2008

Some Effects of Moral Indignation on Law

Cass R. Sunstein

SOME EFFECTS OF MORAL INDIGNATION ON LAW

Cass R. Sunstein[*†]

ABSTRACT

Moral intuitions operate in much the same way as other intuitions do; what makes the moral domain distinctive is the frequent foundation of moral judgments in the emotions, beliefs, and response tendencies that define indignation. The intuitive system of cognition, System 1, is typically responsible for indignation; the more reflective system, System 2, may or may not provide an override. An understanding of indignation helps to explain many phenomena of interest to law and politics: the outrage heuristic, the severity shift, the puzzling centrality of harm, moral framing, and the act–omission distinction. The operation of System 1 also helps to explain moral dumbfounding, understood as intense moral opprobrium that people are unable to justify, and moral numbness, understood as moral indifference that people know on reflection to be unwarranted. Both moral dumbfounding and moral numbness play a significant role in law and politics. Because of the nature of indignation, it is extremely difficult for people to achieve coherence in their moral intuitions, and the absence of coherence appears to be replicated in several areas of law. Legal and political institutions usually aspire to be deliberative, to check intuitions that misfire, and to pay close attention to System 2; but even in deliberative institutions, System 1 can make some compelling demands. A general implication is that judges may not be aware of the actual causes of their moral judgments and of the legal conclusions that rely on them.

INTRODUCTION

The psychological analysis of moral sentiments has witnessed profound changes over the last few decades, from a conception of morality as a system of abstract rules that can be understood and internalized[1] to a view that emphasizes moral emotions and moral intuitions that are not

* Felix Frankfurter Professor, Harvard Law School. The *Vermont Law Review* and the author note that because Professor Sunstein has joined the Office of Management and Budget in the Obama Administration, the author was unavailable to approve editorial changes customary to publication. It was therefore necessary to print the Article with minimal revision. This Article draws on and develops some of the analysis in Daniel Kahneman & Cass R. Sunstein, *Cognitive Psychology of Moral Intuitions*, *in* NEUROBIOLOGY OF HUMAN VALUES 91 (Jean-Pierre Changeux et al. eds., 2005), and Cass R. Sunstein, *Moral Heuristics and Moral Framing*, 88 MINN. L.R. 1556 (2004).

† I am generally grateful to Daniel Kahneman for many years of discussion; he deserves credit for whatever is valuable in this piece but I am solely to blame for confusion and errors.

1. *See* Lawrence Kohlberg, *Stage and Sequence: The Cognitive-Developmental Approach to Socialization*, *in* HANDBOOK OF SOCIALIZATION THEORY AND RESEARCH 347 (D.A. Goslin ed., 1969).

anchored in reasons.[2] On this view, reasons tend to be ex post rationalizations for moral intuitions, rather than causal.[3] My goals here are to sketch an analysis of moral intuitions that builds on the new work, to relate that analysis to a general approach to the study of intuitive thought, and to connect that approach to a set of issues in politics and law.[4]

The central analysis applies to a wide range of moral intuitions, but the emphasis throughout is on the complex of emotions, beliefs, and response tendencies that define *indignation*.[5] As we shall see, indignation is responsible for a number of puzzling practices in politics and law. Recurring themes are that people's moral judgments are often automatic, that their automatic responses play a significant role in both legislatures and in courtrooms, and that it is often valuable but difficult to attempt to constrain automatic responses by reference to more deliberative processes. A general conclusion is that people may not be aware of the actual causes of their indignation and of the legal outcomes that rest on those causes.

It has been suggested that indignation comes in three distinctive varieties: anger, disgust, and contempt.[6] The main concern here is the variant of indignation that involves anger. For a mundane example, imagine

---

2. Joshua D. Greene, *The Secret Joke of Kant's Soul*, *in* MORAL PSYCHOLOGY: THE NEUROSCIENCE OF MORALITY: EMOTION, BRAIN DISORDERS, AND DEVELOPMENT 35, 36 (Walter Sinnott-Armstrong ed., 2008) [hereinafter Greene, *The Secret Joke*]; Joshua D. Greene & J.D. Cohen, *For the Law, Neuroscience Changes Nothing and Everything*, 359 PHIL. TRANSACTIONS ROYAL SOC'Y LONDON B 1775, 1775 (2004); Joshua D. Greene & Jonathan Haidt, *How (And Where) Does Moral Judgment Work?*, 6 TRENDS COGNITIVE SCI. 517, 517 (2002); Jonathan Haidt, *The New Synthesis in Moral Psychology*, 316 SCIENCE 998, 998 (2007) [hereinafter Haidt, *The New Synthesis*]; Jonathan Haidt, *The Emotional Dog and Its Rational Tail: A Social Intuitionist Approach to Moral Judgment*, 108 PSYCHOL. REV. 814, 814 (2001) [hereinafter Haidt, *The Emotional Dog*]; Paul Rozin et al., *The CAD Triad Hypothesis: A Mapping Between Three Moral Emotions (Contempt, Anger, Disgust) and Three Moral Codes (Community, Autonomy, and Divinity)*, 76 J. PERS. & SOC. PSYCHOL. 574, 574 (1999). For a useful synthesis, distinguishing between reactions to moral transgressions and reactions to moral dilemmas, see Benoît Monin et al., *Deciding Versus Reacting: Conceptions of Moral Judgment and the Reason-Affect Debate*, 11 REV. GEN. PSYCHOL. 99 (2007).

3. *See, e.g.*, Marc Hauser et al., *A Dissociation Between Moral Judgments and Justifications*, 22 MIND & LANGUAGE 1, 17 (2007).

4. An illuminating call for an integration of psychology and philosophy can be found in KWAME ANTHONY APPIAH, EXPERIMENTS IN ETHICS 123 (2008). An exploration of related issues can be found in THE AFFECT EFFECT: DYNAMICS OF EMOTION IN POLITICAL THINKING AND BEHAVIOR (W. Russell Neuman et al. eds., 2007).

5. Indignation is not, of course, the only such complex relevant to moral psychology. See, for example, the discussion of disgust in Dan Jones, *Moral Psychology: The Depths of Disgust*, 447 NATURE 768 (2007). Much of what I shall have to say about System 1 and System 2 bears on other pertinent complexes of emotion, beliefs, and response tendencies, including disgust, shame, and hatred. In my view, however, indignation has a kind of primacy, certainly in the domains of policy and law. For an interesting study of the different aspects of the brain that are triggered by disgust and indignation, see Jorge Moll et al., *The Moral Affiliations of Disgust: A Functional MRI Study*, 18 COGNITIVE & BEHAV. NEUROLOGY 68 (2005).

6. Rozin et al., *supra* note 2, at 575.

that you see a bully beat up a weakling without any provocation. You will respond with indignation. Like other intentional states, indignation can be explained in two quite different ways: by referring to reasons, or by invoking psychological causes. As you see the bully assaulting his victim, you are likely to be aware of a reason for your emotion: the action violates an accepted (and in your view justified) social rule that prohibits unprovoked aggression. The categorization of the action provides a reason for indignation, a reason that the observer expects other objective observers to endorse. Classical analyses of moral development were much concerned with people's ability to marshal reasons for their judgments; the reasons were often understood as causing those judgments.[7]

The view that has gained currency in recent years is quite different.[8] In this view, indignation is like a fear of spiders. One does not fear spiders because they are dangerous—one just fears them. Because people tend to attribute their reactions to the objects that evoke these reactions, the feared spider is perceived as a dangerous spider. However, the perception of dangerousness is not the reason for the fear or even its cause; both the fear and the perception are symptoms of an uncontrolled reaction to spiders. Many people who are afraid of spiders know that their fear is objectively groundless and lacks a reason. The equivalent state in the moral domain has been described as "moral dumbfounding": the experience of intense moral reactions, sometimes producing political or legal action, for which no adequate reason can be brought to mind.[9]

Indignation is sometimes not caused by reasons,[10] and people can be dumbfounded when they are asked to explain why they are indignant. For some moral problems, people are unaware of the principles that are motivating their judgments, and they might not endorse those principles on reflection.[11] In fact, some puzzling outcomes, in both politics and law, are a product of indignation that is simultaneously intense and hard to justify.[12] Often the legal system tracks moral judgments for which reasons are hard to give; some political and legal disputes are a direct result.[13] In constitutional

7. *See* Kohlberg, *supra* note 1.

8. *See, e.g.*, Haidt, *The Emotional Dog*, *supra* note 2, at 814.

9. *Id.* at 817.

10. *See* Hauser, *supra* note 3, at 16.

11. *See* Fiery Cushman et al., *The Role of Conscious Reasoning and Intuition in Moral Judgment: Testing Three Principles of Harm*, 17 PSYCHOL. SCI. 1082, 1086–87 (2006) (finding that people appear responsive to a principle that distinguishes between intended harm and foreseen harm, but that they do not appeal to that principle in their own justifications).

12. *See* BRYAN CAPLAN, THE MYTH OF THE RATIONAL VOTER: WHY DEMOCRACIES CHOOSE BAD POLICIES 1 (2007); Cass R. Sunstein, *Moral Heuristics*, 28 BEHAV. & BRAIN SCI. 531, 536 (2005).

