscar Morgenstern, *Theory of Games and Economic Behavior* (2d ed. 1947).

<sup>29</sup> For an outline and critique, see Amartya Sen, *Internal Consistency of Choice*, 61 *Econometrica* 495 (1993).

<sup>30</sup> See Von Neumann & Morgenstern, *supra*. In law, see Richard A. Posner, *Economic Analysis of Law* ch. 1 (5<sup>th</sup> ed. 1999).

<sup>31</sup> See Gary Becker, *Accounting for Tastes* ch. 1 (1999); Leonard J. Savage, *The Foundations of Statistics* (1954).

<sup>32</sup> See *id.*

economists, decision theorists, and psychologists, as well as philosophers and academic lawyers.<sup>33</sup> It is now well-known that people are not in fact perfectly rational.<sup>34</sup> They are better described, in Herbert Simon's phrase, as boundedly rational.<sup>35</sup> Bounded rationality has both cognitive and motivational aspects.<sup>36</sup> Because of the limitations in their ability to process information, boundedly rational agents are not able to maintain a system of beliefs, preferences, and dispositions that is both comprehensive and internally coherent. At best, such agents are locally coherent -- achieving consistency over small regions of the space of possible events and outcomes, but not between more remote regions. Furthermore, boundedly rational agents are also cognitive misers -- they economize on difficult thinking and are not inclined to search for inconsistencies among their ideas, or even to acknowledge inconsistencies, unless they are pressured to do so. In particular, we will develop the claim that people effortlessly achieve local coherence in their rankings of actions and outcomes, but show limited ability and little interest in global coherence. Hence global incoherence can persist for long periods.

As a practical matter, complete consistency of beliefs and preferences is an unattainable ideal for any individual, and probably for any legal system. Failures of consistency are inevitable, but some are easier to avoid than others. People are normally successful at avoiding immediate inconsistencies between statements they make in the same setting, but it is much harder to prevent remote inconsistencies between a judgment one makes now and judgments made, or accepted in the past. It is harder still to prevent situations in which a judgment that one makes now is inconsistent with a judgment that one *would* make if one were asked a different question (or the same question in different words). People who are unable to ensure that their current judgments are consistent with other judgments they accept, or with judgments they would make or would have made under different circumstances, inevitably produce a pattern of outcomes that they would themselves consider incoherent and indefensible. We will see this problem in many legal and policy domains.

---

<sup>33</sup> See, e.g., *Choices, Values, and Frames* (Daniel Kahneman & Amos Tversky eds. 2000); *The Rational Foundations of Economic Behavior* (Kenneth Arrow et al. eds. 1996); Gerd Gigerenzer et al., *Simple Heuristics That Make Us Smart* (1999).

<sup>34</sup> For an overview, see Christine Jolls, Cass R. Sunstein & Richard Thaler, *A Behavioral Approach to Law and Economics*, 50 *Stan L Rev* 1471, 1476-1480 (1998).

<sup>35</sup> See Herbert A. Simon, *A Behavioral Model of Rational Choice*, 69 *Q. J. Econ.* 99 (1955).

<sup>36</sup> For an overview, see Jon Elster, *Sour Grapes* (1983).

### III. Causes of Incoherence

In this section we review in some detail the psychological underpinnings of the analysis to be presented in this Article. Much of our treatment here is an effort both to extend and generalize our earlier empirical research on punitive damages<sup>37</sup> and on contingent valuation,<sup>38</sup> and more generally to build on previous theoretical and empirical analyses, not yet applied to law and policy, of intuitive judgment,<sup>39</sup> attitudes and emotion,<sup>40</sup> and spontaneous categorization.<sup>41</sup> We also offer two empirical results that we report for the first time here. The first is an experiment that demonstrates category-bound incoherence in judgments of punitive damages and in contingent valuation. The second is an examination of category-bound incoherence in administrative penalties. In this section, we explore the psychological mechanisms in some depth.

Our previous studies indicated that the dollar numbers produced by jurors in civil cases involving punitive awards, and by respondents in contingent valuation surveys, can be interpreted as expressions of the intensity of a positive or negative attitude – an emotional evaluation of a defendant, or of a public issue.<sup>42</sup> Specifically, we have argued that punitive damages are an expression of indignation or outrage on a scale of dollars, and we identified the problem of translating outrage onto that unfamiliar scale as a critical cause of unpredictability in punitive awards. We extend this conception of punishment here, by adding that (i) reprehensible actions are naturally categorized, and (ii) outrage and its manifestations – including punitive damages – involve a process that situates any particular case in relation to its category. We will show that this manifestation of category-bound thinking yields systematic incoherence in punitive intentions and in punitive awards.

---

<sup>37</sup> Sunstein, Kahneman & Schkade, *supra*; Daniel Kahneman & Ilana Ritov, Determinants of stated willingness to pay for public goods: A study in the headline method, 9 *Journal of Risk & Uncertainty* 5 (1994); Cass R. Sunstein, David Schkade & Daniel Kahneman, Do People Want Optimal Deterrence? 29 *Journal of Legal Studies* 237 (2000); Kahneman, Schkade & Sunstein, *supra*; David Schkade, Cass R. Sunstein & Daniel Kahneman, Deliberating about dollars: The severity shift, 100 *Columbia Law Review* 101 (2000); Kahneman, Ritov & Schkade, *supra*.

<sup>38</sup> Kahneman & Ritov, *supra*.

<sup>39</sup> Daniel Kahneman & Shane Frederick, Representativeness revisited: Attribute substitution in intuitive judgment, to appear in *Heuristics of Intuitive Judgment: Extensions and Applications* (Thomas Gilovich, Dale Griffin & Daniel Kahneman, eds, 2002).

<sup>40</sup> Kahneman, Ritov & Schkade, *supra*; Kahneman & Frederick, *supra*; Paul Slovic et al., The Affect Heuristic to appear in *Heuristics of Intuitive Judgment: Extensions and Applications* (Thomas Gilovich, Dale Griffin & Daniel Kahneman, eds, 2002).

<sup>41</sup> Kahneman & Miller, *supra*.

<sup>42</sup> We intend to take no stand on the relationship between emotion and cognition. See Jon Elster, *Alchemies of Mind* (1999); Martha Nussbaum, *Upheavals of Thought* (2001).

We also extend the same analysis of punishment to a task that at first glance seems to involve little emotion: the setting of penalties by legislative or administrative bodies. The leap is not as radical as it may appear at first. We argue that members of a society are in wide agreement on the categorization of reprehensible actions, and on the relative outrageousness of actions within any one category. Individuals and commissions that set penalties are likely to respect this ordering both for psychological reasons – it corresponds to their moral intuitions as well as to the intuitions of the public – and for political reasons: setting punishments that transparently violate the common ranking of “punishable merit” within a category will appear unjust and evoke resistance. In this situation as well, we argue, category-bound thinking and the problem of translating punitive intent into dollars or other penalties combine to produce global incoherence.

### **A. Punitive Intent: Determinants and Expressions**

We divide the task of setting punishment into two parts:

1. the determination of *punitive intent*, understood as the desire to punish, which locates the appropriate punishment on a subjective scale that ranges from “no punishment at all” to “extremely severe punishment.”
2. the *translation* of punitive intent into a metric that can actually be used by the legal system, such as dollars of fine or years in jail.

As we shall show, some forms of incoherence arise at the level of punitive intent, while others are caused by features of the translation process.

Social psychologists commonly identify an emotion and a tendency to action as elements of attitudes.<sup>43</sup> In our usage, *outrage* is the emotion and *punitive intent* is the action tendency. We believe that they are directly related. Outrage and punitive intent both are psychological variables – along with other subjective variables such as brightness, loudness, pain, trust and dislike. Any subjective variable can be expressed in multiple ways. For example, outrage may be expressed by appropriate adjectives, by the choice of a number on a rating scale, by loud screams, or by jail sentences.

---

<sup>43</sup> Alice Eagly & Shelley Chaiken. *The Psychology of Attitudes* (1993).

The existence of something like a moral community is assumed in the analysis that follows. When we speak of the outrageousness of an action or the punitive intention that is evoked by it, we have in mind an emotional state and a tendency to act that are widely shared in a relevant community. Consensus in moral attitudes was demonstrated in an earlier study<sup>44</sup> in which respondents (drawn from jury-eligible citizens in Travis County, Texas) were asked to evaluate product liability cases involving physical injuries. Different groups of respondents used different scales in these evaluations. Some evaluated the *outrageousness* of the defendant's behavior (always a firm, also identified by annual profit as an indication of its size) using a rating scale (from "Completely Acceptable" to "Absolutely Outrageous"); others rated their punitive intent, also on a rating scale (from "No Punishment" to "Extremely Severe Punishment"). A third group assessed the appropriate amount of punitive damages, in dollars.

We found substantial consensus in these judgments. Thus, on average, rich and poor; educated and less educated; white, Hispanic, and African-American; old and young; male and female; and all others are likely to agree on how to rank and rate a set of personal injury cases in terms of their "punishable merit,"<sup>45</sup> the outrageousness of the defendant's conduct, and the punitive damages that are appropriate.<sup>46</sup> Note that the agreement is between the *average* judgments for the various social groups. There is some variability in judgments of each case within each group, and especially for judgments made in dollars. In general, however, people appear to agree on what makes one personal injury worse than another (e.g., amount and duration of pain, disability, vulnerability of victims such as the elderly and the very young); they also agree on elements of the defendant's behavior (e.g., intentionality, deception) that make one reprehensible action worse than another; and they agree as well on the severity of the harm inflicted on the plaintiff – provided that comparisons are restricted to harms of a particular type, such as physical injury. .

An analysis of the results of our study showed that the punitive awards assessed for 28 separate scenarios was predicted quite accurately by a very simple formula, which captures the psychological conception of punishment that we apply and extend in this article<sup>47</sup>:

---

<sup>44</sup> The results are reported in Sunstein, Kahneman & Schkade, *supra* note, and in Kahneman, Schkade & Sunstein, *supra* note.

<sup>45</sup> See Breyer, *supra* note 1.

<sup>46</sup> See Kahneman Schkade & Sunstein, *supra*; at 2099. Sunstein et al., *supra*.

<sup>47</sup> The results of this analysis were described by Kahneman & Frederick, *supra*. The severity of the harm associated with each of the scenarios used by Kahneman, Schkade & Sunstein, *supra* was rated by a small group of Princeton students, who were not given any details about the cause of the harm.

$$\text{Punishment (\$)} = \text{Outrageousness of Behavior} \times \text{Severity of Harm} \times \text{Translation factor}$$

What this formula says is that punitive intent is proportional to the outrageousness of the harmful action, and that punitive intent is also proportional to the severity of the harm. The translation factor is required to transform punitive intent – here construed as a state of mind – into actual punishments.

Punitive intent, as this formula indicates, is firmly retributive, in the sense that both the act and its consequences matter.<sup>48</sup> Because of the retributive nature of the underlying intuitions, the public sense of what is just punishment is radically at odds with the idea, popular in some academic circles, that punishment should be grounded largely in its deterrent function, and therefore directly linked to the likelihood that an act will be discovered and punished.<sup>49</sup> Punishments that appear just in this society are based on outrage directed at the action and concern with the degree of harm that the action caused.<sup>50</sup>

## **B. The Translation Problem**

We have described punitive intent as a state of mind – a sense that it is right for a miscreant to suffer some degree of pain. Punitive intent can be expressed in words, such as “severe” or “mild.” To have an effect in the real world, however, the intention to punish must be translated onto a scale that can be used by the legal system, such as dollars of fine or months in jail. The translation factor in the formula above represents this operation of scaling.

For many classes of harms, people lack shared moral intuitions that might specify the translation factor. Except to the extent that they are familiar with existing practice, people do not have a clear, agreed-upon sense that a grossly reckless action should be punished with a punitive award of \$50,000, or \$100,000,

---

<sup>48</sup> Jonathan Baron & Ilana Ritov, Intuitions about penalties and compensation in the context of tort law, 7 *Journal of Risk & Uncertainty* 17 (1993).

<sup>49</sup> Sunstein, Schkade & Kahneman, *supra*, reported that probability of detection, which is central to the law and economics analysis of deterrence, is largely irrelevant to intuitions about just punishment. Punishment that is deliberately calibrated by probability of detection is considered unjust, when it results in similar actions being punished differently in different localities. Baron & Ritov, *supra*, report that known and explicitly specified future impact of decisions do not substantially affect judgment, even of experts, although regulating future behavior is a critical factor in the theory of law and economics.

<sup>50</sup> See the acknowledgment of this point in *Cooper v Leatherman*, *supra*.

or \$1,000,000, or that a case of assault should be punished with a jail term of six months, two years, or five years. As a consequence, the translation factor that is used is often influenced by irrelevant considerations, personal experience, and random circumstances. In the case of punitive damage awards, for example, the plaintiff's demand matters a great deal, simply because juries often have few other relevant dollar figures from which to begin.<sup>51</sup> And because of the influence of these irrelevant considerations and random circumstances, actual punishments for transgressions are often incoherent in our sense. That is, different juries, who agree on punitive intent for a given set of case facts, may impose substantially different punishments solely because of arbitrary differences in their translation factors. These propositions have direct implications for a wide range of tasks relevant to the law. They apply with equal force to the determination of punishment for individual cases, and to the writing of statutes or regulations that specify punishments for particular transgressions.

Optimal deterrence theory offers a way of calculating the appropriate punishments, again by connecting dollars awards to the harm through the translation factor of "likelihood of compensation."<sup>52</sup> But we have already seen that this analysis violates common intuitions about just punishment.<sup>53</sup> For those who seek to come up with dollar amounts that are intuitively appealing, a main difficulty is that harm and punishment do not generally occur in the same or even in commensurable units. What is the appropriate fine for hunting an endangered bird? \$50? \$500? \$5000? Should the punitive damages for a case of employment discrimination be set at three times the value of compensatory damages, or at fifteen times that value?

In an earlier analysis of punitive damages we focused on the arbitrariness of translation factors as a major cause of the notorious unpredictability of punitive damage awards.<sup>54</sup> We expected, and found, large variation in the

---

<sup>51</sup> Reid Hastie, David Schkade & John Payne. Juror judgments in civil cases: Effects of plaintiff's request and plaintiffs identity on punitive damage awards, 23 *Law & Human Behavior* 445 (1999); also Gretchen B. Chapman & Brian H. Bornstein, The more you ask for, the more you get: Anchoring in personal injury verdicts, 10 *Applied Cognitive Psychology* 519 (1996) (finding similar effects for plaintiffs' demand in the context of compensatory awards for personal injury).

<sup>52</sup> See A. Mitchell Polinsky & Steven Shavell, *Punitive Damages: An Economic Analysis*, 111 *Harv. L. Rev.* 869, 870-76 (1998).

<sup>53</sup> See Sunstein, Schkade & Kahneman, *supra*. Indeed, the Supreme Court has acknowledged that punitive awards are not rooted in deterrent judgments. See *Cooper v Leatherman*, *supra*.

<sup>54</sup> Kahneman, Schkade & Sunstein, *supra* at 2106-07. We described the juror's task as a special case of scaling without a modulus. The term is borrowed from psychological experiments in which observers use numbers to indicate the intensity of sensations such as brightness, loudness or pain. The common practice in such scaling experiments is to specify a particular stimulus (e.g., a level of luminance) as a standard. Observers

translation factor that people will apply in judging a particular case, *even when their punitive intent is the same*. In subsequent research we found that jury deliberations are not a cure for the variability of translation factors.<sup>55</sup> The general effect of deliberation was instead a *severity shift*: it appears that when jurors start off with different translation factors, higher punishments that are proposed in the deliberation are more likely to be adopted, and are sometimes even exceeded by the jury's award. Even after deliberation, different mock juries considering the same case often reach radically different decisions – punitive damage awards sometimes vary by a factor of 100 or more.<sup>56</sup>

It is easy to see that free-floating translation factors will often cause failures of our test of coherence. Imagine two juries that consider separate claims A and B against the same defendant, awarding \$1 million for case A and \$5 million for case B. Because the translation factor that they apply may be different, it is entirely possible that both juries, if asked to compare the two cases, would agree that case A is higher in “punishable merit” than case B. The inevitable consequence of variability in the translation factor is incoherence in punishments: different individuals and different juries will make judgments that cannot be reconciled, even in the presence of underlying consensus on punitive intent.

