

Hazardous Heuristics

Cass R. Sunstein†

Heuristics and Biases: The Psychology of Intuitive Judgement, Thomas Gilovich, Dale W. Griffin, & Daniel Kahneman, eds. Cambridge, 2002. Pp xvi, 857.

In the early 1970s, Daniel Kahneman and Amos Tversky produced a series of pathbreaking papers about decisions under uncertainty.¹ Their claim was that in assessing probabilities, “people rely on a limited number of heuristic principles which reduce the complex tasks of assessing probabilities and predicting values to simpler judgmental operations.”² Kahneman and Tversky did not argue that it is irrational for people to use the relevant heuristics. On the contrary, they claimed that as a general rule, the heuristics are quite valuable. The problem is that in some cases, they lead “to severe and systematic errors.”³ It is worth emphasizing the word “systematic.” One of the most striking features of their argument was that the errors were not random—they could be described and even predicted.

The resulting arguments have proved influential in many fields, including law,⁴ where the influence stems from the effort to connect legal analysis to a realistic, rather than hypothetical, understanding of

† Karl N. Llewellyn Distinguished Service Professor of Jurisprudence, The University of Chicago. Thanks to Reid Hastie, Daniel Kahneman, Eric Posner, Richard Posner, Lior Strahilevitz, and Adrian Vermeule for valuable comments. Caryn Campbell provided excellent research assistance.

¹ The key papers can be found in Daniel Kahneman, Paul Slovic, and Amos Tversky, eds, *Judgment under Uncertainty: Heuristics and Biases* (Cambridge 1982) (presenting early studies about human judgment in the face of uncertainty). The heuristics-and-biases literature should be distinguished from the literature on prospect theory, which involves the nature of people’s utility functions under conditions of risk, not mental shortcuts under conditions of uncertainty. See Daniel Kahneman and Amos Tversky, *Choices, Values, and Frames* (Cambridge 2001) (presenting recent studies about prospect theory and related work).

² Amos Tversky and Daniel Kahneman, *Judgment under Uncertainty: Heuristics and Biases*, in Kahneman, Slovic, and Tversky, eds, *Judgment under Uncertainty* 3, 3 (cited in note 1) (setting forth the basic findings).

³ *Id.*

⁴ See, for example, Christine Jolls, Cass R. Sunstein, and Richard Thaler, *A Behavioral Approach to Law and Economics*, 50 *Stan L Rev* 1471, 1518–19 (1998) (exploring implications of judgment biases for demand for environmental regulation, especially Superfund); Roger G. Noll and James E. Krier, *Some Implications of Cognitive Psychology for Risk Regulation*, 19 *J Legal Stud* 747 (1990) (exploring implications of cognitive psychology for the regulation of health and environmental risks).

how human beings think and behave. If human beings use identifiable heuristics, and if they make systematic errors, we might better understand why law is as it is, and we might generate better strategies for ensuring that law actually promotes social goals. Most provocatively, an understanding of heuristics and biases should improve our understanding of the legitimate role of paternalism in law. If people make systematic errors, perhaps government has, more often than anti-paternalists think, good reason to override their choices. The influence of the heuristics-and-biases literature also stems from its obvious connection with particular problems with which lawyers and policymakers are concerned. For example, the system of risk regulation has been said to show a combination of “paranoia and neglect.”⁵ An understanding of systematic errors might help show how and why this is so, and give a sense of what might be done by way of response.

Kahneman and Tversky emphasized three general-purpose heuristics: representativeness, availability, and anchoring. The availability heuristic has probably become the most well-known in law.⁶ When people use this heuristic, they answer a question of probability by asking whether examples come readily to mind.⁷ How likely is a flood, an airplane crash, a traffic jam, a terrorist attack, or a disaster at a nuclear power plant? Lacking statistical knowledge, people try to think of illustrations. Thus, “a class whose instances are easily retrieved will appear more numerous than a class of equal frequency whose instances are less retrievable.”⁸ For people without statistical knowledge, it is far from irrational to use the availability heuristic; the problem is that this heuristic can lead to serious errors of fact, in the form of excessive fear of small risks and neglect of large ones.⁹

Kahneman and Tversky also suggested that in the face of uncertainty, estimates are often made from an initial value, or “anchor,” which is then adjusted to produce a final answer.¹⁰ The initial value seems to have undue influence. What percentage of African countries

⁵ John D. Graham, *Making Sense of Risk: An Agenda for Congress*, in Robert Hahn, ed., *Risks, Benefits, and Lives Saved: Getting Better Results from Regulation* 183, 183 (Oxford 1996) (proposing more balanced approaches toward risk regulation).

⁶ A Lexis search of law reviews found well over 200 references to the availability heuristic (as of February 2003).

⁷ See Tversky and Kahneman, *Judgment under Uncertainty* at 3, 11–14 (cited in note 2) (discussing the availability heuristic).

⁸ *Id.* at 11.