13. *See generally* Lawrence v. Texas, 539 U.S. 558 (2003); LEON KASS, *The Wisdom of Repugnance*, *in* THE ETHICS OF HUMAN CLONING 3 (Leon R. Kass & James Q. Wilson eds., 2001).

law, rationality review might be understood as a response to the risk of moral dumbfounding.[14]

Moral dumbfounding finds its mirror image in moral numbness, in which people are not indignant even though they have reason to be, and know they do. Consider, for example, the generally tepid reaction to natural disasters or even genocide in a distant nation,[15] contrasting with intense responses to incidents involving a single identifiable victim.[16] It is difficult to produce widespread indignation in the face of large numbers of foreign deaths. An important question is whether effective responses might nonetheless be motivated through vivid accounts, triggering indignation and hence action after all, or through legal institutions, producing action even when indignation is absent.[17]

The remainder of this Article is divided into four parts. Part I explores the central psychological points. It sketches two families of operations within the human mind; the first is rapid and automatic whereas the second is slower and more deliberative. It suggests that the automatic system plays a role in discrimination on the basis of race and sex, in judgments about risks, and in assessing fairness. Part II explores the outrage heuristic and its relationship to punishment judgments and to risk regulation. A special point here is a kind of *rhetorical asymmetry* that applies within deliberating groups, heightening outrage among group members. Part III turns to moral framing and its effects on jury behavior and the monetary valuation of human life. Part IV investigates the role of moral intuitions in distinguishing between acts and omissions, in giving special attention to "identifiable victims," and in imposing direct and indirect harm.

I. THE TWO-SYSTEM MODEL OF THE MIND

Consider the expression "17 x 24 = ?". For the great majority of people, the correct answer to the question will come to mind only if it is produced by a voluntary mental activity, which involves deliberate application of a rule, requires several steps of computation, storage, and retrieval, and takes a significant amount of time. For contrast, consider the

14. *See Lawrence*, 539 U.S. at 578 (2003); City of Cleburne v. Cleburne Living Ctr., 473 U.S. 432, 448–50 (1985).

15. *See* Paul Slovic, *"If I Look At the Mass I Will Never Act": Psychic Numbing and Genocide*, 2 JUDGMENT & DECISION MAKING 79, 82–90 (2007), http://journal.sjdm.org/jdm7303a.pdf (presenting a psychological model for understanding public apathy to genocide).

16. Karen Jenni & George Loewenstein, *Explaining the Identifiable Victim Effect*, 14 J. RISK & UNCERTAINTY 235, 236, 253 (1997).

17. Slovic, *supra* note 15, at 91.

word "vomit." For the great majority of people, disgust will come to mind in a completely involuntary process, which is produced very quickly by a process which is itself unconscious—one is aware only of its outcome. The two examples represent different families of cognitive processes.

A. Intuition and Reflection

The ancient idea that cognitive processes can be partitioned into two main families—traditionally called "intuition" and "reason"—is now widely embraced under the general label of dual-process theories.[18] Dual-process theories come in many forms, but all distinguish cognitive operations[19] that are quick and associative from others that are slower, more reflective, and frequently more calculative.[20] The generic labels "System 1" and "System 2" are adopted from Keith Stanovich and Richard West.[21] These terms may suggest the image of autonomous homunculi, and there is growing evidence that the two systems correspond to different locations in the brain,[22] but I do not suggest that the two systems are independent.[23] The

---

18. *See, e.g.*, Shelly Chaiken & Yaacov Trope, *Preface* to DUAL-PROCESS THEORIES IN SOCIAL PSYCHOLOGY, at ix (Shelly Chaiken & Yaacov Trope eds., 1999); Steven A. Sloman, *The Empirical Case for Two Systems of Reasoning*, 119 PSYCHOL. BULL. 3, 3 (1996) (discussing dual reasoning as composed of one associative system and one "rule based" system). For a discussion and partial criticism, see Michael L. Spezio & Ralph Adolphs, *Emotional Processing and Political Judgment: Toward Integrating Political Psychology and Decision Neuroscience*, *in* THE AFFECT EFFECT, *supra* note 4, at 71, 73–76.

19. Thus we do not mean to venture any controversial view about the relationship between emotions and cognition. For varying views, see JON ELSTER, ALCHEMIES OF THE MIND 239–331 (1999); MARTHA NUSSBAUM, UPHEAVALS OF THOUGHT (2001).

20. *See* Daniel T. Gilbert, *What the Mind's Not*, *in* DUAL-PROCESS THEORIES IN SOCIAL PSYCHOLOGY, *supra* note 18, at 3; Elizabeth A. Phelps, *The Interaction of Emotion and Cognition: The Relation Between the Human Amygdala and Cognitive Awareness*, *in* THE NEW UNCONSCIOUS 61 (Ran R. Hassin, James S. Uleman & John A. Bargh eds., 2005).

21. Keith E. Stanovich & Richard F. West, *Individual Differences in Reasoning: Implications for the Rationality Debate?*, *in* HEURISTICS AND BIASES: THE PSYCHOLOGY OF INTUITIVE JUDGMENT 421 (Thomas Gilovich et al. eds., 2002) [hereinafter Stanovich & West, *Individual Differences in Reasoning*].

22. JOSEPH LEDOUX, THE EMOTIONAL BRAIN 163–65 (2003); William D. Casebeer, *Moral Cognition and Its Neural Constituents*, 4 NATURE REVS. NEUROSCIENCE 841, 845 (2003); Joshua D. Greene et al., *The Neural Bases of Cognitive Conflict and Control in Moral Judgment*, 44 NEURON 389, 398 (2004); Carla L. Harenski & Stephan Harmann, *Neural Correlates of Regulating Negative Emotions Relating to Moral Violations*, 30 NEUROIMAGE 313, 317 (2006); Jorge Moll et al., *Morals and the Human Brain: A Working Model*, 14 NEUROREPORT 299, 301–02 (2003); Greene, *The Secret Joke*, *supra* note 2, at 40–41. *See generally* Jana Schaich Borg et al., *Consequences, Action, and Intention as Factors in Moral Judgments: An fMRI Investigation*, 18 J. COGNITIVE NEUROSCIENCE 803, 805 (2006); Michael Koenigs et al., *Damage to the Prefrontal Cortex Increases Utilitarian Moral Judgements*, 446 NATURE 908, 908 (2007).

23. On their mutual dependence, see the different perspectives in D.A. Pizarro & Paul Bloom, *The Intelligence of Moral Intuitions: Comment on Haidt*, 110 PSYCHOL. REV. 193, 194 (2003) (emphasizing the role of reflection in checking intuitions), and Jonathan Haidt, *The Emotional Dog Does*

term "systems" is used as a label for collections of processes that are distinguished by their speed, their controllability, and the contents on which they operate.[24]

System 1 (Intuitive)	**System 2 (Reflective)**
Automatic	Controlled
Effortless	Effortful
Associative	Deductive
Rapid	Slow
Opaque process	Self-aware
Skilled	Rule-following

Table 1
Two Cognitive Systems

Although System 1 is more primitive than System 2, it is not always or necessarily less capable. On the contrary, complex cognitive operations eventually migrate from System 2 to System 1 as proficiency and skill are acquired.[25] A striking demonstration of the intelligence of System 1 is the ability of professional tennis players to know what shot to hit instantly, and in that sense intuitively. For those experts, pattern matching has replaced effortful serial processing. For those who study law as well, there is evident movement, over time, from the controlled and effortful to the rapid and intuitive. Sometimes the movements occur within people, permitting intuition to replace effort; sometimes social changes occur over time, so that for most people, System 1 develops rapid judgments, moral, legal, or otherwise, that differ radically from the intuitive judgments of mere decades before.[26] In the areas of both law and morality, consider the trajectory of the practice of sexual harassment, which often produces intuitive indignation today, but did so far less often in, say, 1960. It is also

---

Learn New Tricks: A Reply to Pizarro and Bloom, 110 PSYCHOL. REV. 197, 197 (2003) (suggesting that deliberate efforts to oppose moral intuitions are rare).

24. There is an evident connection between System 1 and the affect heuristic, discussed in Paul Slovic et al., *The Affect Heuristic*, *in* HEURISTICS AND BIASES, *supra* note 21, at 397. The affect heuristic, as Slovic and his coauthors describe it, can be seen as System 1 in action, and it can account for moral judgments as well as judgments of other kinds. *Id.*

25. *See generally* GARY KLEIN, SOURCES OF POWER: HOW PEOPLE MAKE DECISIONS (1999).