We do not intend to imply that the translation factors that people apply are completely arbitrary. Indeed, data that we have reported demonstrate the opposite. In one of our earlier studies<sup>57</sup> of punitive damages each case was presented in two versions which differed only in the annual profits of the defendant firm. For example this value was “\$10-20 million” in one version and “\$100-200 million” in another. The indication of the size of the defendant firm had a large and systematic effect on punitive awards, which were substantially higher for larger firms. The mechanism that produces this effect is most likely

---

are instructed to assign a particular number (the modulus) to that stimulus, and to assign numbers to other stimuli by comparing them to the subjective intensity associated with the standard. The task may appear meaningless, but in fact it is one that people can carry out with fair agreement – much like the consensus they exhibit in judging the severity of harms or the outrageousness of behaviors. In some experiments, however, a modulus is not supplied, and the observers are requested to assign whatever numbers they feel appropriate to report the subjective intensity of their sensations. In such situations of scaling without a modulus, observers spontaneously adopt their own individual standard and apply it consistently. However, different observers choose different moduli for no apparent reason. Thus, two stimuli may be rated as 10 and 2 by one individual, and as 40 and 8 by another, simply because the second individual picked a modulus that is 5 times larger than the first.

<sup>55</sup> See Schkade, Sunstein & Kahneman, *supra*.

<sup>56</sup> *Id.* at Table 5

<sup>57</sup> Kahneman, Schkade & Sunstein, *supra*.

the same that produces anchoring of awards on irrelevant factors. In this instance, however, the effect of firm size on awards corresponds to the sensible intuition that a larger financial punishment is required to inflict the same level of “pain” if the defendant firm is large than if it is small. Other anchoring effects could be defended as well: punitive awards have been found to be strongly correlated with compensatory awards in the same case,<sup>58</sup> and perhaps anchoring of punitive on compensatory awards could be justified as appropriate, though this would not be an easy task.<sup>59</sup>

### C. Categories and Frames of Reference

In this Article we study two manifestations of *category-bound thinking* that produce incoherence in punishments: (1) the frame of reference for judging harmful actions is linked to a category of harms, and therefore liable to change when cases that involve harms of different kinds are explicitly compared; (2) the commissions that set penalties for categories of misconduct appear to be unconcerned with the penalties already on the books for other categories. In both situations, we suggest that judgments are sensible and coherent within each category separately, but not when the view screen is expanded to include more than one category. To understand this suggestion, and the mechanisms that produce judgment reversals, it is necessary to know something about what categories are and how they operate.

The term “category,” as used in this Article, is borrowed from modern cognitive science, where it serves to explain how people use categories and category labels in informal reasoning and in everyday language.<sup>60</sup> This concept of categories and categorization is quite different from an approach in which a category is defined by a set of necessary and sufficient conditions for membership. In that system, the boundary between categories are sharp and membership is all-or-none. In everyday language, in contrast, the boundaries of categories are fuzzy and membership is graded.<sup>61</sup> A chicken is surely a bird, but is not quite as good a bird as a robin, or an eagle. And a whale is not a fish, but it

---

<sup>58</sup> Theodore Eisenberg et al., *The Predictability of Punitive Damages*, 26 *J. Legal Stud.* 623, 644 (1997).

<sup>59</sup> It would not be easy on any theory of punitive damages. On optimal deterrence theory, a key issue is the multiplier; the compensatory award is in some sense an anchor, but a weak one, because a 100% likelihood of compensation should produce a punitive award of 0, an a 10% likelihood of compensation should produce a punitive award ten times that of the compensatory award. On the retributive view, there is no obvious translation formula.

<sup>60</sup> See, e.g., Daniel Reisberg, *Cognition: Exploring the Science of Mind* (1997), ch. 8; Laurence Barsalou, *Cognitive Psychology: An Overview for Cognitive Scientists* (1992).

<sup>61</sup> Eleanor Rosch et al., *Basic Objects in Natural Categories*, 8 *Cognitive Psychology* 382 (1976).

is a bird even less. Indeed, the habit of thinking in terms of fuzzy categories and graded membership is so deeply ingrained that it explains the documented difficulties of juries in conforming to judges' instructions about verdicts that are defined by necessary and sufficient conditions.<sup>62</sup> Of course, no one would wish to claim that people are *incapable* of thinking in terms of necessary and sufficient conditions for category membership. This mode of thinking is dominant in mathematics and in scientific reasoning, and at least implicitly in some domains of law as well; consider the question of whether speech can be regulated as "obscene."<sup>63</sup>

Any object or event is a member of innumerable categories (for example, the category of things that are not Julius Caesar). But questions that require a judgment reveal that categorization follows rules that are widely shared and quite predictable. Consider the questions "is an eagle large?" and "is a cabin small?" The expected answer to both questions, when separately presented, is "yes." The categories of birds and buildings were not mentioned in the questions, but they are evoked in the respondent's mind, and they are automatically used to provide a frame of reference for the otherwise ambiguous question about the objects' size. We will say that the judgments are spontaneously *normalized* to the frame of reference implied by the category.

There are situations in which a judgment can be normalized to a category that is not associated with a definite category label. Consider the following example: "John is a six-year old child whose pajamas caught fire as he was playing with matches. The boy was badly burned over a significant portion of his body and required several weeks in hospital and months of physical therapy. How would you describe the harm that was caused in this incident? Extremely severe? Severe? Mild?" At first glance, this question may appear impossible to answer without specifying a frame of reference. Remarkably, most people do not find this task meaningless or even particularly difficult. They will readily generate judgments that will cluster around "very severe." Evidently, people evaluate a case as mild or severe relative to an ad hoc frame of reference, which consists of similar examples that spontaneously come to mind. In the case at hand, we suspect that the categories that people construct correspond fairly closely to the class of "injuries that occur in household accidents," and we also suspect that they do not explicitly label the category as such.

---

<sup>62</sup> For an example with punitive damage instructions, see Reid Hastie, David Schkade & John Payne, A study of juror and jury judgments in civil cases: Deciding liability for punitive damages, 22 *Law & Human Behavior* 287 (1998).

<sup>63</sup> See the definition of obscenity in *Miller v. California*, 413 U.S. 15 (1973).

The judgments that people make without a specified frame of reference are nevertheless quite orderly. A story about an elderly woman who slipped in a bathtub and sprained her ankle will elicit substantially milder judgments. Furthermore, the fact that people generally agree closely in their judgments suggests that the informal categories that people construct around any given case are quite similar. The boundaries of these informal categories are undoubtedly fuzzy, and they include cases that differ along many dimensions, such as the permanent loss of four fingers or several months of complete blindness. We can be certain, however, that other harms, such the destruction of a reputation, or a drop in the population of dolphins, will not come to mind when the case of the burned child is considered. Similarly, cases of damage to reputations or to species will evoke their own frame of reference, which will not include personal injuries. The meanings of words such as “mild” or “severe” will be normalized to each of these fuzzy categories.

Normalization in the use of adjectives is a simple convention of language, which usually causes no serious problems of communication. But there is a separate phenomenon, one of considerable importance to psychology and law: *normalization is also found in emotional reactions, in punitive intentions, and in actual punishments.*<sup>64</sup> Consider the behavior of a guest at a formal dinner who pushed his plate away in obvious distaste, while muttering a derogatory comment about the hostess. Casual observation suggests that intense emotions of indignation can be evoked by such rude behavior, not only in the victimized hosts but also in observers. The intensity of these feelings might well be greater than that created by actions such as misdeeds by public officials, which cause much graver harms but are normalized to a different category. In an explicit comparison, of course, everyone will agree that something is inappropriate about reactions of outrage that are more intense for trivial misconduct than for genuine malfeasance. The relativity of the emotion of outrage is important to our story about law, because we argue that outrage and punitive intent are closely linked in the moral attitudes of most people, and are directly expressed in actual punishments.<sup>65</sup> As we now show, the normalization of outrage is a significant source of incoherent punishments.

Recall the minimal criterion for coherence that we apply here: Is the ranking of two or more cases the same when they are directly compared and when they are judged in isolation? Note that this criterion is likely to be met for

---

<sup>64</sup> See Kahneman, Schkade & Sunstein, *supra* note.

<sup>65</sup> See Sunstein, Kahneman & Schkade, *supra* note; Kahneman et al., *supra* note.

objects that belong to the same category. Judged separately, burns over much of the body appear severe and a sprained ankle appears mild. The ranking of the two harms will not change when the two cases are directly compared, because the frames of reference were the same to begin with. When harms of the same general kind are caused by misconduct, we expect outrage and punishments to pass the test of coherence. Normalization occurs, but it does not produce incoherence.

The risk of incoherence arises when the judgments made separately and in direct comparison involve different categories, and therefore different frames of reference. First, comparisons across categories of harms are particularly difficult because they are not easily described in the same language. To appreciate the difficulty, consider an action that led to one person losing four fingers and also to the death of 10,000 migratory birds. The comparison of these harms involves complexities of a different order from within-category comparisons. The difficulties of cross-category comparisons inevitably lead to instability in the judgments of individuals, and to an impairment of consensus, relative to within-category comparisons.<sup>66</sup>

What causes this more systematic form of incoherence? The reason is that comparison across categories introduces features that are relevant to the judgment, but not given adequate weight when cases are considered in isolation.<sup>67</sup> We believe, without formal evidence, that there is substantial agreement on comparisons *between* categories in terms of a vaguely defined but well understood dimension sometimes called “prominence” in the psychological literature.<sup>68</sup> For example, physical injury is more prominent than financial injury, and harm to people is more prominent than harm to other biological species. Because of normalization, the relative prominence of categories of harm plays no role when these harms are judged in isolation. But because the features that distinguish categories become very salient in a comparison, incoherence is an inevitable consequence<sup>69</sup>. We test this proposition in the experiment described below.

---

<sup>66</sup> We have collected data on this topic that confirm the difficulty of cross-category comparisons in a context of punitive damage cases involving different types of harm. See Cass R. Sunstein, Daniel Kahneman, David Schkade & Ilana Ritov, *Legal Coherence and Incoherence* (unpublished draft 2000).

<sup>67</sup> Ilana Ritov, *The Role of Expectations in Comparisons*, 107 *Psychological Review* 345 (2000).

<sup>68</sup> Amos Tversky, Shmuel Sattath & Paul Slovic. *Contingent weighting in judgment and choice*. 95 *Psychological Review* 371 (1988).

<sup>69</sup> Dedre Gentner & Arthur B. Markman, *Structural alignment in comparison: no difference without similarity*, 5 *Psychological Science* 152 (1994); Arthur B. Markman & Douglas L. Medin, *Similarity and alignment in choice*, 63 *Organizational Behavior & Human Decision Processes* 117 (1995). Douglas L. Medin,

## IV. A Study of Normalization and Incoherence

To illustrate the judgment reversals produced by category-bound thinking and by the process of normalization, we consider two different evaluation tasks, involving punitive damage awards and contingent valuation of public goods. Both punitive damage awards and contingent valuation have been highly controversial in recent years.<sup>70</sup> We will present evidence of judgment reversals in each context.

### A. Hypothesis and Design

Our central question was whether people's assessments would differ depending on whether they viewed a problem in isolation (separate evaluation) or in comparison to a problem from a different category (cross-category comparison). The problems we used are drawn from categories of harm that differ in prominence (physical injury vs. financial loss; harm to human beings vs. harm to the ecology<sup>71</sup>). Our central hypothesis was that when a problem is viewed in isolation, judgments are normalized within the relevant category, but that this normalization is prevented when a comparison to a problem from a different category is required. Instead, differences between categories will dominate the consideration. We therefore expected problems drawn from categories involving the worse type of harm to evoke higher judgments (larger punitive awards, higher willingness to pay) in comparative than in isolated evaluations.

We asked a large group of people to consider pairs of legal cases and of public causes (see Table 1).<sup>72</sup> Each pair of legal cases consisted of one case of physical injury and one case of financial loss.<sup>73</sup> Each pair of public causes

---

Robert L. Goldstone & Arthur B. Markman, Comparison and choice: Relations between similarity processes and decision processes, *2 Psychonomic Bulletin & Review* 1 (1995). Ritov, *supra*.

<sup>70</sup> On punitive damages, see, e.g., Symposium, The Future of Punitive Damages, *1998 Wisc. L. Rev.* 1; on contingent valuation, see, e.g., Symposium, Contingent Valuation, *8 J. Econ. Persp.* 3 (1994).

<sup>71</sup> It might be objected here that everything depends on the degree of harm: a trivial physical injury is not as bad as a huge financial loss. We agree. We are suggesting only that the categories, as such, are ranked, and ranked in the way that we describe.

<sup>72</sup> Jury-eligible citizens from Travis County, Texas ( $n = 1,035$ ) were recruited by a survey firm and paid for their participation. All respondents received detailed instructions about the task of jurors in civil cases and the criteria for punitive damages.

<sup>73</sup> Each case was a short summary (200 words) of a realistic incident, in which the defendant was always a large company described as having "annual profits of around \$150 million." Respondents were told to assume that \$500,000 in compensatory damages had already been awarded to the plaintiff, and that the case

consisted of one threat to human health, and one threat to the environment.<sup>74</sup> Some people made their evaluations on a dollar scale and others on rating scales.<sup>75</sup>

After evaluating one case and one cause in isolation, people saw the paired case and cause, which they were asked to compare to the one they had already seen (see Appendices A, B and C). For example, for legal pair A, they were asked “Compared to the manufacturer of childproof safety caps, how much punishment does the company that sold repainted cars as new deserve?” and for Pair D of public causes “Compared to protecting coral reefs, how much money would you be willing to contribute to a fund for farm workers” skin cancer?”). After answering this question by circling “less than,” “more than” or “the same,” participants also evaluated the new item on the same scale they used in making judgments in isolation.

---

had reached the stage of determining whether punitive damages were appropriate, and if so, in what amount.

<sup>74</sup> Each threat and a proposed mitigating intervention was described in a paragraph of text. Each person was randomly assigned to one legal pair and to one public cause pair.

<sup>75</sup> One third of the participants responded on rating scales and the other two thirds on dollar scales. For the punitive damage cases, the rating scale was the appropriate severity of punishment, ranging from 0 (“no punishment”) to 8 (“extremely severe punishment”). The dollar scale was the amount of punitive damages that the defendant should be required to pay. For public causes, the rating scale was the satisfaction that individuals would derive from contributing to the cause, ranging from 0 (“no satisfaction at all”) to 6 (“extreme satisfaction”). The dollar scale was the maximum amount that they would be willing to contribute to the cause (denoted below as WTP – willingness to pay).