⁹ See Timur Kuran and Cass R. Sunstein, *Availability Cascades and Risk Regulation*, 51 *Stan L Rev* 683, 703–05 (1999) (exploring how the availability heuristic and other biases “fuel mass delusions that have large consequences for regulatory policy”); Noll and Krier, 19 *J Legal Stud* at 755 (cited in note 4) (discussing cognitive errors).

¹⁰ See Tversky and Kahneman, *Judgment under Uncertainty* at 14 (cited in note 2) (“In many situations, people make estimates by starting from an initial value that is adjusted to yield the final answer.”).

are in the United Nations? In one study, Kahneman and Tversky spun a wheel of fortune to obtain a number between 0 and 100, and asked subjects to say whether the number that emerged from the wheel was higher or lower than the relevant percentage.¹¹ It turned out that the starting point, though clearly random, greatly affected people's answers. If the starting point was 65, the median estimate was 45 percent; if the starting point was 10, the median estimate was 25 percent. The process of anchoring-and-adjustment has an obvious application to many legal issues, including the setting of damage awards, where anchors play a large role.¹²

When the representativeness heuristic is involved, people answer a question of probability or causation—for example, how likely is it that object A belongs to class B?—by asking about the extent to which A *resembles* B. Suppose, for example, that the question is whether some person, Nick, is a librarian or a farmer. If Nick is described as shy and withdrawn, and as having a passion for detail, most people will think that he is likely to be a librarian—and to ignore the “base-rate,” that is, the fact that the population has far more farmers than librarians. It should be readily apparent that the representativeness heuristic will produce problems whenever people are ignoring base-rates, as they are prone to do. In one study, a majority of judges, in assessing probabilities, fell prey to the representativeness heuristic.¹³

Since the early discussions of heuristics and biases, there has been an explosion of further work, sometimes contesting the basic claims of Kahneman and Tversky,¹⁴ but usually offering more applications, an improved understanding of how the heuristics work, and a discovery of many other heuristics and biases. *Heuristics and Biases: The Psychology of Intuitive Judgment* offers a massive, state-of-the-art treatment of the literature, supplementing a similar book published two decades ago.¹⁵ The book is divided into three parts. The first, called Theoretical and Empirical Extensions, elaborates on the three main heuristics and on several related heuristics and biases, including optimistic bias. The second part, called New Theoretical Directions, discusses the role of emotions and affect, support theory, and alternative perspectives on heuristics. This discussion includes the view, set forth most prominently by Gerd Gigerenzer, that outside the laboratory,

¹¹ Id.

¹² See Chris Guthrie, Jeffrey Rachlinski, and Andrew Wistrich, *Inside the Judicial Mind*, 86 Cornell L Rev 777 (2001) (showing how judges, like juries, fall victim to anchoring).

¹³ See id.

¹⁴ See Gerd Gigerenzer, Peter M. Todd, and the ABC Research Group, *Simple Heuristics That Make Us Smart* 27–28 (Oxford 1999) (contrasting the heuristics-and-biases notion of heuristics with a “fast and frugal” notion based on bounded rationality and adaptation).

¹⁵ See Kahneman, Slovic, and Tversky, eds, *Judgment under Uncertainty* (cited in note 1).

our “fast and frugal” heuristics work very well (p 559). The third part, called Real-World Applications, offers a range of cases in which intuitive judgments go wrong. These judgments include those by ordinary people, who falsely believe, for example, in the “hot hand” phenomenon in basketball (p 601), and those by experts, whose clinical judgments of dangerousness are far less accurate than actuarial judgments—a point with many legal applications.¹⁶

This is an impressive book, full of implications for law and policy. The collection also covers an extraordinary range of problems. I will not be able to come close to doing justice to it here. Instead I have a much narrower purpose: to connect some of the recent research with a set of legal problems, and in particular those relating to risk and litigation behavior. In that connection, two aspects of the book deserve emphasis. The first involves a shift from the strictly cognitive focus of the early work to an effort to see how emotions affect decision and judgment. The second is the emphasis in several of the papers on “dual process” approaches to human thinking. According to these approaches, people have two systems for making decisions. One of these is rapid, intuitive, but sometimes error-prone; the other is slower, reflective, and more statistical. One of the pervasive themes in this collection is that heuristics and biases can be connected with the intuitive system and that the slower, more reflective system might be able to make corrections.¹⁷ This emphasis on correction raises the possibility of “debiasing,” on which several of the papers also focus.

This Review has five parts. Part I discusses some real-world effects of availability and anchoring. Part II examines one of the most important and interesting papers in the book, in which Daniel Kahneman and Shane Frederick offer a rethinking and generalization of

¹⁶ See William Meadow and Cass R. Sunstein, *Statistics, Not Experts*, 51 Duke L J 629, 629–32 (2001) (arguing that statistical data is more accurate than expert reports).