26. See the discussion of "moralization" in Paul Rozin, *The Process of Moralization*, 10 PSYCHOL. SCI. 218, 219–20 (1999). *See also* Paul Rozin & L. Singh, *The Moralization of Cigarette Smoking in the United States*, 8 J. CONSUMER PSYCHOL. 339 (1999); Paul Rozin et al., *Moralization and Becoming a Vegetarian: The Transformation of Preferences Into Value and the Recruitment of Disgust*, 8 PSYCHOL. SCI. 67, 72 (1997).

possible that the reflective judgments produced by System 2 will be based on errors of one or another kind, and that people's intuitive revulsion is telling whatever System 2 might say; many criticisms of utilitarianism are rooted in this view.[27]

In the particular dual-process model explored here, System 1 quickly proposes intuitive answers to judgment problems as they arise, and System 2 monitors the quality of these proposals, which it may endorse, correct, or override.[28] The judgments that are eventually expressed are called "intuitive" if they retain the hypothesized initial proposal without much modification. There is an obvious relationship between this claim and the (controversial) use of intuitions in the search for reflective equilibrium in thinking about justice.[29] When people have a strong intuitive belief that some practice is immoral, that belief may well operate as a fixed point in the search for reflective equilibrium, even if it should not. In an implicit celebration of System 1, ethicist Leon Kass points to the fact that human beings "intuit and feel, immediately and without argument, the violation of things that we rightfully hold dear."[30]

In both politics and law, analogical reasoning also involves System 1 and System 2, as intuitive judgments about analogy, or disanalogy, become tested and refined through more reflective invocation of relevant similarities.[31] An evident possibility is that the legal arguments that are said to support an outcome, or a court's explanation of a judgment of perceived analogousness, are not at all causal; in adjudication as in morality, the motivation for the conclusion may not be well understood even by those

---

27. *See generally* Martha Nussbaum, Foreword, *Constitutions and Capabilities*, 121 HARV. L. REV. 4 (2007); Bernard Williams, *A Critique of Utilitarianism*, *in* UTILITARIANISM: FOR AND AGAINST (J.J.C. Smart & Bernard Williams eds., 1973). An especially vivid plea for respectful use of what we would call System 1 can be found in Leon Kass, *supra* note 13, at 19, and particularly in this suggestion: "We are repelled by the prospect of cloning human beings not because of the strangeness or novelty of the undertaking, but because we intuit and feel, immediately and without argument, the violation of things that we rightfully hold dear. . . . Shallow are the souls that have forgotten how to shudder." *Id.* If my argument here is correct, there is a possibility that even when System 1 is shuddering, System 2 should be called into action.

28. This approach is consistent with claims in Pizarro & Bloom, *supra* note 23, at 195, and the evidence reviewed in Greene & Haidt, *supra* note 2, at 522.

29. *See* JOHN RAWLS, A THEORY OF JUSTICE 42–45 (rev. ed. 1999). On the controversial character of the use of intuitions, see Greene, *The Secret Joke*, *supra* note 2, at 46–48; APPIAH, *supra* note 4, at 75.

30. *See* Kass, *supra* note 13, at 19.

31. *See* EDWARD H. LEVI, AN INTRODUCTION TO LEGAL REASONING 1–8 (1949); Scott Brewer, *Exemplary Reasoning: Semantics, Pragmatics, and the Force of Legal Argument by Analogy*, 109 HARV. L. REV. 923, 950–51 (1996). For a valuable discussion, suggesting the power of what might be called a well-educated System 1, see LLOYD WEINREB, LEGAL REASON: THE USE OF ANALOGY IN LEGAL ARGUMENT 68–77 (2005).

who reach that conclusion.[32] If System 1 plays a large role in the judgments not only of juries but of judges as well, then litigants must find a way to speak directly to its concerns.[33]

The roles of the two systems in determining ultimate judgments depend on features of the task and of the individual, including the topic,[34] the time available for deliberation,[35] the respondent's mood,[36] and intelligence.[37] Without time for deliberation, for example, indignation can be extremely intense; when people have time to reflect, their reaction sometimes diminishes.[38] And when System 1 is not indignant, and people are morally numb, deliberation can heighten moral concern and possibly produce indignation (although this can take a great deal of heavy lifting on the part of System 2). It appears that System 1 and System 2 can be concurrently active,[39] that automatic and controlled cognitive operations compete for the control of overt responses,[40] and that much of the time, deliberative judgments will remain anchored on initial impressions.[41]

B. Accessibility

A defining property of intuitive thoughts is that they come to mind spontaneously, like percepts. The technical term for the ease with which mental contents come to mind is *accessibility*.[42] To understand intuition in

---

32. *See* Cushman et al., *supra* note 11, at 1086; Hauser, et al., *supra* note 3, at 16–17. In this sense, an understanding of moral intuitions does legal realism one better. Where the realists believed that legal reasoning often masked the actual grounds for judicial judgments, see Karl N. Llewellyn, *Some Realism About Realism—Responding to Dean Pound*, 44 HARV. L. REV. 1222, 1228, 1233 (1931), the psychological point is that judges may not even be aware of those grounds.

33. *C.f.* DREW WESTEN, THE POLITICAL BRAIN: THE ROLE OF EMOTION IN DECIDING THE FATE OF THE NATION (2007) (making a claim of this general sort for political campaigns).

34. *See* Borg et al., *supra* note 22, at 808.

35. *See* Melissa L. Finucane et al, *The Affect Heuristic in Judgments of Risks and Benefits*, 13 J. BEHAV. DECISION MAKING 1, 8 (2000).

36. *See* Herbert Bless et al., *Mood and the Use of Scripts: Does a Happy Mood Really Lead to Mindlessness?*, 71 J. PERSONALITY & SOC. PSYCHOL. 665, 665 (1996); Alice M. Isen et al., *Influence of Positive Affect on the Subjective Utility of Gains and Losses: It is Just Not Worth the Risk*, 55 J. PERSONALITY & SOC. PSYCHOL. 710, 716 (1988).

37. *See generally* Stanovich & West, *Individual Differences in Reasoning*, *supra* note 21.

38. *See* Greene, *The Secret Joke*, *supra* note 2, at 45; Haidt, *The Emotional Dog*, *supra* note 2, at 814–15.

39. *See* Greene & Haidt, *supra* note 2, at 522; Dan Cassino & Milton Lodge, *The Primacy of Affect in Political Evaluations*, *in* THE AFFECT EFFECT, *supra* note 4, at 101, 106–07.

40. For evidence, see Greene, *The Secret Joke*, *supra* note 2, at 45; Alan G. Sanfey et al., *The Neural Basis of Economic Decision-Making in the Ultimatum Game*, 300 SCIENCE 1755, 1756–57 (2003).

41. *See* Haidt, *The Emotional Dog*, *supra* note 2, at 814.

42. E. Tory Higgins, *Knowledge Activation: Accessibility, Applicability, and Salience*, *in* SOCIAL PSYCHOLOGY: HANDBOOK OF BASIC PRINCIPLES 133 (E. Tory Higgins & Arie Kruglanski eds., 1996).

general and the operation of indignation in particular, we must understand why some thoughts are accessible and others are not.

Some attributes are more accessible than others, both in perception and in judgment. Attributes that are routinely and automatically produced by the perceptual system or by System 1, without intention or effort, have been called *natural assessments*.[43] For example, experimental evidence shows that when a perceiver is exposed to a set of objects of the same general kind (e.g., a set of lines of different size), attributes of a prototypical member of the set (e.g., the average length of the lines) are computed effortlessly and automatically. Other attributes (e.g., the total length of the lines) are not accessible—they can only be assessed by a deliberate and quite laborious computation. Thus, average length is a natural assessment, but total length is not. Daniel Kahneman and Shane Frederick compiled a partial list of these natural assessments.[44] In addition to physical properties such as size, distance, and loudness, the list includes more abstract properties such as similarity, causal propensity, surprisingness, affective valence, and mood.

The evaluation of stimuli as good or bad is a particularly important natural assessment. The evidence, both behavioral[45] and neuro-physiological,[46] is consistent with the idea that the assessment of whether objects are good (and should be approached) or bad (should be avoided) is carried out quickly and efficiently by specialized neural circuitry.[47] A remarkable experiment reported by John Bargh illustrates the speed of the evaluative process, and its direct link to approach and avoidance.[48] Participants were shown a series of stimuli on a screen, and instructed to respond to each stimulus as soon as it appeared, by moving a lever that blanked the screen. The stimuli were affectively charged words, some positive (e.g., LOVE) and some aversive (e.g., VOMIT), but this feature was irrelevant to the participant's task. Half the participants responded by pulling the lever toward themselves, half responded by pushing the lever away. Although the response was initiated within a fraction of a second,

---

43. Amos Tversky & Daniel Kahneman, *Extensional Versus Intuitive Reasoning: The Conjunction Fallacy in Probability Judgment*, 90 PSYCHOL. REV. 293, 294 (1983).

44. Daniel Kahneman & Shane Frederick, *Representativeness Revisited: Attribute Substitution in Intuitive Judgment*, *in* HEURISTICS AND BIASES, *supra* note 21, at 49, 55.

45. *See* John Bargh, *The Automaticity of Everyday Life*, *in* THE AUTOMATICITY OF EVERYDAY LIFE: ADVANCES IN SOCIAL COGNITION 1 (Robert S. Wyer, Jr. ed., 1997); Robert B. Zajonc, *Emotions*, *in* HANDBOOK OF SOCIAL PSYCHOLOGY 591 (D.T. Gilbert et al. eds., 4th ed. 1998).

46. *See, e.g.*, Joseph E. LeDoux, *Emotional Circuits in the Brain*, 23 ANN. REV. NEUROSCIENCE 155, 155 (2000).

47. For a demonstration with respect to risks, see Michael Siegrist et al., *Implicit Attitudes Toward Nuclear Power and Mobile Phone Base Stations: Support for the Affect Heuristic*, 26 RISK ANALYSIS 1021 (2006).

48. Bargh, *supra* note 45, at 27.

well before the meaning of the stimulus was consciously registered, the emotional valence of the word had a substantial effect. Participants were relatively faster in pulling a lever toward themselves (approach) for positive words, and relatively faster pushing the lever away (avoid) when the word was aversive. The tendencies to approach or avoid were evoked by an automatic process that was not under conscious voluntary control.