**Table 1. Summary of Problems**

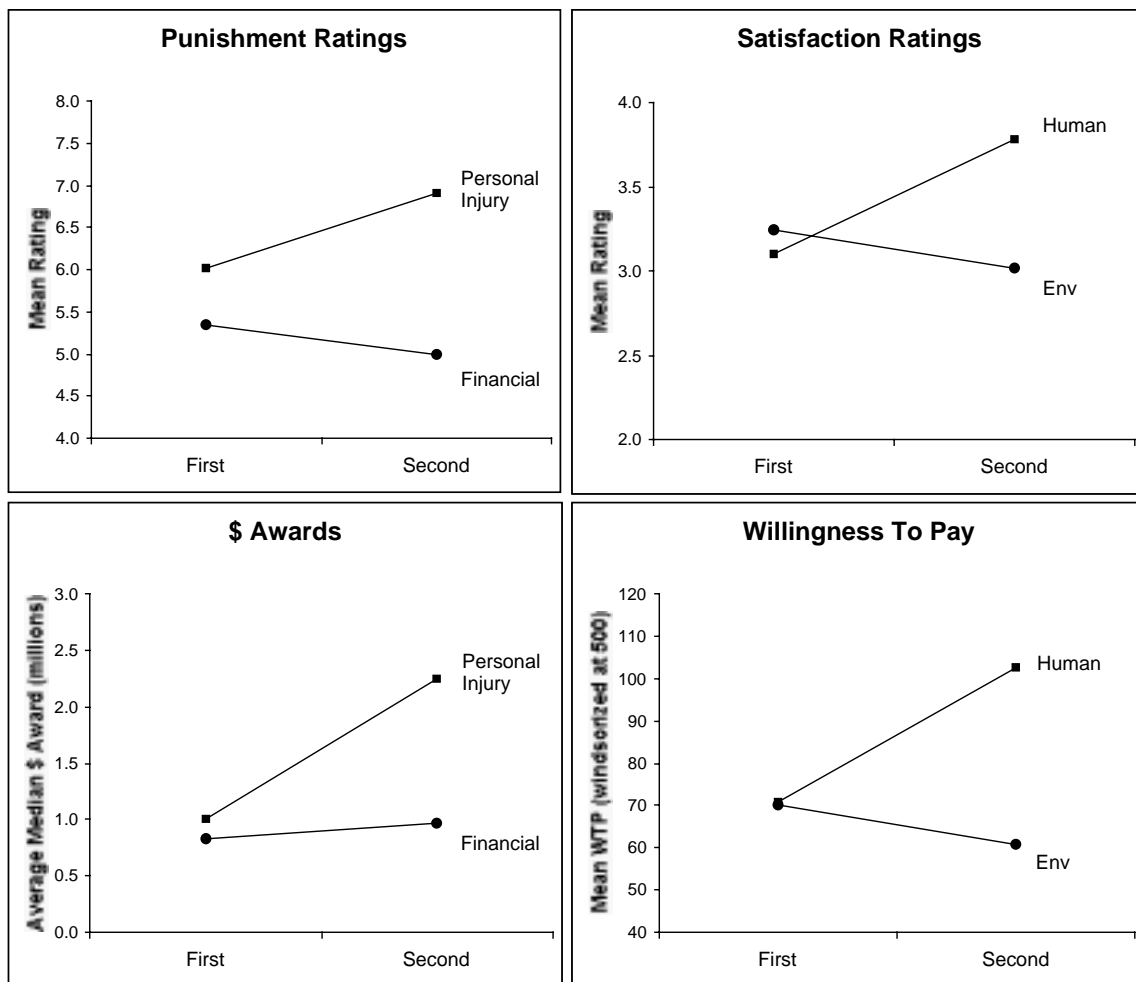
<b>Legal Cases</b>		
<b><u>Pair</u></b>	<b><u>Physical Harm</u></b>	<b><u>Financial Harm</u></b>
A	Childproof safety cap fails; child needs hospital stay	Repainted cars sold as new to leasing agency
B	Patient injured when surgeon ignores standard safety rules	Distributor violates contract, damaging supplier's
business		
C	Driver injured when defective steering system fails	Trustee manipulates trust to enrich a favored client
<b>Public Causes</b>		
<b><u>Pair</u></b>	<b><u>Harm to Humans</u></b>	<b><u>Harm to the Environment</u></b>
D	Program to improve detection of skin cancer in farm workers	Fund to clean up and protect dolphin breeding locations
E	Research on a type of bone marrow cancer in the elderly	Fund to protect coral reefs by banning cyanide fishing
F	Program to replace lead-based paint in low-income areas	Fund to protect elephants from ivory poachers

**B. Findings**

The four panels of Figure 1 present our results, pooled over the three pairs of legal cases and public causes. The Figure shows unambiguous shifts of judgment, depending on whether a problem is viewed in isolation or in the context of a problem from a different category. The pattern of shifts is just the one we predicted from a process of normalization: when cases from categories that differ in prominence are viewed in isolation, the effect of the category is suppressed. As a result, the more prominent harm is assigned a lower rating and a lower dollar value when judged by itself than when directly compared to a harm of a less prominent kind. The pattern of results is strikingly similar for ratings on bounded scales and for dollar responses. It is also precisely the same

for responses to harms that are to be punished (civil suits) and for harms that are to be prevented or remedied (public causes).

Figure 1. Aggregate Evaluations over Cases or Causes for Rating and Dollar Responses



Note also that when cases are compared across categories of harm, the evaluation of the more prominent item rises sharply, while that of the less prominent item declines slightly or stays about the same.<sup>76</sup> The asymmetric

<sup>76</sup> This pattern is confirmed by analysis – the interaction of Evaluation Mode (SE vs CCC) and Item Type (e.g., financial vs personal injury) is evident and statistically significant in all four panels (each  $p < .001$ ).

effects of the comparison on more and on less prominent harms reflects a cognitive phenomenon that is well understood, but not central to our story.<sup>77</sup>

Measured in dollars, the shifts illustrated in Figure 1 are quite substantial. Averaging over the three personal injury cases, the median dollar award rose from \$1 million in separate evaluation to \$2.25 million when they were directly compared to cases of financial harm. The median dollar award increased from separate to cross-category evaluation for all three personal injury cases. At the same time, the median award did not change significantly for any of the three financial cases. The shifts in willingness to pay – which in standard contingent valuation practice would be translated into estimates of the value of the goods – are also substantial.

The similarity of results for ratings and dollar awards is particularly noteworthy. As our discussion of eagles and cabins showed, category-bound shifts in the meaning of adjectives are expected as a convention of language. But dollars, like centimeters, are objective units that are supposed to retain the same meaning across contexts. An individual who said an eagle is large and a cabin is small has no cause for embarrassment, but an individual who estimates the size of eagles and cabins in metric units is supposed to make estimates that are absolute, not category-relative. (We would want to pose some questions to anyone who insists that eagles are *really* bigger than cabins.) Likewise, jurors who assess punishments in dollars are supposed to use an absolute scale, and their judgments should not be different from those they would make in the context of a comparison with a different type of misconduct. Similarly, the standard interpretation of stated willingness-to-pay as a measure of the value of public goods assumes that these statements provide an absolute measure of these goods, but the sensitivity of WTP to the judgment context casts doubt on this essential assumption.

The similarity of the results for ratings and for dollar values also provides support for another claim we have made earlier about both punitive damages and contingent valuation.<sup>78</sup> We have argued that punitive awards and

---

<sup>77</sup> Psychological research indicates that comparison of two objects is not a symmetric process: one object (the subject) is compared to another (the referent). In the present design, the problem presented in the first envelope was the referent to which the second item was compared. Evidence from other studies indicates that the perceived contrast between two objects is greater when the more prominent of them is the topic of the comparison. For example, China is more different from North Korea than North Korea is from China. See Amos Tversky, Features of similarity, 84 *Psychological Review* 327 (1977). In the present instance, this process would cause a greater effect in making the more prominent harm look worse than in reducing the significance of the less prominent one.

<sup>78</sup> Kahneman, Ritov & Schkade, *supra*

statements of willingness to pay are direct expressions of an emotional response: outrage in the case of harm caused by one person to another, and the fear and hope associated with the possibility of preventing or undoing harm to people, or to aspects of the ecology. But we have also added a crucial element here, which we infer from the shifts in dollar responses. Emotional responses, no less than the responses in dollars, are category-bound and context-dependent. The emotion itself will vary, depending on whether a case or problem is considered in isolation or compared to a case or problem from a different category.

### C. Some Speculations

The normalization effect that we documented in this study has other implications as well. In particular, the results suggest that the contingent valuation approach is likely to distort the relative willingness of the public to deal with threats to the environment and with threats to people, unless steps are taken to overcome the problems identified here. And if people's informed judgments are taken to be the criterion, punitive damage awards are likely to be too high in financial injury cases and too low in personal injury cases. Some data supports this suggestion.<sup>79</sup> We expect that similar shifts could be documented with many other pairs of categories. For example, we suspect that punitive damages awards involving libel might well be higher, in isolation, than punitive damage awards involving racial discrimination; but there is likely to be a reversal if the two cases are put together.

In the theory that we have proposed, normalization reduces the differences in the overall level of outrage that is evoked by misconduct that causes different kinds of harms. The comparison of cases drawn from different categories restores differences between categories of harms that differ in prominence – this should be considered an improvement in the quality of judgment. We hypothesize that the comparative situation may alter judgments in another way, by reducing the anchoring effects of compensatory damages on punitive awards.<sup>80</sup> In the real world of punitive damages, unlike our experiment, compensatory awards are generally much larger in financial cases than in cases of physical injury.<sup>81</sup> In this pair of categories, the anchoring effects favor large

<sup>79</sup> See Jonathan Karpoff & John Lott, *On the Determinants and Importance of Punitive Damage Awards*, 42 *J. Law and Ec.* 527, 539 (1999).

<sup>80</sup> See Eisenberg, et al., *supra*. Note that the function of an anchor is similar to the function of a modulus in psychophysical judgments. See not *supra*.

<sup>81</sup> See Karpoff & Lott, at 538-39, showing mean awards of \$14.8 million in fraud cases, and \$20.6 million in business negligence cases, but \$6.2 million in product liability cases, \$1.6 million in malpractice cases, and \$991,000 in motor vehicle accident cases.

judgments for the less prominent harm. As a consequence, a case of financial damage with a large compensatory anchor (say \$10,000,000) is expected to receive a higher punitive damage award than a case of physical injury with a smaller anchor (say \$500,000), when the two are judged in isolation. When cases of the two kinds are directly compared, we expect that many people will be more strongly influenced by the relative prominence of the harms than by the relative size of the anchors. Preliminary evidence<sup>82</sup> that we have collected supports this hypothesis, which suggests that two distinct mechanisms may cause punitive awards for financial cases to be higher in the current system than they would be if jurors were given a richer context: anchoring on high dollar numbers, and masking of the low prominence of the category through the effect of normalization.

## **V. Punitive Damages: Is There A Problem? What Kind of Problem?**

If cross-category judgment reversals are likely to be pervasive (as we believe), the problem of predictable incoherence will appear in many contexts. In what follows, we speculate on the evidence of incoherence in punitive damages and on possible reforms that might address the problem. Our principal message is critical rather than constructive; we claim to have uncovered a serious problem, not to know how to solve it. Our aim is not to endorse a particular proposal, but to explore the implications of incoherence in moral intuitions for legal institutions.

Most of what we say here will apply, with suitable variations, to the assessment of incoherence in other settings, such as contingent valuation, administrative penalties and fines. In Parts VI, VII, and VIII, we turn to some broader speculations.

### **A. Can Coherence Be Achieved?**

We have observed that people's judgments about appropriate punitive awards are very different, depending on whether they see a case in isolation or in comparison to another case. We also believe that most people, in most conditions, prefer a wide view screen 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 than in a narrow context. These points raise serious

---

<sup>82</sup> See Cass R, Sunstein et al., *Legal Coherence and Legal Judgment* (unpublished manuscript 2001).

questions about the current system, and support a presumption against the outcomes of one-shot judgments, which occur by design in punitive damages cases. Furthermore, they suggest a direction for reform. If the purpose of the institution of the jury is both to elicit the judgments of ordinary people and to satisfy their wish for justice, there is good reason to search for ways to prevent or correct outcomes that all or almost all would consider anomalous when placed in the context of the pattern as a whole. More ambitiously, we might seek ways of ensuring judgments that fit into a pattern that ordinary people would find coherent. The achievement of global coherence is exceedingly difficult, but it has considerable appeal, at least as an ideal.

But any effort to achieve this ideal is open to three major objections. The most basic is that there is good reason to doubt the very existence of an underlying system of moral intuitions that is internally consistent. Recent behavioral studies of human decision making have undermined the hope that people's choices can be explained in terms of a comprehensive and coherent 'preference order'.<sup>83</sup> The evidence suggests that people construct preferences on the fly, and that their choices are highly dependent on the immediate context and on inconsequential features of the options they consider. Of course, choices are not chaotic or random – the decisions that people make have some consistency and structure, but this structure is better described by the loose logic of attitude than by the coherent logic of fully rational preferences.<sup>84</sup> This argument extends directly to moral choices, moral preferences and moral intuitions – raising doubts about the feasibility of reflective equilibrium in this domain.<sup>85</sup>

Even if the moral intuitions and sentiments of individuals could be rationalized, there would be a question about the feasibility of a coherent system that is broadly acceptable across society. We have already mentioned the hypothesis that social and cultural groups may differ more in their ranking of categories of harm than in their rankings of harms within each category.<sup>86</sup> Obviously, serious problems will be encountered in attempting to achieve agreement on a system of punishments that incorporates the judgments of different individuals and groups about all categories – recall the epigraph from Justice Breyer.

---

<sup>83</sup> See Kahneman and Tversky, 2000 for multiple sources on this point.

<sup>84</sup> Kahneman, Ritov & Schkade, 1999.

<sup>85</sup> Daniel Kahneman, Unpublished Tanner Lecture, University of Michigan (1994).

<sup>86</sup> For example, members of minority groups may rank racial discrimination as more important than members of the majority, stockholders may rank securities fraud as more important than non-stockholders, etc.

The third objection is that even if the moral intuitions of ordinary people could be represented as a broadly acceptable and coherent system, it might be wrong to adopt this system as the final arbiter of issues in law in general, and of judgments of punitive damages in particular. As we have seen, there is compelling evidence that the popular conception of justice is more concerned with issues of retribution than with issues of deterrence.<sup>87</sup> Those who accept the dominant approach in law and economics, where deterrence is the principal goal of punitive damages,<sup>88</sup> are likely to reject any system in which the probability of detection plays little or no role in the setting of such damages. On the other hand, of course, there is reason to believe that a system of punishments that is exclusively concerned with deterrence will be considered unjust by the public.

Finally, if these all these fundamental objections could be met, issues of feasibility would arise. Even if an underlying and comprehensive system of intuitions exists, the methods of uncovering it do not. There is a limit to the complexity of cognitive tasks that ordinary people can be asked to do, and ranking the punishable merit of different kinds of misconduct that cause different kinds of harms clearly exceeds these limits. Furthermore, the construction of instruments for eliciting these judgments would inevitably be questioned for possible bias and other flaws of method.

Thus, the ideal of seeking a system that is both coherent and reflective of popular moral sentiments can be questioned on both normative and practical grounds. Nevertheless, we submit that incoherence, once acknowledged, is unacceptable, even scandalous. Something should be done about it. We next consider some of the things that may be done. Our emphasis is on finding ways to correct obvious errors, which is a realistic goal, rather than to produce what all would agree to be coherent, a goal that seems, in light of what we have just said, to be unduly ambitious.

## **B. Reforms**

Consider the following reform possibilities: (a) asking the jury to look at cases from more than one category of harms; (b) increasing judicial control over awards, on the theory that judges are likely to have an implicit comparison set in view, and could be required to make explicit comparisons; and (c) transferring

---

<sup>87</sup> See Sunstein, Schkade & Kahneman, *supra*

<sup>88</sup> See note *supra*.

some or all of the relevant tasks to a bureaucracy, which might create and operate a kind of "damage schedule" for conduct calling for punitive awards of some kind.

*1. Comparisons before the jury.* It is tempting to think that the best response to what we have shown is simply to provide jurors with information about cases in more than one category of harms, so as to overcome the global incoherence that results from isolated judgments. But a moment's reflection should suffice to show that this is not a sensible solution.

The first problem involves cognitive overload. Would specific cases be shown to the jury? From which categories? In what detail? Obviously it would be absurd to provide jurors with all cases from all categories. But any small set of comparison cases would seem, and inevitably be, arbitrary. The second problem involves the risk of manipulation. It would be easy, on the basis of our findings here, to select a category that would drive up both outrage and dollar awards, or to drive them down. A lack of a comparative context is indeed a problem, but comparisons before the jury are not a realistic response.