¹⁷ The papers do not discuss the nature of the brain, but suggestive research tends to be supportive of the dual-process idea. Some research suggests that the brain has special sectors for emotions, and that some types of emotions, including some fear-type reactions, can be triggered before the more cognitive sectors become involved at all. See Joseph LeDoux, *The Emotional Brain: The Mysterious Underpinnings of Emotional Life* 157–69, 172–73, 283–96 (Simon & Schuster 1996). A small, almond-shaped region of the forebrain, the amygdala, appears to play a distinctive role in registering fear, with more reflective checks coming from the cerebral cortex. See *id.* at 172–73 (suggesting that stimulation of the amygdala produces “a sense of foreboding danger, of fear,” and that “studies of humans with amygdala damage also suggest that it plays a special role in fear”). Indeed, some “emotional responses can occur without the involvement of the higher processing systems of the brain, systems believed to be involved in thinking, reasoning, and consciousness.” *Id.* at 161. The thalamic pathway, involving the amygdala, “cannot make fine distinctions” but has a strong advantage in speed. *Id.* at 163. It “can provide a fast signal that warns that something dangerous may be there. It is a quick and dirty processing system.” *Id.* An especially interesting finding: A patient with amygdala damage was asked to detect emotional expression on faces, and she succeeded in identifying “most classes of expressions, except when the faces showed fear.” *Id.* at 173.

the whole idea of heuristics.¹⁸ Part III explores the role of emotions and affect. Part IV investigates optimistic bias. Part V goes beyond the book under review to offer some speculative remarks about the possibility of “moral heuristics”—mental shortcuts that generally work well, but that lead to systematic errors in thinking about morality and law.

I. PARLOR GAMES? THE REAL WORLD OF AVAILABILITY AND ANCHORING

The early work on heuristics and biases raised a natural set of doubts:¹⁹ Are people likely to make systematic errors? Are these phenomena important in the real world? Are heuristics an artifact of strange experiments? On one view, the mistakes, often made by undergraduate subjects, are a product of clever manipulations by psychologists, and in real-world environments, or in markets, people do much better.²⁰ These issues receive helpful attention in the introduction (pp 7–15) and elsewhere, but they are not the book’s explicit focus. To answer them, it is important to emphasize that the goal of the heuristics-and-biases literature is emphatically not to show that people are fools, or that they are systematically irrational. On the contrary, Kahneman and Tversky emphasized that the relevant heuristics are efficient and generally work well. But in the laboratory, at least, people who use the heuristics sometimes blunder, and it is the blundering that has attracted the most academic attention. Consider, for example, the fact that when asked how many words, on four pages of a novel, end in “ing,” people will give a larger number than when asked how many words have “n” as their second-to-last letter (p 21)—a clear

¹⁸ See Daniel Kahneman and Shane Frederick, *Representativeness Revisited: Attribute Substitution* (p 81) (generalizing the heuristics and biases approach to cognitive processes of attribute substitution, “not limited to questions about uncertain events”).

¹⁹ See Gigerenzer, Todd, and the ABC Research Group, *Simple Heuristics* at 27–28 (cited in note 14) (arguing that heuristics are adaptive); Richard A. Posner, *Rational Choice, Behavioral Economics, and the Law*, 50 Stan L Rev 1551, 1551–53 (1998) (arguing against abandonment of the rationality assumption). Except for a few brief notations, I do not deal here with the issues raised by Gigerenzer’s numerous criticisms of the heuristics-and-biases literature. I believe, in short, that Gigerenzer’s criticisms suggest that some heuristics work well (which no one should deny) and that problems can be designed so as to reduce or even eliminate errors (which no one should deny). For a helpful discussion of the last point, see Gerd Gigerenzer, *Calculated Risks: How to Know When Numbers Deceive You* 230 (Simon & Schuster 2002) (introducing a three-step educational program for teaching individuals “how to reckon with risk”).

²⁰ For evidence that heuristics and biases operate in the real world, even when dollars are involved, see Werner F.M. Debondt and Richard H. Thaler, *Do Analysts Overreact?* (pp 678–85) (finding systematic overreaction of security analysts); Robert Shiller, *Irrational Exuberance* 136–47 (Princeton 2000) (discussing anchoring and overconfidence in market behavior); Colin Camerer and Robin Hogarth, *The Effects of Financial Incentives in Experiments: A Review and Capital-Labor-Production Framework*, 19 J Risk & Uncertainty 7, 7 (1999) (finding that financial incentives have never eliminated anomalies or persistent irrationalities).

laboratory illustration of the availability heuristic. Several of the papers go well beyond the laboratory and demonstrate that heuristics produce errors in the real world.