Exploring questions related to public policy, several psychologists have investigated the influence of this primordial evaluative system (here included in System 1) on the attitudes and preferences that people adopt consciously and deliberately.[49] The most well-known results come from the implicit-attitude test, designed to measure racial and other biases.[50] The central finding is that most people show an automatic bias against African-Americans, older people, gays and lesbians, and others—even when they are unaware of it, wish to be unbiased, and indeed are stunned to see that they are automatically biased.[51] There is evidence that people's actual behavior is sometimes affected by their automatic biases rather than by their conscious judgments.[52] This evidence bears on many questions in the law of discrimination because it suggests that those who discriminate might not even be aware of that fact, increasing the difficulty of proving unequal treatment even when it has occurred.[53]

The implicit-attitude test has also been used to show that people tend to be intuitively opposed to nuclear power, even when they are not opposed to it consciously or on reflection, and indeed even when they do not believe that they are opposed to it in any way.[54] An evident implication, consistent with recent political reality, is that public officials will have difficulty in convincing the public to support nuclear power because people's affective

---

49. *See, e.g.*, Zajonc, *supra* note 45, at 601–06; Daniel Kahneman et al., *Economic Preferences or Attitude Expressions? An Analysis of Dollar Responses to Public Issues*, 19 J. RISK & UNCERTAINTY 203, 206–10 (1999); Paul Slovic et al., *The Affect Heuristic*, *supra* note 24, at 397; Seymour Epstein, *Cognitive-Experiential Self-Theory of Personality*, *in* 5 HANDBOOK OF PSYCHOLOGY: PERSONALITY AND SOCIAL PSYCHOLOGY 159–84 (T. Millon & M.J. Lerner eds., 2003).

50. A variety of such tests are available through Project Implicit, https://implicit.harvard.edu/implicit/ (follow "Demonstration" hyperlink) (last visited Apr. 8, 2009). *See generally Symposium on Behavioral Realism*, 94 CAL. L. REV. 945 (2006).

51. *See* Brian A. Nosek et al., *Harvesting Implicit Group Attitudes and Beliefs from a Demonstration Website*, 6 GROUP DYNAMICS: THEORY RES. & PRAC. 101, 101–07 (2002). On legal implications, see generally Christine Jolls & Cass R. Sunstein, *The Law of Implicit Bias*, 94 CAL. L. REV. 969 (2006).

52. *See* Anthony G. Greenwald & Linda Hamilton Krieger, *Implicit Bias: Scientific Foundations*, 94 CAL. L. REV. 945, 946 (2006). The relationship between implicit bias and actual behavior is disputed. *See* Gregory Mitchell & Philip E. Tetlock, *Antidiscrimination Law and the Perils of Mindreading*, 67 OHIO ST. L.J. 1023, 1029–30 (2006).

53. *See* Jerry Kang, *Trojan Horses of Race*, 118 HARV. L. REV. 1489, 1513–14 (2005).

54. *See* Michael Siegrist et al., *supra* note 47, at 1025.

systems are opposed to it.[55] A more general lesson is that affective assessments of products, processes, and risks will often drive people's conclusions, and that statistical realities will play a secondary role.[56]

For a striking illustration of the interaction among indignation, System 1, and moral judgments, consider the Ultimatum Game,[57] which has become a staple of analysis of fairness judgments in domains relevant to politics and law.[58] In this game, subjects are randomly assigned to the roles of "proposers" or "responders." Proposers suggest a division of a stated amount of money (say, $10) with responders; responders can answer "yes" or "no." If responders answer yes, both sides receive money in accordance with the proposed division. If responders answer no, neither side receives any money. The standard economic prediction is that proposers will suggest that they receive nearly all of the money and that responders will agree; self-interested behavior on both sides would suggest that outcome.[59] This is not what happens. Responders often reject, with indignation, offers that are worse than 60:40, and 50:50 divisions are common.[60]

For purposes of analyzing the role of System 1 in judgments related to fairness, consider a few recent findings. When responders are provided with unequal offers, identifiable sectors of the brain associated with emotions are unusually active—and when responders do accept unequal offers, there is unusual activity in the sectors of the brain associated with cognitive control.[61] Studies of skin conductance activity, measuring affect, find that such activity is higher for unfair offers and associated with rejection of such offers.[62] Feelings of anger are therefore a more accurate predictor of whether people will reject unfair offers than is the unfairness of the offer itself.[63] Notably, skin conductance activity is not shown for offers that are generated by computers.[64] Indignation drives responders' behavior in the

---

55. *Id.* at 1026.

56. *See* Peter M. Sandman et al., *Communications to Reduce Risk Underestimation and Overestimation*, 3 RISK DECISION & POL'Y 93 (1998) [hereinafter Sandman et al., *Communications to Reduce Risk*]; Peter M. Sandman et al., *Agency Communication, Community Outrage, and Perception of Risk: Three Simulation Experiments*, 13 RISK ANALYSIS 35 (1994); Slovic, *supra* note 15, at 84, 89, 91.

57. For a good overview, see Richard H. Thaler, *The Ultimatum Game*, *in* THE WINNER'S CURSE 21 (1992).

58. *See, e.g.*, WARD FARNSWORTH, THE LEGAL ANALYST: A TOOLKIT FOR THINKING ABOUT THE LAW 209–46 (2007); Christine Jolls et al., *A Behavioral Approach to Law and Economics*, 50 STAN. L. REV. 1471, 1498 (1998).

59. Thaler, *supra* note 57, at 23.

60. *Id.* at 35.

61. *See* Sanfey et al., *supra* note 40, at 1757.

62. *See* Mascha van 't Wout et al., *Affective State and Decision-Making in the Ultimatum Game*, 169 EXPERIMENTAL BRAIN RES. 564, 566 (2006).

63. *Id.* at 567.

64. *Id.* at 566.

Ultimatum Game, and it ensures that people will sacrifice their material self-interest in order to punish unfairness. It is reasonable to speculate that when people punish defectors from a collective endeavor,[65] at the expense of their own self-interest, similar processes are also at work. Voluntary solutions of collective-action problems, making law unnecessary, are made possible in part by the fact that would-be defectors anticipate punishment.

C. Attribute Substitution

To complete this sketch of the operations of System 1, we explore a process of *attribute substitution* that shapes many judgments and choices. The concept was introduced by Kahneman and Frederick as a basic mechanism to explain the results of heuristic judgment.[66] The basic idea is that the reduction of complex tasks to simpler operations, which characterizes such judgments, is achieved by an operation in which an individual assesses "a specified *target attribute* of a judgment object by substituting another property of that object—the *heuristic attribute*—which comes more readily to mind."[67] In the legal domain, as I show later, individuals charged with the task of determining the severity of a punishment appear to solve this difficult problem by consulting the intensity of their outrage.

Several of the processes explored thus far are involved in the explanation of a study in the area of tort law reported by Miller and McFarland in which respondents determined the appropriate compensation for a man who was shot in the arm during the robbery of a grocery store.[68] Some respondents were told that the robbery happened at the victim's regular store. Other respondents were told that the victim was shot in a store that he visited for the first time because his usual store happened to be closed that day. The two versions obviously differ in poignancy, because the counterfactual "undoing" of an unusual event comes more easily to mind than the undoing of a normal occurrence.

The difference of poignancy translated into a remarkable difference of $100,000 in the median award judged appropriate for the two cases.[69] The participants in this experiment apparently answered the difficult question of appropriate question by mapping onto a scale of dollars their answer to a simple question: How much were they emotionally touched by the story?

---

65. Ernst Fehr & Simon Gächter, *Altruistic Punishment in Humans*, 415 NATURE 137, 139 (2002).

66. Kahneman & Frederick, *supra* note 44, at 53.

67. *Id.*

68. D.T. Miller & C. McFarland, *Counterfactual Thinking and Victim Compensation: A Test of Norm Theory*, 12 PERS. & SOC. PSYCHOL. BULL. 513, 514–16 (1986).

69. *Id.* at 515.

It is most unlikely that the respondents deliberately chose to provide this large compensation for poignancy. Indeed, when respondents were presented with both versions of the robbery story and asked whether a compensation board should make different awards in the two cases, ninety percent thought it should not.[70] In the terms of the present discussion, the emotion-anchored process that produced the initial awards is dominated by System 1. The requirement to compare two questions evokes a much more complex activity, here attributed to System 2, which identifies the distinctive element that separates the two versions and is unable to find any moral justification for different awards. This can be seen as an instance of "moral dumbfounding,"[71] in which a strong intuition exists that cannot be anchored in rules that the person consciously accepts.

II. OUTRAGE, PUNISHMENT, AND HARM

Along with several coauthors, I have studied the operation of moral judgments in the particular domain of punitive-damage awards.[72] One of our hypotheses, couched in the language of the present treatment, was that the setting of such awards is mediated by an *outrage heuristic*.[73] Dollar awards are highly variable, and the variability presents its own puzzles;[74] the concern here is the operation of the outrage heuristic and its relationship to System 1. The most general finding is that even if they state a commitment to deterrence, people are intuitive retributivists, and their judgments about appropriate monetary punishment have their origin in outrage.[75]

A. The Role of Harm

Participants drawn from a jury roll in Texas were shown vignettes of cases in which a plaintiff had suffered a personal injury while using a product. For example, one of the scenarios concerned a child who had been burned when his pajamas caught fire as he was playing with matches. The

---

70. *Id.*

71. Haidt, *The New Synthesis*, *supra* note 2, at 317.

72. *See generally* Daniel Kahneman, David Schkade & Cass R. Sunstein, *Shared Outrage and Erratic Awards: The Psychology of Punitive Damages*, 16 J. RISK & UNCERTAINTY 49 (1998) [hereinafter Kahneman et al., *Shared Outrage*]; Cass R. Sunstein, Daniel Kahneman & David Schkade, *Assessing Punitive Damages*, 107 YALE L.J. 2071 (1998).