*2. Increased judicial control: additur, remittitur, and others.* A more promising strategy for reducing incoherence would be to give relatively greater power to the reviewing court and relatively less power to the jury. The basic idea would be that judges are in a better position to move toward global coherence, or at least to avoid the worst anomalies. To be sure, we have said that people are not likely to agree about what pattern of results would really represent "global coherence," and it is possible that judges' views would diverge from those of juries or the public at large. All we are suggesting is that a greater role for judges could produce progress by everyone's lights, if the basic goal is to avoid the most serious misjudgments.

A tempting motivation for this reform would be that by virtue of their experience, judges have a larger implicit comparison set. Having seen a wide range of cases, judges might seem likely to know that a punitive damage judgment of, say, \$400 million for commercial fraud, or \$50,000 for pain and suffering associated with an eight-month hospital stay produced by reckless behavior, is extremely odd when compared to cases falling in other categories. Indeed, the old practice of remittitur, allowing judges to reduce excessive awards, can be understood as an effort to ensure more in the way of global coherence. But there is reason to think that implicit comparison sets are not

enough, and that because judges are human,<sup>89</sup> steps must be taken to counteract the natural tendency to make category-bound judgments. For the process to work, judges must be required to make such comparisons; left to their own devices, they might not do so, or make comparisons only to an idiosyncratic or ad hoc set of cases. In fact courts have expressly suggested that appellate review of awards should be conducted by examining comparison cases, in an effort to promote global coherence.<sup>90</sup> To be sure, these suggestions have been limited to within-category comparisons, but there is good reason to be more ambitious, so as to counteract the problems identified here. We shall return to this point below, in connection with the idea of sentencing guidelines, and the goal of decreasing incoherence in criminal sentencing.

What we have said supports the practice not only of remittitur but of additur as well, through which appellate courts increase awards that are shown to be too low. Undoubtedly some judgments are now insufficiently severe in light of judgments in other categories, and if reducing incoherence is the goal, it would be a good idea to ensure that judges increase those awards. It is most unfortunate that the right to trial by jury is now taken to forbid additur in federal court.<sup>91</sup> We suggest that when they can, judges should use additur and remittitur more self-consciously as efforts to correct judgments that are anomalous when placed in the context of other outcomes. Indeed our findings provide more general support for the view that judges should take a stronger hand in overseeing both high and low awards, on the theory that the predictable incoherence of one-shot jury judgments undermines the populist credentials of those awards. If ordinary people would themselves deem those judgments incoherent, why should judges be unwilling to disturb them? Of course there would be some social cost to keeping the institution of the jury while also

---

<sup>89</sup> There is mounting evidence that judges are subject to the same cognitive processes as everyone else. See, e.g., W. Kip Viscusi; Chris Guthrie et al., *supra*.

<sup>90</sup> *Martell v. Boardwalk Enters, Inc.*, 748 F2d 740, 752-53, 755 (2d Cir 1984); *Geressy v. Digital Equipment*, 980 F Supp 640, 657-60 (EDNY 1997). Cf. *Cooper v Leatherman*, *supra* at 15-16: "Differences in the institutional competence of trial judges and appellate judges are consistent with our conclusion. In *Gore*, 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. 517 U. S., at 574-575. 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."

<sup>91</sup> *Dimick v. Schiedt*, 293 U.S. 474 (1935).

reducing its authority; and those who believe that the civil jury is an important social institution will be skeptical of any effort to reduce its role.

There is an independent concern. If judges have predictable prejudices – if, for example, they are systematically, and improperly, pro-defendant in some cases – the argument for increasing judicial control is greatly weakened. We 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. Coherence is not a trumping value. And indeed there is some (disputed) evidence that judges are less willing to award punitive damages at all, and also that they tend to impose systematically lower awards than juries.<sup>92</sup> Of course this is not evidence that judges are biased; perhaps they are right and juries are wrong. But if the problem lies in judicial bias of some kind or another, there is little reason to approve of the status quo; the remedy for bias may well lie in attempting to discipline jury discretion through exemplars or guidelines, modeled perhaps on the Sentencing Guidelines, designed to impose presumptive floors and ceilings on awards. This suggestion takes us directly to a final reform strategy.

*3. Bureaucratic rationality.* The most radical response to predictable incoherence would be based on the idea of bureaucratic rationality. Here the goal would be to create a kind of "grid" of penalties, matched to the individual facts of diverse cases. As it arises, each case would be fit into the grid. For punitive damages, for example, an agency, overseen perhaps by Congress and working with courts or perhaps entirely independently, might generate a system of financial punishments for different categories of harm. We have seen that there would be significant problems in attracting public agreement on any such system, at least if it attempts to rank harms from diverse categories. But it might well be better to try to ensure that any ranking exercise is explicit rather than implicit, and that the ranking that is observed in practice is something better than an artifact of the cognitive forces that we have discussed here. We will be agnostic here on the substantive judgments that would go into the resulting outcomes. The system might be based on optimal deterrence theory; it might be rooted in retributive intuitions; it might be some kind of mix.

*a. Antecedents.* There is nothing new about the idea of producing more sensible ranking through bureaucratic institutions. Indeed, a great movement of

---

<sup>92</sup> But note that in Theodore Eisenberg & Kevin Claremont, *Trial by Jury or Judge: Transcending Empiricism*, 77 *Cornell L. Rev.* 1124, 1992), it is urged that in product liability and medical malpractice cases, plaintiffs prevail at higher rate before judges than before juries. And in several categories of personal-injury liability, the mean recovery is higher in judge trials.

twentieth century law has been from case-by-case adjudicative decisions toward a system of bureaucratic judgment, cabined by rules laid down in advance. The most familiar example is probably workers' compensation, designed to replace ad hoc judicial judgments with something more systematic.<sup>93</sup> The same kind of reform can be found in the area of social security disability determinations, where a high level of inequality in individual judgments produced a system of "grids" designed to produce greater coherence.<sup>94</sup> The Sentencing Guidelines were similarly responsive to a perception of incoherent judgments,<sup>95</sup> though there the claim was that criminal sentences were incoherent within as well as across categories.

*b. Current initiatives.* Bureaucratic solutions are not foreign to the areas discussed here. Some states have moved in this direction with respect to contingent valuation. Consider, for example, the remarkable approach adopted by the Texas Parks and Wildlife Department. A regulation is designed to establish values of injured or destroyed fish, birds, and animals.<sup>96</sup> These values are determined by assigning a score of 0 to 3 for eight specified "scoring criteria."<sup>97</sup> The criteria include recreation; aesthetic; educational; scarcity; environmental tolerance (that is, the ability of the species to tolerate changes in climate and so forth; economics (economic benefit of the species); recruitment (reproductive and survival capacity); ecological role. After the individual scores are added to produce a "total criteria score," that score is multiplied by a "weighting factor" that relates to the overall demand for the species; the weighting factor ranges from 1.0 (no additional public demand beyond the eight criteria) to 1.5 (for scarcity). The adjusted recovery score produces a recovery value for the species, one that is fully monetized. Special provisions are included for endangered (an additional payment of \$1000) and threatened species (an additional payment of \$500).<sup>98</sup>

<sup>93</sup> See Fishback, *supra* note.

<sup>94</sup> See Jerry Mashaw, *Bureaucratic Justice* (1983).

<sup>95</sup> See Breyer, *supra* note.

<sup>96</sup> See section 69.20.

<sup>97</sup> *Id.* at 69.22.

<sup>98</sup> Pursuant to this rule, the Texas Parks and Wildlife Department has issued a 74-page document listing monetary equivalents; consider, for purposes of illustration, the following excerpt:

Bass, guadalupe, 1 inch	\$ .50
Bass, guadalupe, 21 inch	\$184.19
Mackeral, King, 69 inch	\$404.60
Shark, Sand, 131 inch	\$211.81
Mule Deer , F	\$163
Mule Deer, M	\$525.50

Punitive damages, assessed in court, can be seen as a substitute for administrative penalties, and many agencies impose civil penalties via a method of this kind. As we shall see below, the problem is that agencies tend to operate in isolation, in a way that creates independent risks of global incoherence; the question is whether the existing system of administrative fines might itself be rationalized. Consider, as a model in this regard, state systems for workers' compensation, which explicitly set out damage remedies for various injuries. Perhaps punitive awards could be made part of a general schedule of civil penalties.

The strongest objection to a bureaucratic model would be similar to that raised in the judicial context -- distrust of the relevant bureaucrats (or of the political actors overseeing their operations). If interest-group pressures would distort judgments, or if the resulting "grid" would contain systematic injustice, it may be better to continue with a system that contains a degree of incoherence. But these issues cannot be assessed in the abstract.

Our goal here has not been to argue for any particular reform proposal, but simply to offer a sense of some of the possibilities. With respect to punitive damages, we are confident that we have shown a serious, and thus far unexplored, problem with the existing system. Our emphasis has been on the avoidance of obvious anomalies, not on the achievement of global coherence. We are not confident about a plan for reform, in part because of the sheer range of

---

Sheep, Desert Bighorn	\$2,850
Prairie Chicken, Attwaters	\$7,100.50
Pheasant	\$35.50
Jaguar	\$2,850.50
Mountain Lion	\$525.50
Skunk, stripped	\$15.50
Whopping crane	\$7,100.50
Brown pelican	\$2,850.50
Texas Tortoise	\$163.00
Deer	\$1
Loggerhead Turtle	\$1,050.50
Elk	\$1
Texas Rat Snake	\$3.00
Golden cheeked Warbler	\$7,100.50

The particular judgments may seem a bit arbitrary; why is an elk worth \$1, compared to the \$1,050.50 penalty for killing a loggerhead turtle? What is important is that the Texas provision actually offers an answer to this question, one that is relatively transparent to the public, and one that ensures that the various values line up with one another along the stated criteria.

considerations involved. But we believe that incoherence is a form of injustice, and that it is unquestionably worthwhile to attempt to find solutions.

## **VI. Administrative Penalties**

In the modern administrative state, civil penalties are not assessed in the same way as punitive damages. Congress establishes maximum penalties, generally on a statute-by-statute basis. Congress also establishes enforcement levels via appropriations. Faced with these constraints, agencies have considerable room to maneuver, both in choosing the level of the fine (subject to the statutory maximum) and in deciding whether to bring an enforcement action, and to seek fines, in particular cases. On the basis of our findings here, we hypothesize, not that everyone would agree on what a coherent ranking would look like, but that an investigation of the full range of administrative penalties would show what many people would see as anomalies – that the set of penalties actually sought and imposed could not, in the view of most people, possibly be squared with one another. Those who come up with penalties are not forced to rationalize the system, in part because people do not spontaneously call for rationalization, or even attempt to see whether penalties form a sensible pattern. We attempt to illustrate the problem here, by identifying the statutory maxima and to raise questions about whether they suggest a coherent set of judgments on Congress’ part. We can also suggest two factors that contribute to the existing situation: category-bound judgments and the translation problem.

In the context of administrative penalties, category-bound judgments take a distinctive form. In these situations, legislators or administrators often consider an entire category at once, such as injuries to wild birds. In such cases, the emotional response attached to the category label will be of critical importance. An entire category might be misjudged in separate evaluation (where misjudgment is understood by reference to people’s own beliefs about the sense or nonsense of the resulting overall pattern). Of course political pressure can be influential here. As we will see, the punishments for occupational safety and health violations are relatively light; we are confident that political pressures are a large part of the reason. But the difficulties that we have identified – category-bound thinking and the translation problem – are partly responsible for the incoherence. It would be much harder, politically speaking, to bias judgments toward an entire category if translation were easy and if explicit comparisons to another category were brought into the picture.

## A. Background

One of the most important movements of the post-New Deal state has been the emergence of administrative penalties for wrongdoers. If a company or individual has violated regulations involving occupational safety, or protection of endangered species, some sanction is in order. Often criminal penalties are deemed too severe for regulatory violations, and civil penalties, in the form of fines, take up the slack. But what is the appropriate level of such penalties? What is the proper frame of reference?

1. *Overview.* Our starting point is that administrative penalties are generally assigned at the same time to a set of harms or risks that belong to a single category, in a process that does not encourage the consideration of other categories of harms or risks. Adapting the conclusions of our analysis of punitive damage awards to judgments of categories, we can suggest the following:

- Categories of harms vary in the intensity of the emotions that they evoke;
- moral intuitions do not specify a scaling factor for the task of translating punitive intent into dollars;
- there is almost complete consensus on the relative positions of some categories, and substantial consensus on others;
- because harms from different categories are not transparently commensurable, a misranking of harms from different categories is unlikely to attract attention;
- organizations and individuals setting penalties for categories of harms consider them one at a time, and determine a special modulus for each occasion.

This list of psychological considerations suggests three hypotheses about administrative penalties. First, the ranking of harms and penalties within each category of harms will generally reflect public sentiment. Second, the levels of penalties for *different* categories of harms will not necessarily reflect the ranking of these categories in the public sentiment. Third, incoherence of penalties for different categories of harm is likely to be tolerated and to persist over time. As we shall see, each of these hypotheses is confirmed.

2. *Politics*. The idea that there is no intuitively compelling way to scale punishments suggests that political dynamics will play a central role in driving penalties up or down.<sup>99</sup> Where interest groups opposing the regulatory scheme are well-organized, we should expect low penalties; when such groups are poorly organized, and when those enthusiastic about the regulatory plan have considerable political power, we should expect severe penalties.<sup>100</sup> The translation problem compounds the difficulty. Well-organized groups might well take advantage of the arbitrariness of judgments about appropriate dollar penalties, simply because the public will have a hard time knowing whether one or another number is excessive or insufficient – at least if cross-category comparisons are not made. The translation problem thus creates a lack of transparency in civil fines, including administrative penalties. A wider view screen would of course increase transparency.

There is an independent point. The administrative state is heavily balkanized, with largely independent institutions engaged in independent tasks, and with separate congressional committees engaged in oversight. The Occupational Safety and Health Administration has its own network of duties, under the auspices of the Department of Labor; so too with the Fish and Wildlife Service, under the Department of the Interior; and so too with the Environmental Protection Agency. In an enterprise as enormous as the national government, some such division of labor seems inevitable. But without a mechanism for ensuring coherence, the effect of the balkanization, at the level of committees and regulatory bodies, might well be to produce incoherence, for the reasons we have traced here. Does the system of regulatory penalties show the expected pattern of within-category coherence and cross-category incoherence? In asking that question, we do not mean to require observers to ask complicated normative questions about appropriate penalties, but to focus attention on a question related to that posed in previous sections: If the various penalties are put together, would ordinary people conclude that the system of penalties makes little sense? If so, we claim, consistently with what we have urged above, that the system is unjust, or at least that it contains significant areas of injustice.

---

<sup>99</sup> For dramatic evidence, see Price Fishback & Shawn Kantor, *supra* note, at 172-89.

<sup>100</sup> *Id.* at 12-14.