A. Availability and Risk

1. Availability, health, and safety.

It is reasonable to expect that people's judgments about health and safety risks would be affected by the availability heuristic, and Baruch Fischhoff offers some striking illustrations (p 730). Should women offer physical resistance in cases of assault? In popular publications, experts offer contradictory advice (p 733). Those who claim that resistance is a serious mistake consist disproportionately of people from law enforcement sources, who mostly see bad outcomes in cases of physical resistance. Hence police officers may well be victims of the availability heuristic, at least "if they remembered what they had seen and heard, but lacked an appreciation of what they were not seeing" (p 733). (Fischhoff does not explore the possibility that those who resist and are hurt receive special media attention, producing an additional bias.) More generally, Fischhoff discusses lay estimates of the frequency of forty-one causes of death in the United States. He finds that the errors in these estimates are consistent with the availability heuristic (and hence the errors were predicted before the data were seen) (p 737). Highly publicized causes of death, such as floods and tornadoes, are overestimated, whereas quieter killers, such as strokes and diabetes, are underestimated (p 738). Other studies show a similar pattern.²¹

Apart from surveys, is actual behavior affected by the availability heuristic? There is evidence that it is. If floods have not occurred in the immediate past, people who live on flood plains are far less likely to purchase insurance.²² In the aftermath of an earthquake, the number of people seeking insurance for earthquakes rises sharply—but that number declines steadily from that point, as vivid memories re-

²¹ See W. Kip Viscusi, *Jurors, Judges, and the Mistreatment of Risk by the Courts*, 30 J Legal Stud 107 (2001) (finding similar misestimates). A possible criticism of these findings is that they might show the effect of anchoring. In the relevant surveys, people typically are given a starting number, such as the number of deaths from motor vehicle accidents each year (around 40,000). That starting number is necessary to ensure that numbers, for imperfectly informed respondents, will not be all over the lot. But the starting number, as an anchor, might also compress the range of answers, making high numbers lower and low numbers higher than they would otherwise be.

²² For a vivid demonstration in the context of catastrophes, see Jacob Gersen, *Strategy and Cognition: Regulatory Catastrophic Risk* 57–60, 77 (unpublished manuscript 2001) (presenting and analyzing evidence linking subjective beliefs about the likelihood of floods and their actual occurrence). See also Paul Slovic, et al, *Cognitive Processes and Societal Risk Taking*, in Paul Slovic, ed. *The Perception of Risk* 32, 39–40 (Earthscan 2000) (discussing how individuals' assessment of flooding probability is strongly conditioned by their recent experience).

cede.²³ Notice that the use of the availability heuristic in these contexts strongly suggests that the heuristics operate even when the stakes are large. And it is possible that the use of the availability heuristic in such contexts is fully rational for people who lack statistical knowledge. Perhaps use of that heuristic is the best way of minimizing the sum of decision costs and error costs.²⁴ But it seems less useful to debate the rationality of the availability heuristic than simply to observe that it has a significant effect on actual behavior.

2. The sources of availability.

What in particular produces availability? An interesting essay attempts to test the effects of ease of *imagery* on perceived judgments of risk (p 98). The study asked subjects to read about an illness (Hyposcenia-B) that “was becoming increasingly prevalent” (p 99) on the local campus. In one condition, the symptoms were concrete and easy to imagine—involving muscle aches, low energy, and frequent severe headaches. In another condition, the symptoms were vague and hard to imagine, involving an inflamed liver, a malfunctioning nervous system, and a vague sense of disorientation. Subjects in both conditions were asked both to imagine a three-week period in which they had the disease and to write a detailed description of what they imagined. After doing so, subjects were asked to assess, on a ten-point scale, their likelihood of contracting the disease. The basic finding was that likelihood judgments were very different in the two conditions, with easily-imagined symptoms making people far more inclined to believe that they were likely to get the disease.

There are several implications for policy and law. The public demand for law should be much higher if people can easily imagine the harm in question; in such cases, the law might well reflect a kind of hysteria.²⁵ But if the harm is difficult to imagine, we might well see a pattern of neglect.²⁶ We would therefore predict that easily imaginable

²³ Paul Slovic, et al, *Preference for Insuring Against Probable Small Losses: Insurance Implications*, in Slovic, ed, *The Perception of Risk* at 51 (cited in note 22) (discussing study of one-thousand homeowners in earthquake-prone areas and two-thousand homeowners in flood-prone areas that found that bias from the availability heuristic affected individual decisions to insure).

²⁴ It is reasonable, however, to read Kahneman and Tversky as suggesting that the heuristics cannot entirely be defended in this way—that some of the time, at least, the heuristics operate even though a little thought would improve judgments. Consider the discussion of “ing” as opposed to “n” as the next-to-last letter, above, and consider also the Linda problem, discussed below.

²⁵ I am emphasizing the availability heuristic here, but when people imagine bad events, they might also be vulnerable to “probability neglect,” in which they fail to consider the likelihood of harm. See Cass R. Sunstein, *Probability Neglect: Emotions, Worst Cases, and Law*, 112 *Yale L J* 61, 62 (2002) (illustrating how people “focus on the adverse outcome, not on its likelihood” and discussing possible policy implications).