73. We understand the term "indignation" as interchangeable with the term "outrage." The outrage heuristic might be seen as a special case of the affect heuristic, discussed in Slovic, *supra* note 15, at 82–86.

74. *See* Kahneman et al., *Shared Outrage*, *supra* note 72, at 53.

75. Cass R. Sunstein, Daniel Kahneman, David Schkade & Ilana Ritov, *Predictably Incoherent Judgments*, 54 STAN. L. REV. 1153, 1167 (2002) [hereinafter Sunstein et al., *Predictably Incoherent*].

pajamas were made of fabric that was not adequately fire-resistant, and the defendant firm had been aware of the problem. For some of the scenarios, alternative versions were constructed that differed in the severity of harm. In the high-harm version of the pajamas case, for example, the child was "severely burned over a significant portion of his body and required several weeks in hospital and months of physical therapy." In the low-harm version, "his hands and arms were badly burned and required professional medical treatment for several weeks." Participants were told that the plaintiff had already been awarded compensatory damages. One group of respondents indicated whether *punitive damages* were appropriate, and if so in what amount. Another group rated the *outrageousness* of the defendant's behavior. In a subsequent re-analysis of this study, Kahneman and Frederick also obtained ratings of the *severity of the harm* suffered in each of the fourteen vignettes.[76] Lawsuits were not mentioned in these descriptions of harm. The same basic design was replicated twice, varying the size of the defendant firm.

The results supported the conclusion that assessments of punitive damages (the target attribute in this study) were mediated by an outrage heuristic.[77] In the analysis offered by Kahneman and Frederick, the outrage associated with each case was estimated by the product of the product of the average ratings of outrageousness and of harm.[78] The correlations (over fourteen vignettes) between the estimate of outrage and mean punitive damages were 0.90 in one of the firm-size conditions and 0.94 in the other.

The role of actual harm as a determinant of outrage in this experiment is also of interest as a potential case of moral dumbfounding.[79] The legally recognized distinction between murder and attempted murder is a salient example of the issue. Consider the following scenarios:

1. A wishes B dead but does nothing about it
2. A tries to kill B and fails by chance
3. A tries to kill B and succeeds

---

76. Kahneman & Frederick, *supra* note 44, at 63.

77. To the same general effect, see Kevin M. Carlsmith et al., *Why Do We Punish? Deterrence and Just Deserts as Motives for Punishment*, 83 J. PERS. & SOC. PSYCHOL. 284, 292–93 (2002).

78. Kahneman & Frederick, *supra* note 44, at 64.

79. The effect of the outrage heuristic can also be seen in the finding that contrary to the standard economic account, people do not want to increase punitive awards when the likelihood of detection is low or to decrease such awards when the likelihood of detection is high. They respond to the outrageousness of the underlying conduct; the likelihood of detection is relevant, if at all, only because it bears on that question. *See* Cass R. Sunstein, David Schkade & Daniel Kahneman, *Do People Want Optimal Deterrence?*, 29 J. LEGAL STUD. 237, 246 (2000); Carlsmith et al., *supra* note 77, at 285–89.

It is not so easy to offer a moral distinction between the last two cases. Indeed, it is safe to assume that if people are asked to judge the outrageousness of the *actions*, there will be no difference. But punitive intent reflects the emotional intensity of the response to the event, and the emotion evidently depends on the harm that actually occurred. In the terms of the present analysis, the severity of punishment reflects the intensity of an emotional reaction in System 1. Punishments that are determined in this manner are expected to be crudely retributive, which is what we observe.[80] Note that the argument here is not that it is impossible to defend the distinction, drawn by the criminal law, between murder and attempted murder. There may be good reasons for drawing that distinction.[81] What I am suggesting is that the distinction is not *caused* by those reasons, supposing they exist;[82] it is caused by the fact that moral intuitions, automatic and uncontrolled, are different in the two cases.

B. Outrage and Risk

The outrage heuristic helps to explain a wide range of moral judgments of relevance to policy and law, especially in the domain of risk regulation. Consider, for example, the evident fact that many jurors are outraged by the practice of cost-benefit analysis, in the sense that juries strenuously object to a corporate decision to trade off lives and dollars—even when lives have been highly valued.[83] Legitimate questions can be raised about cost-benefit analysis,[84] but public outrage, to the extent that it exists, is rooted in a strong intuition that people should not act with the knowledge that their action will cause people to die. Cost-benefit analysis is a matter for System 2; System 1 reacts by asking whether the defendant proceeded with knowledge that its actions would lead to human deaths.[85]

Or consider widespread public skepticism about emissions-trading programs, by which polluters are given pollution rights and permitted to trade those rights for a fee.[86] Many people are outraged by such programs,

---

80. *See* Kahneman et al., *Shared Outrage*, *supra* note 72, at 52.

81. *See generally* Michael Davis, *Why Attempts Deserve Less Punishment Than Complete Crimes*, 5 LAW & PHIL. 1 (1986).

82. For a similar conclusion in the context of choices among political candidates, see Cassino & Lodge, *supra* note 39, at 101, 119.

83. *See* W. Kip Viscusi, *Corporate Risk Analysis: A Reckless Act*, 52 STAN. L. REV. 547, 550 (2000).

84. *See* FRANK ACKERMAN & LISA HEINZERLING, PRICELESS 8–12 (2005).

85. *See* Greene, *The Secret Joke*, *supra* note 2, at 45. This claim is consistent with the findings in Viscusi, *supra* note 83, at 571.

86. For a valuable discussion, see generally A. DENNY ELLERMAN, MARKETS FOR CLEAN AIR (2000).

and sometimes their outrage has been developed into elaborate critiques of trading programs.[87] I speculate that their outrage is founded in a simple heuristic, to the effect that people should not be paid for agreeing not to commit wrongs. That heuristic generally works well, but it misfires as applied to the context of emissions-trading programs, which often appear to be the most effective and efficient means of handling many environmental problems.[88]

Several studies have attempted to explore whether outrage operates as an amplifier with respect to people's perceptions of risks.[89] These studies hypothesized that certain low-probability risks, such as those associated with nuclear waste radiation, produce outrage, whereas other low-probability risks, such as those associated with radon exposure, do not. (Recall that people tend to show an automatic aversion to nuclear power.[90]) A central finding is consistent with the account offered here: a large difference in probability had no effect in the "high outrage" condition, with people responding the same way to a risk of 1/100,000 as to a risk of 1/1,000,000.[91]

More striking still: even when the risk was identical in the nuclear waste (high outrage) and radon (low outrage) cases, people in the nuclear waste case reported a much greater perceived threat and a much higher intention to act to reduce that threat.[92] Indeed, "the effect of outrage was practically as large as the effect of the 4000-fold difference in risk between the high-risk and low-risk conditions."[93] Efforts to communicate the meaning of differences in risk levels, by showing comparisons to normal risk levels, reduced the effect of outrage, but even after those efforts, outrage had nearly the same effect as a 2000-fold increase in risk.[94] More generally, choices among political candidates have been found to turn on affect and on System 1, even when people are quite unaware of that fact.[95]

---

87. *See generally* STEVEN KELMAN, WHAT PRICE INCENTIVES? (1981).

88. *See* RICHARD B. STEWART & JONATHAN WIENER, RECONSTRUCTING CLIMATE CHANGE POLICY: BEYOND KYOTO 66 (2003). On paying people not to commit wrongs, see Guido Calabresi & A. Douglas Melamed, *Property Rules, Liability Rules, and Inalienability: One View of the Cathedral*, 85 HARV. L. REV. 1089, 1115–24 (1972).

89. *See, e.g.*, Sandman et al., *Communications to Reduce Risk*, *supra* note 56, at 106.

90. Siegrist et al., *supra* note 47, at 1025.

91. Sandman et al., *Communications to Reduce Risk*, *supra* note 56, at 102.

92. *Id.* at 106.

93. *Id.*

94. *Id.*

95. *See* Cassino & Lodge, *supra* note 39, at 119.

C. The Severity Shift and Rhetorical Asymmetry

What happens when indignant people deliberate with one another? It might be tempting to suppose that they would converge on the judgment of the group's median member. In fact, however, deliberating groups end up more indignant than their median member, and the consequence can be especially severe punishment.[96]

Over 500 deliberating juries, consisting of six people, were asked to record their judgments in advance of deliberation on three different questions: the outrageousness of the defendant's conduct on an eight-point scale; the appropriate punishment, also on an eight-point scale; and the appropriate dollar award.[97] As we would predict, the correlation between outrage judgments and punishment judgments was quite close.[98] In both cases, juries whose members began with a high degree of outrage (four or higher) produced "verdicts" that were systematically higher than those of the jury's median member—in a general "severity shift."[99] By contrast, low-outrage jurors (three or lower) ended up being more lenient, as juries, than their median member—in a general "leniency shift."[100]

With dollars, the result was even more dramatic: for the overwhelming majority of positive awards, the jury's verdict was higher than that of the median juror.[101] And in twenty-seven percent of the cases, the jury's verdict was at least as high as that of the highest juror.[102] For punitive-damage awards, deliberation produces a systematic shift in the direction of greater severity.