## B. The Pattern of Penalties

Consider the following table, capturing a subset of the universe of administrative penalties (see Appendix D for a fuller version):

**Table 2: Selected Civil Penalties from Administrative Agencies**

Agency	Civil/ Criminal	Type of violation	Penalty
OSHA	Civil	Violation of a serious nature	Up to 7000
	Civil	Violation of non-serious nature	Up to 7000
FDA	Civil	Egg products inspection act- violation	Up to 1000
	Civil	Above + intent to defraud or distribute adulterated eggs	Up to 10000
Fish and Wildlife Service	Civil	Violation of Wild Bird Conservation Act	Up to 25000
	Civil	Violate WBCA 111 a1 or 2 (import in violation of ban )or permit under 112 (exemptions for sci research and ltd breeding)	Up to 25000
	Civil	Civil violation of African Elephant Conservation Act	Up to 5000
	Civil	Violation of Bald Eagle Protection Act	Up to 5000
EEOC	Civil	Give public information about unlawful employment practice under T7 during proceeding	Up to 1000
	Civil	Discriminatory practice with malice	Punitive damages
	Civil	Total damages under t7 for small business (under 100 employees)	Up to 50,000
	Civil	Total damages under T7 for medium business (over 100 less than 200 employees)	Up to 100,000
	Civil	Total damages under T7 for large business (Over 200 less than 500)	Up to 200,000
	Civil	Total damages under T7 for big business (more than 500)	Up to 300,000

	Civil	Willful violation of section 15 of equal pay act	Up to 10,000
	Criminal	See above	Up to 6 mo. prison

Notice that within categories, penalties seem extremely sensible, at least in the sense that the more serious harms are punished more severely. For occupational safety and health violations, the largest penalties are for repeated violations, the next largest for violations that are both willful and serious; the least serious for failures to engage in the requisite record-keeping. The Wild Bird Conservation Act imposes its most serious penalties for unlawful imports of wild birds or members of endangered species; penalties are less severe if violations are not undertaken as part of import-export business, or if they are not knowing and willful. The most serious acts of discrimination are subject to more serious penalties. In general, each of the regulatory arrangements shows a high degree of internal consistency.

At the same time, it is clear that this system contains serious anomalies. The penalties do not fit together as a coherent whole, and it is most unlikely that most people would approve of all of what appears here. The most obvious set of anomalies involve OSHA. A serious violation of OSHA calls for a maximum penalty of \$7000; unlawful importation of a wild bird calls for a penalty of more than triple that amount (\$25,000). In fact the penalties imposed by the Fish and Wildlife Service are systematically higher than those imposed by OSHA. It is extremely doubtful that the public would support that pattern of penalties, taken as a whole.

Of course it is possible that agency enforcement patterns create sense out of nonsense. Perhaps those statutes with low maximum fines are frequently enforced, and perhaps enforcement activity, in such cases, involves maximum fines – whereas statutes with anomalously high maximum fines might be enforced less frequently, and agencies might seek fines toward the low end. We would be surprised if this were true. But for present purposes, the key point is that no one seems even to have asked whether the resulting pattern of fines, on the books or in the world, makes any sense at all.

### **C. Difficult Cross-Category Comparisons**

It is relatively easy to compare OSHA violations with OSHA violations; it is much harder to compare OSHA violations with Endangered Species Act

violations or with acts of unlawful discrimination. To be sure, categories themselves can probably be compared. Offhand, most people are likely to think that it is worse to subject someone to life-threatening working conditions than to discriminate against them. It is reasonable to suppose that the category of occupational safety and health violations ranks higher than the category of unlawful discrimination, Hence it is odd that an act of discrimination can receive a penalty of \$50,000 and more, whereas an OSHA violation is subject to a \$7000 ceiling. But many cross-category comparisons are much harder to make. An egregious act of discrimination (say, firing a long-time employee on the basis of racist motivations, and attempting to ruin that employee's career) is difficult to compare with a milder case of occupational safety violation (say, failing to ensure that ladders and toilets comply with OSHA regulations).

Illustrations could easily be multiplied. A general conclusion follows. A consensus is hard to achieve on the content of full coherence (return to the epigraph from Justice Breyer); but it would be much easier to achieve general agreement that certain outcomes are truly anomalous. This point is directly relevant to the question of institutional reform.

#### **D. A Civil Penalties Commission?**

The data alone cannot show how and to what extent it would be appropriate to revise the set of administrative penalties, once they are taken as a whole. The initial problem is that to our knowledge, the pattern of penalties has not, until now, even been presented as such. The oddity to which we draw attention is the remarkable tolerance of the system for a set of outcomes that do not fit together and that show significant anomalies. Legislators and administrators appear to have made judgments within categories without investigating whether the resulting judgments can be criticized as palpably incoherent. Whatever the difficulties in producing a rational pattern – and they are formidable – at least it seems clear that some effort should be made to correct the worst outliers. Because courts must take statutes and regulations as they stand, there is little that judges can do about the problem. Instead reforms must come from executive and legislative spheres.

In these circumstances we suggest three possible routes to reform. First, the Office of Information and Regulatory Affairs (OIRA), within the Office of Management and Budget (OMB), should provide a full accounting of regulatory penalties, publicize it, and evaluate the existing pattern of outcomes, with recommendations both to agencies and to Congress. The longstanding missions

of OIRA in particular and OMB in general have been to produce more in the way of global rationality, with respect to regulation and the budget generally.<sup>101</sup> This role should be extended to better rationalization of the system of administrative penalties. Note in this regard that even without legislative change, agencies have considerable room to maneuver. Seeing that occupational safety violations are low, they might choose penalties at the high end of the permissible range. Seeing that penalties for violations of the Wild Bird Conservation Act are relatively high, the Fish and Wildlife Service might select penalties at the low end of the permissible range. The purpose of executive oversight would be to move agency practice in the direction of a sensible overall pattern of penalties. Of course we have suggested that coherence is difficult to achieve, in part because people do not agree on what it requires, because they do not find it easy to rank cases from diverse categories. But at the very least, a coherence commission could correct palpable errors; and perhaps it could try to be more ambitious.

The second route for reform involves Congress. On both the House and Senate sides, we suggest that a new subcommittee should be created on administrative penalties, growing out of existing committees on regulation and administrative practice. The purpose of the subcommittees should be to gather information about administrative penalties, as described in statutes and regulation, and also about actual agency practices – with the goal of publicizing the results and ultimately enacting corrective legislation, at least in response to the most obvious anomalies. Nothing said here establishes the appropriate content of that legislation. But we believe that we have established that if the overall pattern is not presented before citizens and representatives, incoherence and injustice are inevitable results.

A third route would involve a mixture of legislative and administrative action. Acting on the model of the Sentencing Commission, Congress should create a Civil Penalties Commission, whose responsibility would be to give a public accounting of civil penalties, system-wide, and to propose standards for rationalizing them. The simplest task of the Civil Penalties Commission would be to correct the most obvious errors; as we have seen, it would be much harder to produce agreement on what would make the full set of outcomes coherent. More generally, the Civil Penalties Commission should give a sense of the relationship between the law on the books (including statutory maxima and minima) and actual practice (for example, the frequency, median, average, and range of penalties). Like the Sentencing Commission, it should be composed of

---

<sup>101</sup> See Thomas McGarrity, *Rationalizing Regulation* (1992).

diverse people; it might well include judges. The Civil Penalties Commission would have the authority to issue guidelines to discipline administrative agencies, so as to increase the likelihood of coherence. Perhaps the guidelines should be binding rather than merely advisory. In the interest of public disclosure, the Civil Penalties Commission should maintain a website containing information about current practice, allow comparisons, and discussing proposals. Of course the actions of the Civil Penalties Commission would be subject to congressional review.

We conclude this section by commenting on similarities and differences between this discussion of administrative penalties and the earlier discussion of punitive damages and contingent valuation. The critical difference is that penalties are now set by category, not by case. The main source of incoherence is not normalization, but the arbitrariness inherent in the use of dollar values (or indeed, of other numerical scales such as time in jail) to express the intensity of the response to harms and to the people who perpetrate them. What is common to findings in both types of problems is that people's thinking does not spontaneously cross boundaries between categories; when setting penalties for violations of the fish and wildlife statutes, people do not naturally ask about penalties for violation of the tax and pollution laws. Indeed, in the current arrangement, an administrator charged with attention to one category of problems who actively considered problems from other categories might be accused of wasting resources or even of poaching on the turf of other officials. In addition, a full understanding of administrative penalties would require an exploration of political dynamics and, in particular, the role of interest groups in assigning dollar values to regulatory violations. Interest group power is undoubtedly magnified by the difficulty of knowing what different amounts mean in the context of the system as a whole. As we have suggested, interest group power would be greatly reduced if comparisons were routine – if the penalties established under one statute were routinely evaluated against a number of penalties from other categories. We think that such comparisons would be highly likely to increase the sense and rationality of a system that, on both counts, currently falls short.

## **VII. Predictable Incoherence: Extensions, Speculations, Remedies**

If one-shot judgments are predictably incoherent in the contexts discussed thus far, it is likely that the legal system is beset by similar problems elsewhere. We now broaden the viewscreen, not to resolve the issues with which we will

deal, but to show that related difficulties can be found in other domains. The overarching point is that incoherence is likely to be created, and to persist, wherever cases are considered one at a time, because of the translation problem, category-bound thinking, or both. Our goal here, as elsewhere, is not to suggest that agreement on full coherence is easy or even possible to achieve, but instead to emphasize the possibility of correcting the worst abuses once the viewscreen is widened and comparisons are made.

### **A. Criminal Penalties: A Cognitive Perspective on Guidelines**

To say the least, the topic of sentencing guidelines has received massive attention.<sup>102</sup> Our findings cast new light on the value of such guidelines – not only to overcome differences among judges (an original inspiration for the guidelines<sup>103</sup>), but also to reduce the predictable problems created when people are selecting penalties one case, or one category, at a time (a difficulty that the guidelines do not attempt to solve).<sup>104</sup> If we are right, judicial sentencing without guidelines is likely to produce incoherence, simply because people do not naturally seek to put individual decisions into a sensible overall pattern. The same problems will arise for legislatures imposing criminal punishments by categories. And if those who create sentencing guidelines are not alert to the problems, they will not solve them.

To be sure, it is not impossible that judicial experience with a wide range of cases could reduce the problem of incoherent sentences. It is conceivable that by virtue of their experience, judges engaged in sentencing have a kind of “sentencing menu” in their heads, reducing the difficulties we have found here. On the other hand, our evidence suggests that the difficulties will persist unless they are explicitly drawn to people’s attention. Whether and how judges can overcome the relevant problems remains an empirical question, one that we discuss in more detail below. But even the most experienced judge must engage in isolated sentencing decisions, and it is highly likely that without guidelines, within-category coherence and global incoherence will be the result. To the extent that sentencing guidelines can reduce this problem, they should provide large improvements.

---

<sup>102</sup> See, e.g., Albert Alschuler, *The Failure of Sentencing Guidelines*, 58 U Chi L Rev 901 (1991).

<sup>103</sup> See *id.*

<sup>104</sup> The reason is that the Sentencing Commission built on past practice, without attempting to engage in any more ambitious effort at rationalization. See Stephen Breyer, *The Federal Sentencing Guidelines and the Key Compromises Upon Which They Rest*, 17 Hofstra L. Rev. 1, 14-19 (1988).

But we do not mean only to provide support for the idea of federal sentencing guidelines. Our discussion suggests that there is good reason for much more in the way of rationalization at the state and federal levels. Within the states, there continue to be many anomalies. In Illinois, for example, an assault – defined as knowingly or recklessly causing serious bodily injury – is a misdemeanor with a penalty of up to thirty days imprisonment<sup>105</sup>; an assault against a public servant is also a misdemeanor, with a maximum penalty of one year in jail.<sup>106</sup> So far so good. But possession of an endangered plant for commercial use is treated the same way, with a maximum one year sentence,<sup>107</sup> and destruction of a nest of wild birds is a felony, treated more seriously than many violent crimes, with a maximum sentence of two to five years.<sup>108</sup> As far as we are aware, no state has made even a minimal effort to rationalize its system of criminal punishment. In fact we are aware of no work, academic or otherwise, that attempts to explore whether state penalties fit into a sensible whole.

At the same time, the federal sentencing guidelines do not avoid the problem of global incoherence. As we have noted, the Sentencing Commission built on existing judicial practice, and did not attempt to overcome the problems discussed here.<sup>109</sup> Justice Breyer's suggestion, quoted above, is that the Commission's relative lack of ambition was a product of the sheer difficulty of the task of achieving agreement on rankings across categories. Most of the ongoing work of the Sentencing Commission has involved the achievement of within-category coherence, a much easier task. The existence of extremely high penalties for drug offenses, alongside lower penalties for crimes of violence, attests to the problem. But the problem goes deeper still, involving not only incoherence but also complexities in defining categories.

The most highly publicized anomalies involve the treatment of powder and crack cocaine.<sup>110</sup> Under federal law, it takes one hundred times as much powder cocaine as crack cocaine to trigger equivalent mandatory penalties, so that the mandatory five-year penalty applies to someone with 500 grams of powder cocaine and five grams of crack cocaine.<sup>111</sup> Perhaps powder cocaine and crack cocaine should be seen as different categories, rather than subcategories of "cocaine," because of the fact that crack cocaine is more often associated with

---

<sup>105</sup> 720 ILCS 5/12-1.

<sup>106</sup> 720 ILCS 5/12-2.

<sup>107</sup> 520 ILCS 10/9.

<sup>108</sup> 520 ILCS 5/2, 36a.

<sup>109</sup> See Breyer, *supra* note.

<sup>110</sup> United States Sentencing Commission, *Cocaine and Federal Sentencing Policy* (April 1997).

<sup>111</sup> *Id.* at 2-3.

systemic crime.<sup>112</sup> For current purposes, what is most important is that the comparison with the two has produced the widespread view that penalties for crack cocaine are far too severe.<sup>113</sup> This view is itself evidence for our basic claim here: It is easy to suggest that powder cocaine and crack cocaine belong to the same category, and hence easy to generate a public outcry against the within-category inconsistency. In the abstract, it might be hard to know whether one or another mandatory minimum is best, perhaps because of the difficulty of finding a “modulus” with which to scale. The comparative exercise produces public concern.

This is simply one example. The Sentencing Commission, and Congress itself, should engage in far more in the way of cross-category comparisons, designed to reduce the largest anomalies, in the form of both excessive severity and unwarranted leniency.

## **B. Contingent Valuation**

The practice of contingent valuation raises many complex questions, and it is not clear that the practice can be defended.<sup>114</sup> Recall that contingent valuation is a procedure in which people are asked to scale their attitude to different harms, and to public actions to correct or mitigate harms, on a scale of hypothetical willingness to pay. As we have seen, the psychological difficulties associated with the task of a respondent in a contingent valuation survey are precisely analogous to the difficulties faced by a juror in determining punitive damages.<sup>115</sup> The translation problem – the absence of a principled or intuitively compelling way to use the dollar scale – arises in two different ways. First, the numbers that people provide for any specific problem are arbitrary, and therefore not a sensible guide for policy. Second, the normalization process tends to distort the valuations of items from different categories. We have seen that ecological problems gain a great deal from being presented in isolation, and that health problems lose from that style of presentation. At the very least, people answering contingent valuation questions should not be restricted to a small set of problems from one or more categories.

The most modest reform would attempt to ensure that any responses to contingent valuation questions are given after exposure to sufficient information

---

<sup>112</sup> *Id.* at 4.

<sup>113</sup> *Id.* at 9-10.

<sup>114</sup> See Peter Diamond & Jerry Hausman, *Contingent Valuation: Is Some Number Better Than No Number?*, 8 *J Econ Persp* 45, 51-52 (1994)

<sup>115</sup> Kahneman, Ritov & Schkade, *supra*.

to counteract the difficulties discussed here. Some people have attempted to design a contingent valuation study that offers a wide range of categories, so as to diminish the risk of incoherence.<sup>116</sup> Respondents could be asked about problems from a number of cases at once, so as to ensure that any particular answer follows from cross-category comparisons. There are, however, problems with this approach, akin to those discussed in the context of punitive damages: information overload, framing, and manipulation.<sup>117</sup> People can process a small set of cases, but any such set risks framing effects; a full range of cases, sufficient to overcome those effects, might overload people's cognitive capacities. It is not clear that this modest reform could ultimately succeed.

A much less modest reform, based on the approach in Texas, would be to move away from contingent valuation and toward a damage schedule for environmental harms, based on a combination of technocratic and democratic judgments.<sup>118</sup> Such a schedule would be produced by people considering the full range of cases. The most serious risk of such a reform is that coherence would be purchased at the price of increasing interest-group power over the determination of penalties. The extent of this risk cannot be assessed in the abstract.