²⁶ Compare the finding that teens’ rates of risk behaviors—smoking, driving after drinking,

harms would lead to relatively greater private precautions and relatively greater governmental concern.²⁷ Well-organized private groups should, and do, take advantage of this heuristic, attempting to publicize visible examples of harms to which they seek to draw attention.²⁸ The point also offers implications about public informational campaigns. If the government wants people to take protective steps, it should provide information about symptoms in a vivid rather than statistical way (p 102), relying on examples that can later be brought to mind. (Terrorists show a good intuitive understanding of the availability heuristic, simply because a highly publicized terrorist act is likely to create far more fear than the bare statistics warrant; consider the September 11 attack and the acts of the Washington, DC snipers in the fall of 2002.²⁹)

But there is an interesting puzzle for those interested in the real-world uses of this heuristic: In many contexts, multiple images are literally “available.” Consider the problem of gun violence. It is not hard to find cases in which the presence of guns led to many deaths, and also cases in which the presence of guns allowed law-abiding citizens to protect themselves against criminals.³⁰ In the face of conflicting instances, which cases are especially available? Why should one or another kind of case be available? The behavior of the media, and of relevant interest groups, is undoubtedly important here. Many perceived “epidemics” are in reality no such thing, but instead a product

unsafe sex—can be reduced by addressing heuristics and biases, in part by explaining that the availability heuristic leads teens to overestimate the risk behavior of their peers. See Baruch Fischhoff, *Heuristics and Biases in Application* (p 747) (arguing that better education about biases leads teens to reduce undesirably risky behaviors).

²⁷ See Noll and Krier, 19 J Legal Stud at 767–71 (cited in note 4) (analyzing the correlation of availability with the political demands for risk policy).

²⁸ See Kuran and Sunstein, 51 Stan L Rev at 715–35 (cited in note 9) (discussing availability campaigns).

²⁹ Although only ten people were killed by the snipers, extensive and even extraordinary precautions were taken. For example, October SAT testing at several area schools was canceled. See Erik Brady, *Weekend Plans Fall Victim to Sniper Fears*, USA Today A01 (Oct 14, 2002) (reporting on the extensive precautions taken in the wake of the Washington DC, Virginia, and Maryland sniper attacks). The size of the panic is not easy to explain on the numbers alone. The statistical risk of being killed by the snipers was, at every stage, quite low. I believe that the extreme fear was produced by a combination of the availability heuristic and probability neglect. Compare the fact that, on average, twenty-five people die in auto accidents during a two-week time span in an area the size of Washington DC. Such deaths hardly cause a panic. See Mark Memmott, *Fear May Be Overwhelming, But So Are the Odds*, USA Today A06 (Oct 18, 2002) (drawing several statistical comparisons with the sniper attacks in order to highlight the public's overreaction).

³⁰ See Donald Braman and Dan M. Kahan, *More Statistics, Less Persuasion: A Cultural Theory of Gun-Risk Perceptions* 26–27, 26 n 63, working paper (2002), online at <http://www.law.uchicago.edu/academics/circulation5.pdf> (visited Feb 18, 2003) (discussing the cultural determinants of gun attitudes given the ready availability of competing cases of offensive or defensive uses of guns).

of media coverage of gripping, unrepresentative incidents.³¹ But this does not provide the whole picture. Beliefs and orientations are a product of availability, to be sure; but what is available is also a product of antecedent beliefs and orientations. In other words, availability may be endogenous to individual predispositions.

Social processes are quite important here, for apparently representative anecdotes and gripping examples can move rapidly from one person to another.³² Once several people start to take an example as probative, many people may come to be influenced by their opinion, giving rise to cascade effects.³³ In the domain of risks, “availability cascades” help to account for many social beliefs, and here local variations are likely, with different examples becoming salient in different communities. Indeed, processes of deliberation typically lead like-minded people to accept a more extreme version of their original views,³⁴ making it likely that the effects of certain available examples will become greatly amplified through group discussion. The problem might well be aggravated by certain media and new technologies. Because of the internet, incidents can be made widely available almost instantly, creating “availability cybercascades” in which misleading or false information is rapidly spread.³⁵ With the same technologies, group discussion can amplify these processes of fear-mongering.³⁶ And undoubtedly different cultural orientations play a large role in deter-

³¹ See Michelle Cottle, *Summer Scare*, The New Republic Online (July 31, 2002), online at <http://www.tnr.com/docprint.mhtml?i=life&s=cottle073102> (visited Nov 15, 2002) (“[T]ake these much-hyped abductions, add in the half dozen other cases mentioned by the national media since the first of the year . . . [it] still doesn’t qualify as a new crime wave.”).

³² See Chip Heath, Chris Bell, and Emily Sternberg, *Emotional Selection in Memes: The Case of Urban Legends*, 81 *J Personality & Soc Psych* 1028, 1028 (2001) (discussing three studies finding a more ready willingness to pass along stories that evoked relatively stronger emotional responses, especially disgust); Chip Heath, *Do People Prefer to Pass Along Good or Bad News? Valence and Relevance as Predictors of Transmission Propensity*, 68 *Org Beh & Human Dec Processes* 79, 89 (1996) (finding that whether news is good, bad, or neither, individuals will propagate it rapidly so long as it remains within a congruent domain of social action).