These findings might be explained in two different ways. Because of the robust phenomenon of group polarization,[103] it would be predicted that outraged *juries* would be more outraged than outraged *jurors*. A key reason involves the exchange of information. Such exchange, within a group of people antecedently inclined to show outrage, tends to produce an intensification of their antecedent inclination.[104] But peer pressure is also important: jurors are not likely to want to seem to be unconcerned with serious wrongdoing, even within a group of strangers, and for that reason deliberation among outraged jurors is likely to fuel outrage.[105]

---

96. *See* David Schkade et al., *Deliberating About Dollars: The Severity Shift*, 100 COLUM. L. REV. 1139, 1164 (2000).
97. *Id.* at 1149.
98. *Id.* at 1152.
99. *Id.* at 1153.
100. *Id.* at 1155–56.
101. *Id.*
102. *Id.*
103. *See* ROGER BROWN, SOCIAL PSYCHOLOGY 200 to 243 (2d ed. 1986).
104. *See id.* at 217–22.
105. *See id.* at 213–17.

With respect to both moral judgments and dollar awards, indignation can also be intensified as a result of *rhetorical asymmetry*.[106] In some domains, one or another position has an automatic upper hand, in the sense that people find it easier to support that position in the face of social conflict. Evidence supports the view that when a group of people disagree about the appropriate monetary punishment for corporate wrongdoing, it is simply easier to argue in favor of the higher award.[107] Existing social norms are responsible for the existence of rhetorical asymmetry. It is reasonable to speculate that the asymmetry is likely to be present, and to be especially severe, when the issue is simple rather than complex, in the particular sense that people do not perceive tradeoffs to be present. If, for example, a large damage award were thought to have adverse effects on innocent employees or consumers, the asymmetry might well be diminished or eradicated.[108] And if one side is able to appeal to "core values" of one or another kind, such as the protection of human life, then a rhetorical asymmetry is more likely to be in play. But much work remains to be done on this subject.

An understanding of the effects of social interactions on the operation of the outrage heuristic and System 1 has many implications. It is reasonable to think that a rhetorical asymmetry helps explain why it is easier, in familiar times and places, to argue for stiffer punishments for drug offenders and murderers—and for decreases rather than increases in tax rates. When freedom of association leads certain groups to be especially outraged about past or present treatment, group polarization and rhetorical asymmetry provide at least part of the picture.[109] "Moral panics," involving epidemics of outrage directed against certain practices and groups, are much influenced by the mechanisms sketched here.[110] In addition, social interactions, including rhetorical asymmetry, help to explain the wellsprings of terrorism, which is typically a product not of poverty, poor education, or mental illness,[111] but of social networks that attempt to fuel outrage.[112]

---

106. *See* Schkade et al., *supra* note 96, at 1161.

107. *Id.* at 1162.

108. *Cf.* Jonathan Baron & Susan Leshner, *How Serious are Expressions of Protected Values?*, 6 J. EXPERIMENTAL PSYCHOL.: APPLIED 183 (2000) (finding that commitments to protected values, as immune from tradeoffs, become weaker when tradeoffs are made explicit).

109. *Cf.* Timur Kuran, *Ethnic Norms and Their Transformation Through Reputational Cascades*, 27 J. LEGAL STUD. 623 (1998) (exploring use of social pressures in producing ethnic antagonism).

110. *See* STANLEY COHEN, FOLK DEVILS AND MORAL PANICS 197–98 (2d ed. 1987). *See generally* KENNETH THOMPSON, MORAL PANICS (1998).

111. *See* ALLAN KRUEGER, WHAT MAKES A TERRORIST? 6 (2007) (emphasizing that poor education and poverty do not appear to contribute to terrorism).

112. *See generally* MARC SAGEMAN, UNDERSTANDING TERROR NETWORKS (2004).

III. MORAL FRAMING AND INDIGNATION

A framing effect is said to occur when two extensionally equivalent statements evoke different judgments or preferences when presented singly, yet appear transparently equivalent when shown together.[113] Framing effects arise because statements that are extensionally equivalent may nevertheless evoke different associations and different emotional responses. Consider the question whether to undergo a risky medical procedure. When people are told, "Of those who have this procedure, ninety percent are alive after five years," they are far more likely to agree to the procedure than when they are told, "Of those who have this procedure, ten percent are dead after five years."[114] Experience might be expected to solve this problem, but doctors too are vulnerable to this framing effect.[115] Similarly, a cold cut described as ninety percent fat-free is more attractive than if it is described as ten percent fat, and more likely to be purchased. Framing effects are a manifestation of the associative and emotional processes of System 1. There have been several demonstrations of framing effects in the domain of moral judgments.[116]

A. Losses and Gains

Consider the valuation of injuries to health, and an experiment in which the same difference between two states of health was caused to be coded either as a loss or as a gain.[117] The experiment was concerned with lay assessments of appropriate monetary compensation for the pain and suffering associated with personal injuries, such as "losing mobility in one knee for four years."[118] Separate samples of respondents were given different jury instructions describing the thought experiment they should conduct to determine fair compensation. One of the instructions suggested a positive choice between two desirable options. The respondents were instructed to imagine that the victim had very recently suffered the injury and was now offered a choice between a complete and immediate cure and an amount of money. Fair compensation was to be set at the highest amount

113. *See generally* Amos Tversky & Daniel Kahneman, *The Framing of Decisions and the Psychology of Choice*, 211 SCIENCE 453 (1981).

114. *See* Donald A. Redelmeier, Paul Rozin & Daniel Kahneman, *Understanding Patients' Decisions*, 270 JAMA 72, 73 (1993).

115. *Id.*

116. *See, e.g.*, Shane Frederick, *Measuring Intergenerational Time Preference:* Are *Future Lives Valued Less?*, 26 J. RISK & UNCERTAINTY 39, 41–47 (2003).

117. Edward J. McCaffery, Daniel J. Kahneman & Matthew L. Spitzer, *Framing the Jury: Cognitive Perspectives on Pain and Suffering Awards*, 81 VA. L. REV. 1341, 1354–57 (1995).

118. *Id.* at 1408.

for which the victim would still prefer the cure. In contrast, the selling instruction required the respondent to assume that the victim considered an ex ante proposition to accept the injury in return for a payment of money. Fair compensation was to be set at the lowest payment for which the victim would have accepted the offer.

The difference between health and injury is coded as a gain in the former case and as a loss in the latter—this is the pattern of an endowment effect.[119] In terms of final states, of course, the two versions of the problem are not distinguishable. As expected, the average judgment of fair compensation was about twice as high with the selling than with the choice instruction.[120] This is also a framing effect: when the participants in each experimental condition were shown the instruction given to the other group, they thought both instructions were fair and did not notice that they were likely to evoke discrepant responses.[121]

The legal system typically uses a version of the buying instruction rather than the selling instruction. Jurors are asked what amount would place plaintiffs in the position that they would have occupied if the injury had never occurred, and it is impermissible for plaintiff's lawyers to ask jurors to focus on the amount that the plaintiff would have had to be paid to accept the injury in the first instance.[122] But courts do not undertake a great deal of reflective thinking about why the buying instruction should be preferred, and in general, there is much dispute about whether goods should be valued by reference to willingness to accept or willingness to pay.[123] In any event, creative lawyers are sometimes able to frame the problem so as to ensure that a selling instruction comes before the jury, in a way that produces predictably higher dollar awards.[124]

B. Valuing Life

Moral framing has been demonstrated in the important context of obligations to future generations,[125] a much-disputed question of morality, politics, and law[126] with particular importance for the issue of climate

119. *See* RICHARD H. THALER, QUASI RATIONAL ECONOMICS 8 (1991); Russell Korobkin, *The Endowment Effect and Legal Analysis*, 97 NW. U. L. REV. 1227, 1228 (2003).

120. McCaffery et al., *supra* note 117, at 1359.

121. *Id.* at 1371.

122. W. Kip Viscusi, *Pain and Suffering in Product Liability Cases: Systematic Compensation or Capricious Awards?*, 8 INT'L REV. L. & ECON. 203, 203 (1988).

123. *See generally* Korobkin, *supra* note 119.

124. McCaffery et al., *supra* note 117, at 1399.

125. *See* Frederick, *supra* note 116, at 46–47.

126. *See generally* Symposium, *Intergenerational Equity and Discounting*, 74 U. CHI. L. REV. 1 (2007); Richard Revesz, *Environmental Regulation, Cost–Benefit Analysis, and the Discounting of*

change.[127] Most people have not given a great deal of thought to the appropriate discount rate for those yet to be born, and hence their judgments are highly susceptible to different frames. With some frames, lower weighting of future generations will seem natural and unexceptionable. With other frames, people will find it outrageous to suggest that future people should be given less attention than current people. The reason is that some frames will trigger System 1, producing indignation by suggesting that some people are "worth less" than others.

From a series of surveys, Maureen Cropper and her coauthors suggest that people are indifferent between saving one life today and saving 44 lives in 100 years.[128] They make this suggestion on the basis of questions—asking people whether they would choose a program that saves "100 lives now" or a program that saves a substantially larger number "100 years from now." It is possible, however, that people's responses depend on uncertainty about whether people in the future will otherwise die (perhaps technological improvements will save them?); other ways of framing the same problem yield radically different results.[129] For example, most people consider "equally bad" a single death from pollution next year and a single death from pollution in 100 years.[130] This finding implies no preference for members of the current generation. The simplest conclusion is that people's moral intuitions about obligations to future generations are very much a product of framing effects.[131]

The same point holds for the question whether government should consider not only the number of "lives" but also the number of "life-years" saved by regulatory interventions.[132] If the government focuses on life-years, a program that saves children will be worth far more money that a similar program that saves senior citizens. Is this immoral? People's intuitions, and their tendency toward indignation, depend on how the question is framed.[133] If people are asked whether they would favor a policy that saves 105 old people or 100 young people, many will favor the latter, in a way that suggests a willingness to pay considerable attention to the

---

Human Lives, 99 COLUM. L. REV. 941 (1999).