Because the evaluation of these reforms raises many of the same questions as in the context of punitive damages, and because there is an emerging literature on the point,<sup>119</sup> we will not discuss them in any detail here. What we have added is a fresh reason for skepticism about many contingent valuation studies: They do not require people to engage in cross-category comparisons, and hence they generate outcomes that people would not themselves endorse, taken as a whole.

### **C. Regulatory Crazy-Quilts**

Related issues arise in the continuing debates about expenditures on life-saving expenditures. The central point here is that agencies generally decide on the appropriate extent of regulation without being required to engage in cross-

---

<sup>116</sup> There are, however, problems with this approach, akin to those discussed in the context of punitive damages, including information overload, framing, and manipulation.

<sup>117</sup> For example, there are large anchoring effects when several environmental goods are evaluated in sequence – when a highly valued good comes first, the willingness to pay for the goods that follow is higher than when a lower valued good comes first. See John Payne, David Schkade, William Desvousges, & Chris Aultman. Valuation of multiple environmental programs, 21 *Journal of Risk & Uncertainty* 95 (2000).

<sup>118</sup> Kahneman & Ritov, *supra* note.

<sup>119</sup> See Murray Rutherford et al., *Assessing Environmental Losses*, 22 *Harv J Env L.* 51 (1998); Richard B. Stewart & Richard Revesz, *Reassessing Superfund* (1995).

category comparison (and indeed with little in the way of constraints from Congress<sup>120</sup>). Agencies generally act in isolation, and without specifying the regulatory expenditure per life saved; this institutional fact has increased the difficulty of ensuring global coherence. For many years, agency actions were categorized by subject matter, and no one attempted to ensure cross-category comparisons, let alone to see if the relevant actions could be seen as falling, in some sense, within the same category.

In 1986, an official at the Office of Management and Budget had an extremely clever, psychologically acute idea, and a simple table was compiled, allowing comparisons of agency practice.<sup>121</sup> That table has come to dominate many discussions of these problems.<sup>122</sup> (See the Appendix E for a recent version). The table seemed to have such an effect in part because it made the underlying numbers explicit, and in part because it makes comparisons possible across programs. Indeed, the dramatic effect of the table seemed to stem from the fact that it put a wide range of seemingly separate problems under a single “category”: government efforts to save lives through regulatory protection. Once the table was constructed, unnoticed anomalies became highly visible, because they seemed, all of a sudden, to involve within-category nonsense, to which people are extremely sensitive. It should be no surprise that critics of the table have urged that separate categories are involved here, and that people should not be so quick to think that the current situation is senseless.<sup>123</sup> This response makes perfect psychological sense.

A prominent study of five hundred life-saving interventions goes much further.<sup>124</sup> The study shows systematic disparities in expenditures across agencies, with some problems receiving apparently disproportionate attention. Consider, for example, the fact that the median cost per life year is \$23,000 for the Federal Aviation Administration; \$68,000 for the Consumer Product Safety Commission; \$78,000 for the National Highway Traffic Safety Administration; \$88,000 for the Occupational Safety and Health Administration; and several million dollars for the Environmental Protection Agency.<sup>125</sup> Here too cross-

---

<sup>120</sup> See Cass R. Sunstein, *Cost-Benefit Default Principles*, *Mich L Rev* (forthcoming 2001).

<sup>121</sup> See John Morrall, *A Review of the Record*, *REGULATION*, Nov./Dec. 1986, at 25, 30 tbl. 4.

<sup>122</sup> A skeptical overview is Lisa Heinzerling, *Regulatory Costs of Mythic Proportions*, 107 *Yale LJ* 1981 (1998).

<sup>123</sup> See Heinzerling, *supra* note.

<sup>124</sup> See Tammy O. Tengs et al., *Five Hundred Life-Saving Expenditures and Their Cost-Effectiveness*, 15 *Risk Analysis* 369 (1995).

<sup>125</sup> *Id.* at 371.

category comparisons has played a large role in reform debates,<sup>126</sup> suggesting that American government could save the same number of lives it now saves with \$31 billion in savings – or that we could use the same level of resources we now use, and end up saving 60,000 more lives every year

To be sure, this is not a “pure” case of the kind that we have been emphasizing throughout: Many of the relevant government actions were apparently undertaken without explicit decisions about the relevant numbers, or about how much should be spent in different areas. Hence the evidence shows apparently inexplicable disparities within agencies. We believe that an investigation of those anomalies, and an attempt to correct them, would be likely to alter government behavior, and perhaps to save money and lives in the process.

## **VIII. Cognitive Notes on Legal Theory**

Our findings here bear on two intensely debated issues in legal theory. The first involves the aspiration to fully coherent law. The second involves the idea of “incommensurability.” We will not engage the normative issues here, but we will suggest that an understanding of cognitive issues illuminates both debates.

---

<sup>126</sup> The role of cross-category comparisons is sharply contested in this domain. In a prominent article Lisa Heinzerling, for example, has deplored “quantification,” suggesting that the table printed above gives an illusion of precision. See Lisa Heinzerling, *Regulatory Costs of Mythic Proportions*, 107 *Yale LJ* 1981 (1998). In her view, this is an illusion because many of the numbers depend on contestable value judgments, involving (for example) the appropriate discount rate, and because quantitatively identical risks should be treated differently if, for example, they involve involuntary exposure and uncontrollable harm.

Even if Heinzerling makes some legitimate criticisms here – and we believe that she does – we think that her skepticism about quantification ignores some important points about human cognition, of central relevance to regulatory policy. First, people will often have a hard time in assessing the appropriate degree of regulation without some assistance from seeing the pattern of expenditures used in other cases. With respect to regulatory expenditures, legislators, regulators, and others are too often in the position of those attempting to impose punitive damage awards in an empirical vacuum. Second, and of special relevance to our claims here, an attempt to see the overall pattern of expenditures, in monetary terms, can greatly aid the process of producing coherent regulation, simply by bringing into view the existence of other categories and a sense of how problems within those categories are treated. Cross-category comparisons, even fairly crude ones, can help people to generate more sensible patterns even if the “bottom line” numbers are not treated as determinative, and even if it makes sense to say that some problems deserve more attention than others, even if they are quantitatively identical. In short, we think tables of this sort can help to produce more rational regulation for those who insist that it would make no sense to insist on a mechanical number for every regulatory program.

## A. Categories, Coherence, Cognition

The last years have seen an intense debate about the aspiration to global coherence in law.<sup>127</sup> Much of the debate has involved the appropriate conception of legal reasoning, an issue with both normative and descriptive dimensions. Ronald Dworkin has been the most prominent advocate of ambitious thinking, in which judges do not always restrict themselves to small pockets of problems, but sometimes attempt to ensure that all parts of law fit together as a principled whole.<sup>128</sup> Others (including one of the current authors) have argued against this idea, on the ground that it would strain judicial capacities, and perhaps tend to produce errors of its own.<sup>129</sup> In a discussion of particular relevance to our topic here, Dworkin discusses the “compartmentalization of law into separate departments” and sees that as a “prominent feature of legal practice.”<sup>130</sup> Hercules, Dworkin's idealized judge, makes judgments that “expand out from the immediate case before him in a series of concentric circles,” increasingly far afield from the particular case and category at hand.<sup>131</sup> Hercules does not take the law's compartments for granted, and he is willing to reject any idea of “local priority” where “traditional boundaries between departments have become mechanical and arbitrary.”<sup>132</sup> It seems clear that Hercules is willing to attempt global coherence, as, for example, through “a new unification of private law that blurs even the long-established and once much firmer boundary between contract and tort.”<sup>133</sup> Advocates of less ambitious thinking stress that judges are not in a good position to make global sense of multiple areas of the law, and that the effort to try, especially if it is early, might well overwhelm the cognitive capacities of judges.<sup>134</sup>

Our findings here do not resolve these issues, which have arisen in a much broader context than those discussed here. We have assessed judgments within and across categories, but in the relatively simple setting of judgments

---

<sup>127</sup> See Ronald Dworkin, *Law's Empire* (1985); Cass R. Sunstein, *Legal Reasoning and Political Conflict* (1996); Cass R. Sunstein, *One Case At A Time* (1999); Edward McCaffery, *Inside-Out Freedom's Law*, 85 Cal. L. Rev. 1043 (1997).

<sup>128</sup> See Ronald Dworkin, *In Praise of Theory*, 29 Ariz State L J 353 (1997).

<sup>129</sup> See Sunstein, *supra* note.

<sup>130</sup> Dworkin, *Law's Empire*, at 251.

<sup>131</sup> *Id.* at 250.

<sup>132</sup> *Id.* at 253.

<sup>133</sup> *Id.* at 254.

<sup>134</sup> See Sunstein, *supra* note; see also Edward McCaffery, *Inside-Out Freedom's Law*, 85 Cal. L. Rev. 1043 (1997) (arguing that cognitive limitations might justify a less ambitious approach from courts, and that this less ambitious approach is consistent with Dworkin's general understanding of legal interpretation as integrity).

about penalties rendered in terms of dollars or years. We have not explored how the process of category-bound judgments would affect free speech principles in categories involving pornography, commercial speech, and libel, or tort and contract principles involving medical malpractice, building construction, and prescription drug companies. It would be important to know whether (for example) people's judgments about the appropriate treatment of false commercial speech and libelous speech would be different in isolation from what they would be if the two categories were considered together. We do not have evidence on that issue. Nonetheless, we think that what we have found casts some fresh light on this debate, giving a more detailed account of why it is hard to produce global coherence, but also providing new support for Dworkin's view, on the ground that local pockets of coherence might well produce an overall pattern that is senseless, or that at least contains what everyone would see as senselessness.

To be sure, global coherence would be a significant strain on judicial capacities, in part because the mental operations involved do not come naturally. Judicial efforts to provide more limited forms of "local coherence," through relatively unambitious reasoning by analogy, can be understood as a good way of avoiding cognitive overload – by focusing on the cognitively manageable category, and by failing to investigate problems from other categories. But our larger point cuts the other way, giving strong cognitive ammunition to Dworkin's plea for global coherence. We have provided fresh reason to believe that in law, the various categories showing internal coherence will not fit together -- and that the pattern of outcomes, generated by unambitious judges, will contain what they themselves would see as error and confusion. Referring precisely to this risk, Dworkin suggests that we "must strive, so far as we can, not to apply one theory of liability to pharmaceutical companies and a different one to motorists, not to embrace one theory of free speech when we are worried about pornography and another when we are worried about flag burning."<sup>135</sup> If law is to be coherent, Dworkin is entirely correct. Judges who seek only local coherence, or who proceed one case at a time, are highly likely to produce a pattern of outcomes of which they themselves would disapprove.

## **B. Commensurability and Incommensurability**

Are values or options commensurable?<sup>136</sup> In what sense? Those who object

---

<sup>135</sup> *Id.*

<sup>136</sup> See Joseph Raz, *The Morality of Freedom* (1985); Elizabeth Anderson, *Value in Ethics and Economics* (1993).

to the idea of commensurability claim that people lack a shared metric by which to assess an array of qualitatively distinctive options, and that the use of such a metric can do violence to our considered judgments about how such options should be assessed.<sup>137</sup> On this view, ideas like “utility” and “efficiency” are quite inadequate, as a way of capturing the operation of practical reason in law or daily life; these ideas are inadequate because they elide qualitative differences that matter when people reason well.<sup>138</sup>

We do not attempt to take a stand on any normative issues here, nor do we attempt to give anything like an account of the operation of practical reason. But we do suggest that our findings here help to establish a cognitive basis for part of the experience of “incommensurability” in both law and ordinary thinking. People lack confident judgments about how to rank cases from different categories, and their judgments on this point are not widely shared. Recall the difficulty of comparing a case of tax fraud with an occupational safety and health violation, or of ranking a case of outrageous commercial fraud with one involving a relatively minor physical injury. For those using ordinary intuitions, there is no readily available metric by which to make the relevant comparisons.<sup>139</sup> It is in this particular respect that incommensurability is a concrete psychological phenomenon.<sup>140</sup>

## **IX. Conclusion**

We have attempted here to identify some pervasive features of human judgment and cognition – features that, we believe, account for significant anomalies in both private and public law. In making many judgments in law, people must translate a moral judgment into numerical terms, involving dollars or years. Moreover, people’s judgments are insistently category-bound. They do not naturally seek coherence across categories. Their assessment of problems, taken in isolation, are often different from their judgments about problems, taken in the context of cases from other categories. This is largely because any judgment, in isolation, is made against a background of a “natural” comparison set, consisting of problems from the same basic category. Much of the time, people will look at problems from other categories only when forced to do so. When a problem from a different category is introduced, the isolated judgment is

---

<sup>137</sup> See Anderson, *supra* note.

<sup>138</sup> See *id.* at 203-15.

<sup>139</sup> Here economics can overcome the incommensurability problem by using the metric of dollars.

<sup>140</sup> See also the discussion of ‘taboo tradeoffs’ in Philip E. Tetlock, *Coping With Trade-Offs: Psychological Constraints and Political Implications*, in *Elements of Reason* 239 (Arthur Lopia et al. eds 2000).

unsettled, and people's judgments will shift, sometimes quite dramatically. The reason is that the introduction of the new problem alters the set of comparison cases, and shifts in judgment are a common consequence of that alteration.

The most important implication of this phenomenon is that judgments in isolation will predictably produce incoherence from the standpoint of the very people asked to make those judgments. This is true of judgments about punitive damage awards; it is also true of willingness to pay for public goods. Thus judgment shifts are easy to generate in experimental settings. Outside of those settings, we have seen similar results in the domains of regulatory penalties. The pattern of within-category coherence, and global incoherence, is a nearly inevitable product of adjudication, which is defined by one-shot judgments; but the same pattern is embedded in many domains of law and policy.

These are descriptive points. It is far less clear what to do about the situation. Introduction of a single problem from one separate category may make things worse rather than better. There is a serious risk of manipulation and strategic behavior here; careful selection of the comparison case can drive judgment in predictable directions. Deaths of buffaloes may seem a relatively small problem when presented alongside deaths of human beings; but deaths of buffaloes may seem a relatively large problem when presented alongside injuries to plants. In these circumstances, the ideal solution, for a legal system committed to obtaining people's reflective judgments, is to move in the direction of ensuring, not attention to one category or two, but simultaneous appreciation of the large number of categories of cases to which any particular case might be compared.

Coherence is important; it seems to be a minimum requirement of rationality. But coherence is not a trumping value, and a system displaying incoherence may well be better than one that is coherent but pervasively unjust. An incoherent system in which penalties fit together, but are three times as high, or one-third as high, as they ought to be. Nonetheless, we think that any domain of law should aspire to coherence, at least as a presumption, in order to prevent the kinds of arbitrariness and injustice that we have found in both experimental and real-world settings. 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.

We have emphasized throughout that for any single person, or jury, the achievement of coherence is an exceptionally difficult cognitive task. But steps in the direction of coherence are far less difficult at the level of institutional design.<sup>141</sup> We close with the suggestion that the practical remedies for predictably incoherent judgments are institutional; they involve the creation of frameworks for decision that ensure a pattern of judgments that, when taken as a whole, reflective people could endorse. Perhaps this seems an unrealistically ambitious aspiration. But something of just this sort underlies many of the most impressive institutional innovations of the twentieth century. It should not be too much to expect twenty-first century institutions to build on these precedents, bringing a measure of rationality and sense to areas of the law that now lack them.

---

<sup>141</sup> This is a lesson of the movement to workers' compensation, as described in Fishback & Kantor, *supra* note.

## **Appendix A. Excepts from Instructions**

### **Background on Civil Legal Cases and Your Role**

In this part of the study, we would like you to imagine that you are a juror for a legal case in a civil court. Civil law suits involve disputes between private individuals, companies, or individuals and companies, in which the plaintiff alleges that the defendant harmed them in some way. The primary purpose of a civil suit brought by a plaintiff is to seek compensation from the defendant for the alleged harm.