³³ See Shiller, *Irrational Exuberance* at 148–68 (cited in note 20) (discussing social pressures leading to information cascades); Kuran and Sunstein, 51 *Stan L Rev* at 715–36 (cited in note 9) (proposing a mechanism for rapid, self-reinforcing formation of beliefs); Sushil Biikhchandani, et al, *Learning from the Behavior of Others: Conformity, Fads, and Informational Cascades*, 12 *J Econ Persp* 151, 168 (Summer 1998) (“[I]nformational cascades theory . . . implies pervasive but fragile herd behavior. This occurs because cascades are triggered by a small amount of information.”).

³⁴ See Cass R. Sunstein, *Conformity and Dissent* 8, University of Chicago Law & Economics Olin Working Paper No 164 (2002), online at <http://papers.ssrn.com/id=341880> (visited Feb 18, 2003) (illustrating and explaining how individual members’ suppression of more modest views produces group deliberative bias toward extremism, especially for groups of like-minded members); Cass R. Sunstein, *Deliberative Trouble? Why Groups Go To Extremes*, 110 *Yale L J* 71, 88–94 (2000) (explaining how deliberation may polarize the views of groups if their members are composed of primarily like-minded individuals).

³⁵ See Cass R. Sunstein, *Republic.com* 80–84 (Princeton 2001).

³⁶ See *id.* at 64–69.

mining what turns out to be available.³⁷ A great deal of work, normative and empirical, remains to be done on this topic.

B. Anchors and Damages

The original studies of anchoring-and-adjustment were memorable in part because they were so amusing. They suggested that when people lack information about an appropriate value, they are highly suggestible, even by apparently irrelevant numbers.³⁸ But the original studies left open many questions about the necessary conditions for anchoring, and also about the role of anchoring outside of the laboratory.³⁹

Gretchen Chapman and Eric Johnson offer a great deal of help in answering these questions (p 120). Chapman and Johnson show that an anchor is often operating even when people think that it is not (*id*); that anchors have effects even when people believe, and say they believe, that the anchor is uninformative (*id*); and that making people aware of an anchor's effect does not reduce anchoring (*id*). Very extreme or ludicrously implausible anchors also seem to have an effect: Estimates of the year that Albert Einstein first visited the United States are greatly affected by asking people to begin by considering anchors of 1215 or 1992 (p 124). Chapman and Johnson also show that economic incentives do not eliminate the effects of anchors (p 125); anchoring is not a result of casualness about the underlying task. It follows that "debiassing" is very difficult in this context.

Anchors have major effects on legal outcomes. The plaintiff's demand influences jury verdicts, in terms of both liability judgments and amounts awarded (p 137). Even implausibly low and implausibly high demands operate as anchors (*id*). Opening offers in negotiation have a significant influence on settlements. An ingenious study finds that anchors affect judges, too. Judges were asked to come up with appropriate awards in a personal injury case.⁴⁰ The study involved two conditions. The "no anchor" condition involved a simple statement of the facts. The "anchor" condition was the same as the first, but with one critical difference: The defendant filed an obviously meritless motion to dismiss the case on the ground that the \$75,000 jurisdictional

³⁷ See Braman and Kahan, *More Statistics, Less Persuasion* at 1, 18 (cited in note 30) (setting forth the cultural determinants of heuristics about gun ownership).

³⁸ See Tversky and Kahneman, *Judgment under Uncertainty* at 14–16 (cited in note 2) (discussing studies showing anchoring effect prevents proper scaling adjustment).

³⁹ For a discussion of market behavior and anchoring, see Shiller, *Irrational Exuberance* at 135–42 (cited in note 20) (discussing quantitative anchors and moral anchors in the context of the stock market).

⁴⁰ See Guthrie, Rachlinski, and Wistrich, 6 *Cornell L Rev* at 790–91 (cited in note 12) (presenting judges with a description of a serious personal injury suit with or without a damage award anchor).

minimum had not been met. Almost all of the judges denied the motion, which nonetheless served as an anchor, with large effects on ultimate judgments. In the no-anchor condition, the average award was \$1.24 million, while it was \$882,000 in the anchor condition.⁴¹

Anchors also play a role in “contingent valuation” studies, an influential method of valuing regulatory goods, such as increased safety and environmental protection (p 137).⁴² Perhaps the most striking, and in a way hilarious, evidence to this effect comes from a study of willingness to pay to reduce annual risks of death and injury in motor vehicles.⁴³ The authors attempted to elicit both maximum and minimum willingness to pay for safety improvements. People were presented with a risk and an initial amount, and asked whether they were definitely willing or definitely unwilling to pay that amount to eliminate the risk, or “not sure.” If they were definitely willing, the amount displayed was increased until they said that they were definitely unwilling; if they were unsure, the number was moved up and down until people could identify the minimum and maximum.