127. *See* NICHOLAS STERN, THE ECONOMICS OF CLIMATE CHANGE 35 (2007).

128. Maureen Cropper et al., *Preferences for Life-Saving Programs: How the Public Discounts Time and Age*, 8 J. RISK & UNCERTAINTY 243, 244 (1994).

129. *See* Frederick, *supra* note 116, at 41–47.

130. *Id.* at 42.

131. For a similar result, see Jonathan Baron, *Can We Use Human Judgments to Determine the Discount Rate?*, 20 RISK ANALYSIS 861, 866 (1993).

132. For a brief overview of the controversy, see Jocelyn Kaiser, *How Much Are Human Lives and Health Worth?*, 200 SCIENCE 1836, 1836–37 (2003).

133. *See* Cass R. Sunstein, *Lives, Life-Years, and Willingness to Pay*, 104 COLUM. L. REV. 205, 244 (2004).

number of life-years at stake. At the same time, people will predictably reject as outrageous an approach that would count every old person as "worth less" than what every young person is worth.

If people are asked whether safety and health policies should adopt a "senior death discount," or assign a monetary value, for those over sixty years of age, that is worth some fraction of the monetary value assigned for all others, there will be a high degree of indignation. System 1 rebels against the idea that older people are worth (say) sixty percent of what younger people are worth (even if System 2 might ultimately be persuaded that the life-years approach is the right one). Facing such indignation, the national government eventually retreated from a suggestion that federal agencies should adopt a "senior death discount" to take account of the fact that some policies mostly helped people who were already old and thus had relatively few years left.[134]

C. Coherence and Incoherence

Framing effects present a large difficulty for the achievement of coherent judgments and preferences. The normal process of comprehension takes a given message to a state of the world, but the correspondence of messages and states is not one-to-one. Ambiguity arises when a single message is compatible with multiple states of the world. Framing effects arise when a single state of the world may be described in multiple ways, and when a relevant response is description-dependent. Thus, the avoidance of framing effects requires a search through the set of descriptions that are extensionally equivalent to the original message. Unfortunately, the human mind is not equipped to solve this problem.

In the context of both punitive-damage awards and valuation of environmental amenities, my coauthors and I have found that incoherence in both moral and legal judgments is predictable.[135] The problem is especially severe for judgments rooted in indignation, the intensity of which depends on the relevant comparison set. To take a mundane example, rude behavior by a guest at a dinner table can produce extremely intense indignation on the part of a host, even if such behavior is trivial compared to (say) theft, assault, and child abuse. System 1 is deeply offended by rude behavior at dinner in part because such behavior is automatically compared to standard conduct at dinner—not to a wide range of inappropriate or bad behavior in which human beings engage. If a host at dinner takes the guest's rude behavior in the

134. Cindy Skryzycki, *Under Fire, EPA Drops the "Senior Death Discount,"* WASH. POST, May 13, 2003, at E1.

135. Sunstein et al., *Predictably Incoherent*, *supra* note 75, at 1165.

context of much worse conduct, indignation is likely to be greatly dampened. It turns out that punishment judgments have a similar structure.

The basic point is that such judgments about cases, taken one at a time, are very different from judgments about the same cases, taken in the context of a problem from another category. In the relevant experiments, people were asked to assess a case involving a personal injury on a bounded scale and also on a dollar scale.[136] People were also asked to assess a case involving financial injury on a bounded scale and also on a dollar scale. When the two cases were judged in isolation, the financial injury case received a more severe rating and a higher dollar award. But when the two cases were seen together, there was a significant *judgment shift*, in which people tried to ensure that the financial award was not higher than the personal injury award.[137] People's decisions about the two cases were very different, depending on whether they saw a case alone or in the context of a case from another category.[138]

An explanation for the shift starts with the suggestion that when people see a case in isolation, they naturally "normalize" it by comparing it to a set of comparison cases that it readily calls up. People easily normalize judgments about size, and the normalization is mutually understood. (Steve Nash, who is a little more than six feet tall, is a very small basketball player.) What happens, in ordinary communication, is innocuous. It does not breed error or confusion. In the context of legally relevant moral judgments, something similar happens, but it is far from innocuous. When evaluating a case involving financial injury, people apparently normalize the defendant's conduct by comparing it with conduct in other cases from the same category. This is a species of "narrow framing," which is a basic property of the human mind.[139] People examine problems in the context of narrow frames, including small sets of similar problems.

It follows that jurors (and in all probability judges too[140]) do not easily or naturally compare that defendant's conduct with conduct from other

136. *Id.* at 1174.

137. *Id.* at 1175–77.

138. Exactly the same kind of shift was observed for judgments about two problems calling for government regulation and expenditures: skin cancer among the elderly and protection of coral reefs. Looking at the two cases in isolation, people wanted to pay more to protect coral reefs and registered more satisfaction from doing that. But looking at the two cases together, people were quite disturbed at this pattern and generally wanted to pay more to protect elderly people from skin cancer. Here too there was a significant shift in judgment. *Id.* at 1174.

139. *See, e.g.*, Nicholas Barberis & Ming Huang, *The Loss Aversion/Narrow Framing Approach to the Equity Premium Puzzle*, *in* THE HANDBOOK OF THE EQUITY RISK PREMIUM 199 (Rajnish Mehra ed., 2008), *available at* http://papers.ssrn.com/sol3/papers.cfm?abstract_id=912776.

140. See the argument for comparison of cases in Martell v. Boardwalk Enters., Inc., 748 F.2d 740, 752–53, 755 (2d Cir. 1984).

categories. Because of the natural comparison set, people are likely to be quite outraged by the misconduct if it is far worse than what springs naturally to mind. When a case from another category is introduced, this natural process of comparison is disrupted. Rather than comparing a financial injury case to other cases of business misconduct, people now compare it to a personal injury case, which (in most people's view) involves more serious wrongdoing. As a result of the wider viewscreen, judgments shift, often dramatically.

This finding helps to explain a serious problem with current practice in many domains of law.[141] The problem is that when people assess cases in isolation, their viewscreen is narrow, indeed limited to the category to which the case belongs, and that as a result, people produce a pattern of outcomes that makes no sense by their own light. In other words, the overall set of outcomes is one that people would not endorse, if they were only to see it as a whole. Their considered judgments reflect the very pattern that they have produced, because of a predictable feature of human cognition. The result is a form of incoherence.

We can find such incoherence not only in jury verdicts, but also in administrative fines, where no serious effort has been made to ensure that the overall pattern of outcomes makes the slightest sense.[142] Indeed there is reason to believe that the pattern, in many domains, is quite senseless. And it may not be too much of a stretch to suggest that the same is true of reactions, some of the time, by both individuals and institutions—that people are quite outraged about behavior that, in a broader or different comparison set, would outrage them little or not at all.

These observations have obvious relevance both for the attempt to reach coherence in law and for the idea of reflective equilibrium in ethical judgments.[143] For law, the basic lesson is that judgments made one at a time are likely to produce incoherent patterns, and hence it would be useful to attempt to systematize outcomes by seeing them as part of larger comparison sets.[144] In the world of punitive-damages awards, comparisons should produce real improvements over isolated judgments, in the sense that the isolated judgments yield patterns rejected by the very people who are responsible for them. In the world of administrative penalties, established pursuant to congressional guidelines, it would make a great deal

---

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

142. *See* Sunstein et al., *Predictably Incoherent*, *supra* note 75, at 1186–96.

143. *See* RAWLS, *supra* note 29, at 42–45. For relevant discussion, see APPIAH, *supra* note 4.

144. *See* Randall R. Bovbjerg et al., *Valuing Life and Limb in Tort: Scheduling "Pain and Suffering,"* 83 NW. U. L. REV. 908, 924–25 (1989).

of sense to try to produce broader coherence, for example by seeing if fines for Occupational Safety and Health Act violations fit with fines for Clean Air Act violations and for violations of the tax law.

For ethical judgments, the search for reflective equilibrium might seem all the more important in this light, because a wide viewscreen can help to control indefensibly intense reactions to particular cases. On the other hand, the attempt to achieve equilibrium between Systems 1 and 2 can produce real difficulties, the resolution of which we cannot attempt here.[145]

IV. ACTS, OMISSIONS, AND RELATED PROBLEMS

To say the least, there has been much discussion of whether and why the distinction between acts and omissions might matter for morality, law, and policy.[146] In one case, for example, a patient might ask a doctor not to provide life-sustaining equipment, thus ensuring the patient's death. In another case, a patient might ask a doctor to inject a substance that will immediately end the patient's life. Many people seem to have a strong moral intuition that the failure to provide life-sustaining equipment, and even the withdrawal of such equipment, is acceptable and legitimate—but that the injection is morally abhorrent. And indeed American constitutional law reflects judgments to exactly this effect: people have a constitutional right to withdraw equipment that is necessary to keep them alive, but they have no constitutional right to physician-assisted suicide.[147] But what is the morally relevant difference?