Civil suits involve two different types of potential damages that a defendant could be required to pay. Compensatory damages are intended to fully compensate a plaintiff for the harm suffered as a result of the defendant's actions. Punitive damages are intended to achieve two purposes: (1) to punish the defendant for unusual misconduct, and (2) to deter the defendant and others from committing similar actions in the future.

Punitive damages should be awarded if a preponderance of the evidence shows that the defendant acted either maliciously or with reckless disregard for the welfare of others. Defendants are considered to have acted maliciously if they intended to injure or harm someone or their property. Defendants are considered to have acted with reckless disregard for the welfare of others if they were aware of the probable harm to others or their property but disregarded it, and their actions were a gross deviation from the standard of care that a normal person would use.

The case you will consider involves a special procedure that is sometimes used in civil trials, and which requires two different juries: (1) a trial jury, which decides whether the defendant should pay compensatory damages to the plaintiff, and if so in what amount, and (2) if compensatory damages are awarded, a separate punishment jury decides whether the defendant's conduct also warrants punitive damages.

Please imagine that you are a member of the punishment jury for this case. Your job is to decide whether and how much the defendant should be punished, in addition to paying compensatory damages.

In the case you will consider, the trial jury has already ordered the defendant to pay compensatory damages to the plaintiff. This does not necessarily mean that punitive damages must also be awarded. Whether or not punitive damages should be awarded and if so how large they should be is completely separate from compensatory damages.

## Appendix B. Example Pair of Legal Cases<sup>142</sup>

### B.1. Marking v. Royal Motors (financial harm)

Among his other large investments, William Marking owns Canyon Luxury Leasing, a company that leases luxury cars. He bought 1000 cars from Royal Motors, a manufacturer of high-priced automobiles with annual profits of around \$150 million. Royal Motors did not disclose the fact that the paint on all 1000 cars was not the original factory paint: all cars had been repainted from an ugly shade of green to the current colors of red and black.

One month after the purchase, an employee of Royal Motors reported this fact to a local newspaper, and the repaintings received a great deal of national publicity. Mr. Marking believes that people who lease luxury cars demand perfection, and that this negative publicity caused a sharp decrease in business. He has sued Royal Motors for compensation.

Internal documents of Royal Motors produced at the trial included a management memo: "Marking will not be happy at all when he finds out that the cars have been repainted. He will feel that we took advantage of him, but it was his responsibility to check the cars. Anyway we run almost no risk: it will not be in Marking's interest to publicize the problem, so he will not dare sue us, no matter how mad he is. The worst that will happen is that he won't buy from us again." The trial jury ordered Royal Motors to pay \$500,000 in compensatory damages.

### B.2. Glover v. General Assistance (physical harm)

Joan Glover, a five-year-old child, ingested a large number of pills of a non-prescription allergy medicine called Allerfree. The Allerfree bottle carried a label reading "Childproof Cap," but it did not meet federal regulations for the use of that label.

Joan's parents testified that they had been very careful in ensuring that all of their medications had childproof safety caps. Joan found the pills in a kitchen drawer and ingested most of the bottle. The overdose permanently weakened her respiratory system, which will make her more susceptible to breathing-related diseases such as asthma and emphysema for the rest of her life, and may reduce her life expectancy.

Joan's parents sued the manufacturer of Allerfree, the General Assistance company, a drug manufacturer with annual profits of \$150 million. Internal company documents showed that General Assistance chose to ignore federal regulations about standards for using the label "childproof cap." An internal memo presented at trial says that "this

---

<sup>142</sup> Pair A from Table 1.

stupid, unnecessary federal regulation is a waste of our money"; it acknowledges the risk that Allerfree might be punished for violating the regulation but says 'the punishments are extremely mild; basically we'd be asked to improve the safety caps in the future.' The trial jury ordered General Assistance to pay the Glovers \$500,000 in compensatory damages.

**B.3. Response Scales: Isolation**

What amount of punitive damages (if any) should Royal Motors be required to pay as punishment and to deter them and others from similar actions in the future? Compensatory damages do not count as part of the punishment. Please write the appropriate amount of punitive damages in the blank below.

\$ \_\_\_\_\_

**OR**

How severely should Royal Motors be punished because of their actions, and to deter them and others from similar actions in the future? Note that the compensatory damages that they must pay do not count as part of the punishment. Please circle the number that best expresses your judgment of the appropriate level of punishment.

None		Mild		Substantial		Severe		Extremely Severe
0	1	2	3	4	5	6	7	8

**.4. Response Scales: Comparison**

We would like you to compare this case (Glover v. General Assistance) to the legal case you saw earlier (Marking v. Royal Motors, which involved repainted cars).

**Compared to Royal Motors, how much punishment does General Assistance deserve? (please circle your answer)**

Less than Royal Motors	The same as Royal Motors	More than Royal Motors
---------------------------	-----------------------------	---------------------------

What amount of punitive damages (if any) should General Assistance be required to pay as punishment and to deter them and others from similar actions in the future? Compensatory damages do not count as part of the punishment. Please write the appropriate amount of punitive damages in the blank below.

\$ \_\_\_\_\_

**OR**

How severely should General Assistance be punished because of their actions, and to deter them and others from similar actions in the future? Note that the compensatory damages that they must pay do not count as part of the punishment. Please circle the number that best expresses your judgment of the appropriate level of punishment.

None		Mild		Substantial		Severe		Extremely Severe
0	1	2	3	4	5	6	7	8

## **Appendix C. Example Pair of Public Causes<sup>143</sup>**

### **C.1. Environmental Harm**

The next question moves outside the legal context. It involves problems of a more general nature, which concern society as a whole. In this part we are interested in your reaction as a member of society, rather than as a potential juror.

Dolphins in many breeding locations are threatened by pollution. The threat to breeding locations is expected to result in a decline of the dolphin population. A special fund is needed to clean up and protect dolphin breeding locations.

Increased funding to provide pollution free breeding locations for dolphins must be supported mostly by private contributions. Consider the possibility of making a voluntary contribution of money to this fund.

### **C.2. Harm to Humans**

The next question moves outside the legal context. It involves problems of a more general nature, which concern society as a whole. In this part we are interested in your reaction as a member of society, rather than as a potential juror.

Farm workers, who are exposed to the sun for many hours, have a higher rate of skin cancer than the general population. Frequent medical checkup can reduce the risk. Increased funds are needed to establish programs for more frequent checkups of the threatened groups.

Increased funding for these programs must be supported mostly by private contributions. Consider the possibility of making a voluntary contribution of money to this fund.

### **C.3. Response Scales: Isolation**

What amount of money (if any) would you be willing to contribute to the fund to protect dolphins? Please write your amount in the blank below.

\$ \_\_\_\_\_

**OR**

How much personal satisfaction would you expect to get from making a contribution a fund to protect dolphins? (please circle your answer)

---

<sup>143</sup> Pair D from Table 1.

no satisfaction at all 0	a little 1	less than for most other contributions 2	a moderate amount 3	a significant amount 4	a great deal 5	extreme satisfaction 6
-----------------------------------	------------------	---	------------------------------	---------------------------------	-------------------------	------------------------------

#### C.4. Response Scales: Comparison

We would like you to compare this problem (farm workers' skin cancer), to the problem of public concern that you saw earlier (protection of dolphins).

Compared to protecting dolphins, how much money would you be willing to contribute to a fund for farm workers' skin cancer? (circle your choice)

less than for  
dolphins

the same as for  
dolphins

more than for  
dolphins

What amount of money (if any) would you be willing to contribute to a fund to reduce farm workers' skin cancer? Please write your amount in the blank below.

\$ \_\_\_\_\_

**OR**

How much personal satisfaction would you expect to get from making a contribution a fund to reduce farm workers' skin cancer? (please circle your answer)

no satisfaction at all 0	a little 1	less than for most other contributions 2	a moderate amount 3	a significant amount 4	a great deal 5	extreme satisfaction 6
-----------------------------------	------------------	---	------------------------------	---------------------------------	-------------------------	------------------------------

### Appendix D. Selected Civil Penalties from Administrative Agencies

Agency	Civil/ Criminal	Type of violation	Penalty
OSHA	Civil	Violation of a serious nature	Up to 7000
	Civil	Violation of non-serious nature	Up to 7000
	Civil	Violation of posting requirement	Up to 7000
	Civil	Willful violation if non-serious or regulatory	No less than 5000
	Civil	Willful violation if serious	No less than 25000
	Civil	Repeated violations	Up to 70000
	Civil	Violation of record keeping- reporting	Up to 7000
	Civil	Fail to make requested forms available	Up to 3000
FDA	Civil	Egg products inspection act- violation	Up to 1000
	Civil	Above + intent to defraud or distribute adulterated eggs	Up to 10000
	Civil	Federal Meat Inspection act violation	Up to 1000
	Civil	Above + intent to defraud or distribute adulterated meat	Up to 10000
	Civil	Violation of sanitary food transportation act	Same as hazardous material transport act
	Criminal	Violation of sanitary food transportation act	Same as hazardous material transport act
	Civil	Violation of Filled milk act	Up to 1000
	Criminal	Violation of Filled milk act	Up to 1yr prison
EPA	Civil	Administrative penalty for CERCLA and EPCRA- class 1	Up to 25000 a day
	Civil	“ “ “ class 2	Up to 75000 a day
	Civil	Violation of right to know reporting under EPCRA	Up to 10000 a day
	Civil	Frivolous trade secret claim under EPCRA	Up to 25000
	Civil	Administrative violation of CWA- class 1	Up to 10000 a day 25000 max
	Civil	Administrative violation of CWA- class II	Up to 10000 a day 25000 max
	Civil	Hazardous substance discharge violation of CWA	Up to 25000 a day

	Civil	Above + gross negligence of willful misconduct	No less than 100000
	Civil	Violation of FIFRA if for hire applicator or dealer	Up to 1000
	Civil	Violation of FIFRA if non hire applicator	Up to 500 1 <sup>st</sup> offense then up to 1000
	Civil	Knowingly violate FIFRA if registrant or producer	Max 50,000
	Civil	Knowingly violate FIFRA if other	Max 25,000
	Civil	Knowingly violate FIFRA if private applicator	Max 1000
	Civil	Violation of RCRA	Up to 25000 a day
	Civil	Knowingly violate financial disclosure for solid waste disposal	Up to 2500
	Criminal	See above	Up to 1 yr prison
	Civil	Violation of compliance requirement for solid waste	Up to 25000 a day
	Criminal	Knowingly transport, treats or stores haz waste w/o permit	Up to 50000 a day or 2 yr prison
	Criminal	See above (except permit) and knows that puts person in imminent danger of death or serious bodily injury	Up to 250,000 or 15 yr prison
	Criminal	See above and defendant is organization	Up to 1,000,000
	Civil	Violates requirement of solid waste disposal	Up to 25000
	Civil	Owner fails to notify or submits false info about underground storage tanks	Up to 10,000
Fish and Wildlife Service	Civil	Violation of Wild Bird Conservation Act	
	Civil	Violate WBCA 111 a1 or 2 (import in violation of ban )or permit under 112 ( exemptions for sci research and ltd breeding)	Up to 25000
	Civil	Knowingly violate and import under WBCA 11a3 ?	Up to 12000
	Civil	If non business related and violate WBCA	Up to 500
	Civil	Civil violation of African Elephant Conservation Act	Up to 5000
	Civil	Violation of Bald Eagle Protection Act	Up to 5000
	Civil	Violation of ESA if knowingly violate/in business of im/export, taking, possessing etc	Up to 25000

	Civil	Violation of ESA if knowingly violate or ex/import and not enumerated above	Up to 12000
	Civil	Otherwise violate ESA	Up to 500
EEOC	Civil	Give public info about unlawful employment practice under T7 during proceeding	Up to 1000
	Criminal	See above	Up to 1 yr prison
	Civil	Intentional engagement in unlawful employment practice under T7	Reinstate, hire, backpay or other approp equit relief
	Civil	Willful failure to post pertinent provisions of T7	Up to 100
	Civil	Same remedies available as under T7	
	Civil	Discriminatory practice with malice	Punitive damages
	Civil	Total damages under t7 for small business (under 100 employees)	Up to 50,000
	Civil	Total damages under T7 for medium business (over 100 less than 200 employees)	Up to 100,000
	Civil	Total damages under T7 for large business (Over 200 less than 500)	Up to 200,000
	Civil	Total damages under T7 for big business (more than 500)	Up to 300,000
	Civil	Willful violation of section 15 of equal pay act	Up to 10,000
	Criminal	See above	Up to 6 mo prison
	Civil	Violation of section 6 or 7 of equal pay act	Liable for wages and liquidated damages
	Civil	Violation of section 212 (child labor) of equal pay act	Up to 10,000
	Civil	Repeated or willful violation of section 6 or 7 of EPA	Up to 1000

### Appendix E. Cost Per Life Saved

<b>Cost-Effectiveness of Selected Regulations</b>		
Regulation	Agency	Cost per premature death averted (\$ millions 1990)
Unvented Space Heater Ban	CPSC	0.1
Aircraft Cabin Fire Protection Standard	FAA	0.1
Auto Passive Restraint/Seat Belt Standards	NHTSA	0.1
Steering Column Protection Standard	NHTSA	0.1
Underground Construction Standards	OSHA-S	0.1
Trihalomethane Drinking Water Standards	EPA	0.2
Aircraft Seat Cushion Flammability Standard	FAA	0.4
Alcohol and Drug Control Standards	FRA	0.4
Auto Fuel-System Integrity Standard	NHTSA	0.4
Standards for Servicing Auto Wheel Rims	OSHA-S	0.4
Aircraft Floor Emergency Lighting Standard	FAA	0.6
Concrete & Masonry Construction Standards	OSHA-S	0.6
Crane Suspended Personnel Platform Standard	OSHA-S	0.7
Passive Restraints for Trucks & Buses (Proposed)	NHTSA	0.7
Side-Impact Standards for Autos (Dynamic)	NHTSA	0.8
Children's Sleepwear Flammability Ban	CPSC	0.8
Auto Side Door Support Standards	NHTSA	0.8
Low Altitude Windshear Equipment & Training Standards	FAA	1.3
Electrical Equipment Standards (Metal Mines)	MSHA	1.4
Trenching and Excavation Standards	OSHA-S	1.5
Traffic Alert and Collision Avoidance (TCAS) Systems	FAA	1.5
Hazard Communication Standard	OSHA-S	1.6
Side-Impact Standards for Trucks, Buses, and MPVs (Proposed)	NHTSA	2.2
Grain Dust Explosion Prevention Standards	OSHA-S	2.8
Rear Lap/Shoulder Belts for Autos	NHTSA	3.2
Standards for Radionuclides in Uranium Mines	EPA	3.4
Benzene NESHAP (Original: Fugitive Emissions)	EPA	3.4
Ethylene Dibromide Drinking Water Standard	EPA	5.7
Benzene NESHAP (Revised: Coke Byproducts)	EPA	6.1