The authors were not attempting to test the effects of anchors. On the contrary, they were alert to anchoring only because they “had been warned” of a possible problem with their procedure, in which people “might be unduly influenced by the first amount of money that they saw displayed.”⁴⁴ To solve that problem, the authors allocated people randomly to two subsamples, one with an initial display of £25, the other with an initial display of £75. The authors hoped that the anchoring effect would be small, with no significant consequences for minimum and maximum values. But their hope was dashed. For every level of risk, the minimum willingness to pay was higher with the £75 starting point, than the maximum willingness to pay with the £25 starting point!⁴⁵ For example, a reduction in the annual risk of death by 4 in 100,000 produced a *maximum* willingness to pay of £149 with the £25

⁴¹ Id at 791. There is a possible response to the authors’ claim to have shown the effects of anchoring: Perhaps the motion to dismiss suggested that the injury was less serious than was apparent. Why would a lawyer file a totally frivolous motion to dismiss? But the abundant evidence of effects from anchors suggests that this is unlikely to explain all or even much of the authors’ finding. See also W. Kip Viscusi, *Corporate Risk Analysis: A Reckless Act?*, 52 Stan L Rev 547, 558 (2000) (finding an anchoring effect from monetary value of life on jury awards, so much so that companies that placed a high value on human life ended up paying higher punitive awards).

⁴² See, for example, Ian J. Bateman and Kenneth G. Willis, eds, *Valuing Environmental Preferences: Theory and Practice of the Contingent Valuation Method in the US, EU, and Developing Countries* (Oxford 1999) (presenting several studies of contingent valuation).

⁴³ See Michael Jones-Lee and Graham Loomes, *Private Values and Public Policy*, in Elke U. Weber, et al, eds, *Conflict and Tradeoffs in Decision Making* 205, 210–12 (Cambridge 2000) (presenting data for strong starting point effects on individual’s willingness to pay for reducing auto accident injuries and deaths).

⁴⁴ Id at 210.

⁴⁵ Id at 211.

starting value, but a *minimum* willingness to pay of £232, with the £75 starting value (and a maximum, in that case, of £350).⁴⁶

The most sensible conclusion is that whenever people are uncertain about appropriate values, anchors have a significant effect, and sometimes a startlingly large one. Clever negotiators, lawyers, and policymakers should be able to exploit those effects, sometimes even by providing an outlandish or apparently irrelevant anchor. There is a real opportunity for legal reform here, in part because anchors might well produce results that are not easy to defend, and in part because different anchors will ensure that similarly situated people are not treated similarly. Perhaps lawyers should not be permitted to inform jurors of potentially effective anchors, such as the annual profits of the firm or even the plaintiff's demand, at least in cases involving punitive awards or hard-to-monetize compensatory awards. Or perhaps judges should be asked to review jury awards carefully and by reference to comparison cases, so as to weaken the effect of arbitrary anchors.

This point raises a related one: Are groups able to avoid the judgment errors made by individuals? The evidence is mixed.⁴⁷ In general, groups tend to polarize: They tend to end up in a more extreme position in line with their predeliberation tendencies.⁴⁸ At the same time, groups have been found to make better decisions than individuals with respect to certain statistical problems.⁴⁹ There is some evidence that groups are slightly better at avoiding the problems created by use of the availability heuristic.⁵⁰ On the other hand, some evidence suggests that the use of the representativeness heuristic is actually amplified in groups.⁵¹ It seems clear that group processes do not eliminate the use of heuristics, and it remains to be seen whether and when they reduce or increase the resulting errors.

⁴⁶ *Id.*

⁴⁷ See Norbert L. Kerr, Robert J. MacCoun, and Geoffrey P. Kramer, *Bias in Judgment: Comparing Individuals and Groups*, 103 *Psych Rev* 687, 688–93 (1996) (reviewing several studies of different types of group and individual judgment errors and finding no clear pattern).

⁴⁸ See Sunstein, 110 *Yale L J* at 88–94 (2000) (cited in note 34) (finding that homogeneity, a common characteristic of groups, reinforces group members' tendencies to go to extremes).

⁴⁹ See Alan S. Blinder and John Morgan, *Are Two Heads Better Than One?: An Experimental Analysis of Group vs. Individual Decisionmaking* 1, NBER Working Paper No 7909 (2000), online at <http://www.nber.org/papers/w7909.pdf> (visited Feb 18, 2003) (finding that groups responded as rapidly and more effectively than individuals to a statistical urn problem and a monetary policy experiment).

⁵⁰ Kerr, MacCoun, and Kramer, 103 *Psych Rev* at 692 (cited in note 47), citing M.F. Stasson, et al, *Group Consensus Processes on Cognitive Bias Tasks: A Social Decision Scheme Approach*, 30 *Japanese Psych Rsrch* 68 (1988).

⁵¹ See *id.*, citing G. Stasser and W. Titus, *Effects of Information Load and Percentage Shared Information on the Dissemination of Unshared Information during Discussion*, 53 *J Personality & Soc Psych* 81 (1987).

II. TWO SYSTEMS

What, exactly, is a heuristic? When will a heuristic be overridden by cognitive processes that produce a more accurate understanding of the problem in question? In a highlight of this collection, Daniel Kahneman and Shane Frederick make real progress on these questions (p 49). Their discussion is packed with new material, and I touch here only on the points of particular relevance for policy and law.