It is worth considering the possibility that the act–omission distinction is rooted in System 1, and is in some cases very hard to defend in principle.[148] The moral puzzles arise when life, or a clever interlocutor, comes up with a case in which there is no morally relevant distinction between acts and omissions, but when moral intuitions strongly suggest that there must be such a difference. As an example, consider the question whether to vaccinate one's children; many people show a persistent omission bias, favoring inaction over statistically preferable action.[149] The widespread acceptance of withdrawal of life-saving equipment,

---

145. Paul Slovic and Kwame Anthony Appiah offer relevant discussion from different standpoints. *Compare* APPIAH, *supra* note 4, at 78, 122–23, *with* Slovic et al., *The Affect Heuristic*, *supra* note 24.

146. *See, e.g.*, Judith Jarvis Thomson, *Individuating Actions*, 68 J. PHIL. 774, 777 (1970). *See generally* 1 F.M. KAMM, MORALITY, MORTALITY: DEATH AND WHOM TO SAVE FROM IT 4 (1993).

147. Washington v. Glucksberg, 521 U.S. 702, 724–25 (1997).

148. *See* APPIAH, *supra* note 4.

149. Ilana Ritov & Jonathan Baron, *Reluctance to Vaccinate: Omission Bias and Ambiguity*, 3 J. BEHAV. DECISION MAKING 263, 275 (1993).

alongside persistent doubts about euthanasia, may be another demonstration of the point.

Compare the dispute over two well-known problems in moral philosophy.[150] These problems do not involve the act–omission distinction; no omission is involved. But the problems implicate closely related concerns. The first, called the trolley problem, asks people to suppose that a runaway trolley is headed for five people, who will be killed if the trolley continues on its current course. The question is whether you would throw a switch that would move the trolley onto another set of tracks, killing one person rather than five. Most people would throw the switch. The second, called the footbridge problem, is the same as that just given, but with one difference: the only way to save the five is to throw a stranger, now on a footbridge that spans the tracks, into the path of the trolley, killing that stranger but preventing the trolley from reaching the others. Most people will not kill the stranger; in fact they are indignant at the suggestion that they ought to do so. But what is the difference between the two cases, if any? A great deal of philosophical work has been done on this question, much of it trying to suggest that our firm intuitions can indeed be defended in principle.[151]

Without engaging these arguments, let us suggest the possibility of a simpler answer.[152] As a matter of principle, there may or may not be a difference between the two cases. But people's different reactions are based on automatic moral intuitions that condemn the throwing of the stranger but support the throwing of the switch. As a matter of intuition, it is worse to throw a human being in the path of a trolley than to throw a switch that (indirectly?) leads to a death. People also struggle heroically, and by reference to System 2, to rescue their intuitions and to establish that the two cases are genuinely different in principle, whether or not this is so. But System 1, and indignation about brutal acts of commission, are responsible for the underlying intuitions.

Consider a suggestive experiment designed to see how the human brain responds to the two problems.[153] The authors do not attempt to answer the moral questions in principle, but they find "that there are systematic variations in the engagement of emotion in moral judgment," and that brain areas associated with emotion are far more active in contemplating the

150. *See* Judith Jarvis Thomson, *The Trolley Problem*, *in* RIGHTS, RESTITUTION, AND RISK 94–116 (William Parent ed., 1986).

151. *See id. See also* Philippa Foot, *The Problem of Abortion and the Doctrine of Double Effect*, 5 OXFORD REV. 5, 12–14 (1967).

152. A detailed discussion can be found in Greene, *The Secret Joke*, *supra* note 2, at 41–46.

153. Joshua D. Greene et al., *An fMRI Investigation of Emotional Engagement in Moral Judgment*, 293 SCIENCE 2105, 2106 (2001).

footbridge problem than in contemplating the trolley problem.[154] An implication of the authors' finding is that human brains are hard-wired to distinguish between bringing about a death "up close and personal" and doing so at a distance.[155] It follows that acts, especially brutal acts, would be far more likely to produce reactions from the brain areas associated with emotions than omissions that cause identical harms.

A related study finds that certain forms of brain damage, dampening the social emotions, lead people to accept utilitarian approaches to certain problems, and to reject deontological inclinations that help distinguish between the trolley problem and the footbridge problem.[156] Patients with damage to the ventromedial prefrontal cortex (VMPC) show reduced emotional sensitivity and reduced social emotions, such as compassion, shame, and guilt. Such patients were asked to resolve certain moral dilemmas, including the trolley and the footbridge problems, and other problems asking whether one person should be sacrificed to save several. A control group, consisting of people without VMPC damage, produced the normal responses, with considerable skepticism about utilitarian balancing. By contrast, the VMPC patients were far more likely to be willing to sacrifice one person for the benefit of a larger number.

The authors conclude that VMPC patients, lacking "an emotional reaction" to the relevant harm, are more willing to "rely on explicit norms endorsing the maximization of aggregate welfare."[157] It is possible to understand this finding as a demonstration that when System 1 is damaged, people will rely on System 2, which leads in the direction of welfarism.[158]

Compare the case of fear, where an identifiable region of the brain makes helpfully immediate but not entirely reliable judgments,[159] in a way that suggests a possible physical location for some of the operations of System 1. In the context of risk-related judgments, similar findings have been made, in a way that suggests that those with brain damage can actually do far better in investment decisions.[160] Putting the normative issues to one side, we think that something analogous is true in the context of morality, politics, and law.[161] A clear implication involves moral numbness: many

154. *Id.* at 2107.

155. *Id.* at 2106.

156. *See* Koenigs et al., *supra* note 22, at 908.

157. *Id.* at 910.

158. This understanding is defended in Joshua D. Greene, *Why Are VMPFC Patients More Utilitarian? A Dual-Process Theory of Moral Judgment Explains*, 11 TRENDS COGNITIVE SCI. 322, 322 (2007).

159. *See* LEDOUX, *supra* note 22, at 163–65.

160. *See* Baba Shiv et al., *Investment Behavior and the Negative Side of Emotion*, 16 PSYCHOL. SCI. 435, 436–37 (2005).

161. *See generally* Greene & Haidt, *supra* note 2, at 522.

acts and even more omissions do not trigger indignation on the part of System 1, but might well be subject to moral criticism from the standpoint of System 2, if only it can become or be made sufficiently active.

Consider in this regard the "identifiable victim effect."[162] People will devote substantial resources to save an identifiable victim, and they will be indignant at the failure to make large efforts to assist such a victim. By contrast, "statistical victims" or large groups of nameless people, at serious risk from some harm, often occasion little attention or concern.[163] A potentially beneficial function of some practices, such as cost-benefit analysis, is to bring a System 2 check to bear, ensuring that statistical victims receive serious attention even if people are not indignant about their plight.[164]

I have not suggested that System 2 generally outperforms System 1. People's automatic judgments might be quite good from the moral point of view,[165] and the judgments yielded by System 2 might be erroneous or worse. But in some domains, the intuitive system is likely to be activated when there is little justification for indignation, and to be passive in the face of serious suffering. A potential virtue of institutional safeguards, including efforts to ensure some kind of accounting of actual consequences, is to provide a deliberative check in cases in which System 1 reacts excessively or not at all.[166]

CONCLUSION

Moral intuitions operate in much the same way as other intuitions do; what makes the moral domain distinctive is its frequent foundation in the emotions, beliefs, and response tendencies that define indignation. System 1 is typically responsible for indignation; System 2 may or may not provide an override. Moral dumbfounding and moral numbness are often a product of moral intuitions that people are unable to justify. Both of these have consequences for public policy and law. Thus, for example, some legal outcomes and prohibitions are rooted in automatic, intensely held intuitions that people find hard to justify. Individual and collective inaction, in the face of widespread suffering and distress, often persists because System 1 is difficult to activate.[167] A large task is to produce institutional safeguards to

---

162. *See, e.g.*, Jenni & Loewenstein, *supra* note 16, at 236.

163. *See* Slovic, *supra* note 15, at 86; Lisa Heinzerling, *The Rights of Statistical People*, 24 HARV. ENVTL. L. REV. 189, 189–90 (2000).

164. *See* Allan Gibbard, *Risk and Value*, *in* VALUES AT RISK 94 (Douglas MacLean ed., 1986).

165. An argument to this effect can be found in GERD GIGERENZER, GUT FEELINGS: THE INTELLIGENCE OF THE UNCONSCIOUS (2007).

166. Slovic, *supra* note 15, at 82–88.

167. *Id.*

ensure against the risk that legal and political outcomes will respond to unjustifiably intense indignation, or that democratic societies will remain passive simply because the relevant harms are not of the sort that stir System 1.

An understanding of indignation helps to explain the operation of the outrage heuristic, the centrality of harm, the severity shift, the role of reference states, moral framing, and the use of the act–omission distinction. And because of the nature of indignation, it is extremely difficult for people to achieve coherence in their moral intuitions; the problem of incoherence besets legal outcomes as well, in the areas of jury awards and administrative penalties. A general implication is that people are sometimes unaware of the causes of their moral judgments, which may stem from System 1 rather than System 2. So too judges, and others involved in law, may be quite oblivious to the causes of the moral judgments that underlie their legal conclusions, and may sincerely but mistakenly believe that their ex post explanations were causal.

The intuitions described here play an important role in multiple domains, including families, labor unions, workplaces, student groups, sporting events, and religious organizations. But as many of the examples suggest, they also influence the decisions of legal and political institutions. Such institutions are usually intended to be deliberative, to override error-prone intuitions, and to pay close attention to System 2; but even in the most deliberative institutions, System 1 can make some compelling demands.