Asbestos Occupational Exposure Limit	OSHA-H	8.3
Benzene Occupational Exposure Limit	OSHA-H	8.9
Electrical Equipment Standards (Coal Mines)	MSHA	9.2
Arsenic Emission Standards for Glass Plants	EPA	13.5
Ethylene Oxide Occupational Exposure Limit	OSHA-H	20.5
Arsenic/Copper NESHP	EPA	23.0
Hazardous Waste Listing for Petroleum Refining Sludge	EPA	27.6
Cover/Move Uranium Mill Tailings (Inactive Sites)	EPA	31.7
Benzene NESHP (Revised: Transfer Operations)	EPA	32.9
Cover/Move Uranium Mill Tailings (Active Sites)	EPA	45.0
Acrylonitrile Occupational Exposure Limit	OSHA-H	51.5
Coke Ovens Occupational Exposure Limit	OSHA-H	63.5
Lockout/Tagout	OSHA-S	70.9
Asbestos Occupational Exposure Limit	OSHA-H	74.0
Arsenic Occupational Exposure Limit	OSHA-H	106.9
Asbestos Ban	EPA	110.7
Diethylstilbestrol (DES) Cattlefeed Ban	FDA	124.8
Benzene NESHP (Revised: Waste Operations)	EPA	168.2
1,2 Dichloropropane Drinking Water Standard	EPA	653.0
Hazardous Waste Land Disposal Ban (1st 3rd)	EPA	4,190.4
Municipal Solid Waste Landfill Standards (Proposed)	EPA	19,107.0
Formaldehyde Occupational Exposure Limit	OSHA-H	86,201.8
Atrazine/Alachlor Drinking Water Standard	EPA	92,069.7
Hazardous Waste Listing for Wood-Preserving Chemicals	EPA	5,700,000

Readers with comments should address them to:

Professor Cass R. Sunstein  
 Karl N. Llewellyn Distinguished Service Professor of Jurisprudence  
 University of Chicago Law School  
 1111 East 60th Street  
 Chicago, IL 60637  
 773.702.9498  
 csunstei@midway.uchicago.edu

Chicago Working Papers in Law and Economics  
(Second Series)

1. William M. Landes, Copyright Protection of Letters, Diaries and Other Unpublished Works: An Economic Approach (July 1991).
2. Richard A. Epstein, The Path to *The T. J. Hooper*: The Theory and History of Custom in the Law of Tort (August 1991).
3. Cass R. Sunstein, On Property and Constitutionalism (September 1991).
4. Richard A. Posner, Blackmail, Privacy, and Freedom of Contract (February 1992).
5. Randal C. Picker, Security Interests, Misbehavior, and Common Pools (February 1992).
6. Tomas J. Philipson & Richard A. Posner, Optimal Regulation of AIDS (April 1992).
7. Douglas G. Baird, Revisiting Auctions in Chapter 11 (April 1992).
8. William M. Landes, Sequential versus Unitary Trials: An Economic Analysis (July 1992).
9. William M. Landes & Richard A. Posner, The Influence of Economics on Law: A Quantitative Study (August 1992).
10. Alan O. Sykes, The Welfare Economics of Immigration Law: A Theoretical Survey With An Analysis of U.S. Policy (September 1992).
11. Douglas G. Baird, 1992 Katz Lecture: Reconstructing Contracts (November 1992).
12. Gary S. Becker, The Economic Way of Looking at Life (January 1993).
13. J. Mark Ramseyer, Credibly Committing to Efficiency Wages: Cotton Spinning Cartels in Imperial Japan (March 1993).
14. Cass R. Sunstein, Endogenous Preferences, Environmental Law (April 1993).
15. Richard A. Posner, What Do Judges and Justices Maximize? (The Same Thing Everyone Else Does) (April 1993).
16. Lucian Arye Bebchuk and Randal C. Picker, Bankruptcy Rules, Managerial Entrenchment, and Firm-Specific Human Capital (August 1993).
17. J. Mark Ramseyer, Explicit Reasons for Implicit Contracts: The Legal Logic to the Japanese Main Bank System (August 1993).
18. William M. Landes and Richard A. Posner, The Economics of Anticipatory Adjudication (September 1993).
19. Kenneth W. Dam, The Economic Underpinnings of Patent Law (September 1993).
20. Alan O. Sykes, An Introduction to Regression Analysis (October 1993).
21. Richard A. Epstein, The Ubiquity of the Benefit Principle (March 1994).
22. Randal C. Picker, An Introduction to Game Theory and the Law (June 1994).
23. William M. Landes, Counterclaims: An Economic Analysis (June 1994).
24. J. Mark Ramseyer, The Market for Children: Evidence from Early Modern Japan (August 1994).
25. Robert H. Gertner and Geoffrey P. Miller, Settlement Escrows (August 1994).
26. Kenneth W. Dam, Some Economic Considerations in the Intellectual Property Protection of Software (August 1994).
27. Cass R. Sunstein, Rules and Rulelessness, (October 1994).

28. David Friedman, More Justice for Less Money: A Step Beyond *Cimino* (December 1994).
29. Daniel Shaviro, Budget Deficits and the Intergenerational Distribution of Lifetime Consumption (January 1995).
30. Douglas G. Baird, The Law and Economics of Contract Damages (February 1995).
31. Daniel Kessler, Thomas Meites, and Geoffrey P. Miller, Explaining Deviations from the Fifty Percent Rule: A Multimodal Approach to the Selection of Cases for Litigation (March 1995).
32. Geoffrey P. Miller, Das Kapital: Solvency Regulation of the American Business Enterprise (April 1995).
33. Richard Craswell, Freedom of Contract (August 1995).
34. J. Mark Ramseyer, Public Choice (November 1995).
35. Kenneth W. Dam, Intellectual Property in an Age of Software and Biotechnology (November 1995).
36. Cass R. Sunstein, Social Norms and Social Roles (January 1996).
37. J. Mark Ramseyer and Eric B. Rasmusen, Judicial Independence in Civil Law Regimes: Econometrics from Japan (January 1996).
38. Richard A. Epstein, Transaction Costs and Property Rights: Or Do Good Fences Make Good Neighbors? (March 1996).
39. Cass R. Sunstein, The Cost-Benefit State (May 1996).
40. William M. Landes and Richard A. Posner, The Economics of Legal Disputes Over the Ownership of Works of Art and Other Collectibles (July 1996).
41. John R. Lott, Jr. and David B. Mustard, Crime, Deterrence, and Right-to-Carry Concealed Handguns (August 1996).
42. Cass R. Sunstein, Health-Health Tradeoffs (September 1996).
43. G. Baird, The Hidden Virtues of Chapter 11: An Overview of the Law and Economics of Financially Distressed Firms (March 1997).
44. Richard A. Posner, Community, Wealth, and Equality (March 1997).
45. William M. Landes, The Art of Law and Economics: An Autobiographical Essay (March 1997).
46. Cass R. Sunstein, Behavioral Analysis of Law (April 1997).
47. John R. Lott, Jr. and Kermit Daniel, Term Limits and Electoral Competitiveness: Evidence from California's State Legislative Races (May 1997).
48. Randal C. Picker, Simple Games in a Complex World: A Generative Approach to the Adoption of Norms (June 1997).
49. Richard A. Epstein, Contracts Small and Contracts Large: Contract Law through the Lens of Laissez-Faire (August 1997).
50. Cass R. Sunstein, Daniel Kahneman, and David Schkade, Assessing Punitive Damages (with Notes on Cognition and Valuation in Law) (December 1997).
51. William M. Landes, Lawrence Lessig, and Michael E. Solimine, Judicial Influence: A Citation Analysis of Federal Courts of Appeals Judges (January 1998).
52. John R. Lott, Jr., A Simple Explanation for Why Campaign Expenditures are Increasing: The Government is Getting Bigger (February 1998).

53. Richard A. Posner, *Values and Consequences: An Introduction to Economic Analysis of Law* (March 1998).
54. Denise DiPasquale and Edward L. Glaeser, *Incentives and Social Capital: Are Homeowners Better Citizens?* (April 1998).
55. Christine Jolls, Cass R. Sunstein, and Richard Thaler, *A Behavioral Approach to Law and Economics* (May 1998).
56. John R. Lott, Jr., *Does a Helping Hand Put Others At Risk?: Affirmative Action, Police Departments, and Crime* (May 1998).
57. Cass R. Sunstein and Edna Ullmann-Margalit, *Second-Order Decisions* (June 1998).
58. Jonathan M. Karpoff and John R. Lott, Jr., *Punitive Damages: Their Determinants, Effects on Firm Value, and the Impact of Supreme Court and Congressional Attempts to Limit Awards* (July 1998).
59. Kenneth W. Dam, *Self-Help in the Digital Jungle* (August 1998).
60. John R. Lott, Jr., *How Dramatically Did Women's Suffrage Change the Size and Scope of Government?* (September 1998)
61. Kevin A. Kordana and Eric A. Posner, *A Positive Theory of Chapter 11* (October 1998)
62. David A. Weisbach, *Line Drawing, Doctrine, and Efficiency in the Tax Law* (November 1998)
63. Jack L. Goldsmith and Eric A. Posner, *A Theory of Customary International Law* (November 1998)
64. John R. Lott, Jr., *Public Schooling, Indoctrination, and Totalitarianism* (December 1998)
65. Cass R. Sunstein, *Private Broadcasters and the Public Interest: Notes Toward A "Third Way"* (January 1999)
66. Richard A. Posner, *An Economic Approach to the Law of Evidence* (February 1999)
67. Yannis Bakos, Erik Brynjolfsson, Douglas Lichtman, *Shared Information Goods* (February 1999)
68. Kenneth W. Dam, *Intellectual Property and the Academic Enterprise* (February 1999)
69. Gertrud M. Fremling and Richard A. Posner, *Status Signaling and the Law, with Particular Application to Sexual Harassment* (March 1999)
70. Cass R. Sunstein, *Must Formalism Be Defended Empirically?* (March 1999)
71. Jonathan M. Karpoff, John R. Lott, Jr., and Graeme Rankine, *Environmental Violations, Legal Penalties, and Reputation Costs* (March 1999)
72. Matthew D. Adler and Eric A. Posner, *Rethinking Cost-Benefit Analysis* (April 1999)
73. John R. Lott, Jr. and William M. Landes, *Multiple Victim Public Shooting, Bombings, and Right-to-Carry Concealed Handgun Laws: Contrasting Private and Public Law Enforcement* (April 1999)

74. Lisa Bernstein, The Questionable Empirical Basis of Article 2's Incorporation Strategy: A Preliminary Study (May 1999)
75. Richard A. Epstein, Deconstructing Privacy: and Putting It Back Together Again (May 1999)
76. William M. Landes, Winning the Art Lottery: The Economic Returns to the Ganz Collection (May 1999)
77. Cass R. Sunstein, David Schkade, and Daniel Kahneman, Do People Want Optimal Deterrence? (June 1999)
78. Tomas J. Philipson and Richard A. Posner, The Long-Run Growth in Obesity as a Function of Technological Change (June 1999)
79. David A. Weisbach, Ironing Out the Flat Tax (August 1999)
80. Eric A. Posner, A Theory of Contract Law under Conditions of Radical Judicial Error (August 1999)
81. David Schkade, Cass R. Sunstein, and Daniel Kahneman, Are Juries Less Erratic than Individuals? Deliberation, Polarization, and Punitive Damages (September 1999)
82. Cass R. Sunstein, Nondelegation Canons (September 1999)
83. Richard A. Posner, The Theory and Practice of Citations Analysis, with Special Reference to Law and Economics (September 1999)
84. Randal C. Picker, Regulating Network Industries: A Look at *Intel* (October 1999)
85. Cass R. Sunstein, Cognition and Cost-Benefit Analysis (October 1999)
86. Douglas G. Baird and Edward R. Morrison, Optimal Timing and Legal Decisionmaking: The Case of the Liquidation Decision in Bankruptcy (October 1999)
87. Gertrud M. Fremling and Richard A. Posner, Market Signaling of Personal Characteristics (November 1999)
88. Matthew D. Adler and Eric A. Posner, Implementing Cost-Benefit Analysis When Preferences Are Distorted (November 1999)
89. Richard A. Posner, Orwell versus Huxley: Economics, Technology, Privacy, and Satire (November 1999)
90. David A. Weisbach, Should the Tax Law Require Current Accrual of Interest on Derivative Financial Instruments? (December 1999)
91. Cass R. Sunstein, The Law of Group Polarization (December 1999)
92. Eric A. Posner, Agency Models in Law and Economics (January 2000)
93. Karen Eggleston, Eric A. Posner, and Richard Zeckhauser, Simplicity and Complexity in Contracts (January 2000)
94. Douglas G. Baird and Robert K. Rasmussen, Boyd's Legacy and Blackstone's Ghost (February 2000)
95. David Schkade, Cass R. Sunstein, Daniel Kahneman, Deliberating about Dollars: The Severity Shift (February 2000)
96. Richard A. Posner and Eric B. Rasmusen, Creating and Enforcing Norms, with Special Reference to Sanctions (March 2000)

97. Douglas Lichtman, Property Rights in Emerging Platform Technologies (April 2000)
98. Cass R. Sunstein and Edna Ullmann-Margalit, Solidarity in Consumption (May 2000)
99. David A. Weisbach, An Economic Analysis of Anti-Tax Avoidance Laws (May 2000)
100. Cass R. Sunstein, Human Behavior and the Law of Work (June 2000)
101. William M. Landes and Richard A. Posner, Harmless Error (June 2000)
102. Robert H. Frank and Cass R. Sunstein, Cost-Benefit Analysis and Relative Position (August 2000)
103. Eric A. Posner, Law and the Emotions (September 2000)
104. Cass R. Sunstein, Cost-Benefit Default Principles (October 2000)
105. Jack Goldsmith and Alan Sykes, The Dormant Commerce Clause and the Internet (November 2000)
106. Richard A. Posner, Antitrust in the New Economy (November 2000)
107. Douglas Lichtman, Scott Baker, and Kate Kraus, Strategic Disclosure in the Patent System (November 2000)
108. Jack L. Goldsmith and Eric A. Posner, Moral and Legal Rhetoric in International Relations: A Rational Choice Perspective (November 2000)
109. William Meadow and Cass R. Sunstein, Statistics, Not Experts (December 2000)
110. Saul Levmore, Conjunction and Aggregation (December 2000)
111. Saul Levmore, Puzzling Stock Options and Compensation Norms (December 2000)
112. Richard A. Epstein and Alan O. Sykes, The Assault on Managed Care: Vicarious Liability, Class Actions and the Patient's Bill of Rights (December 2000)
113. William M. Landes, Copyright, Borrowed Images and Appropriation Art: An Economic Approach (December 2000)
114. Cass R. Sunstein, Switching the Default Rule (January 2001)
115. George G. Triantis, Financial Contract Design in the World of Venture Capital (January 2001)
116. Jack Goldsmith, Statutory Foreign Affairs Preemption (February 2001)
117. Richard Hynes and Eric A. Posner, The Law and Economics of Consumer Finance (February 2001)
118. Cass R. Sunstein, Academic Fads and Fashions (with Special Reference to Law) (March 2001)
119. Eric A. Posner, Controlling Agencies with Cost-Benefit Analysis: A Positive Political Theory Perspective (April 2001)
120. Douglas G. Baird, Does Bogart Still Get Scale? Rights of Publicity in the Digital Age (April 2001)
121. Douglas G. Baird and Robert K. Rasmussen, Control Rights, Priority Rights and the Conceptual Foundations of Corporate Reorganization (April 2001)
122. David A. Weisbach, Ten Truths about Tax Shelters (May 2001)

123. William M. Landes, What Has the Visual Arts Rights Act of 1990 Accomplished? (May 2001)
124. Cass R. Sunstein, Social and Economic Rights? Lessons from South Africa (May 2001)
125. Christopher Avery, Christine Jolls, Richard A. Posner, and Alvin E. Roth, The Market for Federal Judicial Law Clerks (June 2001)
126. Douglas G. Baird and Edward R. Morrison, Bankruptcy Decision Making (June 2001)
127. Cass R. Sunstein, Regulating Risks after ATA (June 2001)
128. Cass R. Sunstein, The Laws of Fear (June 2001)
129. Richard A. Epstein, In and Out of Public Solution: The Hidden Perils of Property Transfer (July 2001)
130. Randal C. Picker, Pursuing a Remedy in *Microsoft*: The Declining Need for Centralized Coordination in a Networked World (July 2001)
131. Cass R. Sunstein, Daniel Kahneman, David Schkade, and Ilana Ritov, Predictably Incoherent Judgments (July 2001)