A. Dual Processing and Attribute Substitution

Much of their argument turns on drawing a connection between heuristics and dual-process theories.⁵² Recall that those theories distinguish between two families of cognitive operations, sometimes labeled System I and System II. System I is intuitive; it is rapid, automatic, and effortless. System II, by contrast, is reflective; it is slower, self-aware, and deductive. Kahneman and Frederick are careful to disclaim the view that the two systems operate as “autonomous homunculi”; in their view, they represent “collections of processes that are distinguished by their speed, controllability, and the contents on which they operate” (p 51). They suggest that System I proposes quick answers to problems of judgment, and that System II operates as a monitor, confirming or overriding those judgments. Consider, for example, someone who is flying from Chicago to New York in the month after an airplane crash. This person might make a rapid, barely conscious judgment, rooted in System I, that the flight is quite risky, but there might well be a System II override, bringing a more realistic assessment to bear. In making a distinction between System I and System II, Kahneman and Frederick announce a theme that plays a significant role in this book.⁵³

Kahneman and Frederick also offer a general claim about the nature of heuristics: They operate through a process of *attribute substitution* (p 53). In this process, people are interested in assessing a “target attribute,” and they do so by substituting a “heuristic attribute” of the object, which is easier to handle. Consider the question whether more people die from suicides or homicides. Lacking statistical information, people might respond by asking whether it is easier to recall cases in

⁵² See generally Shelly Chaiken and Yaacov Trope, eds, *Dual-Process Theories in Social Psychology* (Guilford 1999) (reviewing role of dual-process theories in social cognition and individual judgment).

⁵³ See, for example, Steven Sloman, *Two Systems of Reasoning* (p 379) (reviewing evidence for an “associative system” and a “rule-based system” of computation); Paul Slovic, et al, *The Affect Heuristic* (p 416) (advancing the distinction between a heuristic that orders reactions into “good” and “bad” and deliberate behavior meant to manipulate this effect); Robyn Dawes, et al, *Clinical versus Actuarial Judgment* (p 716) (distinguishing between error-prone clinical evaluation and data-based actuarial judgment).

either class (the availability heuristic). The response might well be sensible, but it might also lead to errors.

B. Amending the Theory

With these understandings, Kahneman and Frederick offer some significant amendments to the original presentation by Kahneman and Tversky. They suggest that anchoring should not be seen as a heuristic at all; anchoring operates not by substituting an attribute, but by making a particular value seem more plausible (p 56). They also argue that the third general-purpose heuristic, to replace anchoring, is the *affect heuristic* (id), discussed below. Kahneman and Frederick urge that punitive damage awards are mediated by an *outrage heuristic* (p 63), which we might see as an example of the affect heuristic in action. Jurors do not have a good sense of how to set punitive damage awards (a hard question), and they begin the process by asking about the outrageousness of the defendant's conduct (an easier question).⁵⁴ Something like an outrage heuristic undoubtedly plays a role in punishment judgments of many different kinds; there is a large research agenda here.

Now turn to the authors' focus, the representativeness heuristic, which has led to some large controversies.⁵⁵ The most famous of these involves questions about the likely career of a hypothetical woman named Linda (p 62), described as follows: "Linda is 31 years old, single, outspoken, and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice and also participated in antinuclear demonstrations." Subjects were asked to rank, in order of probability, eight possible futures for Linda. Six of these were fillers (like psychiatric social worker or elementary school teacher); the two crucial ones were "bank teller" and "bank teller and active in the feminist movement." Most people said that Linda was less likely to be a bank teller than to be a bank teller and active in the feminist movement. This is an obvious logical mistake, called a *conjunction error*, in which characteristics A and B are thought to be more likely than characteristic A alone. The error

⁵⁴ Here, Kahneman and Frederick draw on work in which I have been involved. See, for example, Cass R. Sunstein, et al, *Predictably Incoherent Judgments*, 54 Stan L Rev 1153, 1157–59, 1167–70 (2002) (discussing how juries move from a determination of punitive intent to a metric for damages or punishment).

⁵⁵ For a treatment of the representativeness heuristic and investment behavior, see Schiller, *Irrational Exuberance* at 144 (cited in note 20) (reviewing models of expectational feedback in the stock market). For some of the controversy, see Barbara Mellers, Ralph Hertwig, and Daniel Kahneman, *Do Frequency Representations Eliminate Conjunction Effects?*, 12 Psych Sci 269 (2001) (presenting disagreement over whether testing anomalies explain Kahneman and Tversky's finding of conjunction effects and whether frequency formats would eliminate such effects).

stems from the representativeness heuristic: Linda's description seems to match "bank teller and active in the feminist movement" far better than "bank teller."

As Kahneman and Frederick note, people's answers to the Linda problem have been explained on numerous grounds (p 67), with critics arguing that the structure of the problem increased or perhaps even generated logical mistakes. Kahneman and Frederick urge that this point should be taken not as a challenge to the claim that people use the representativeness heuristic, but as evidence that under certain circumstances, people will overcome the errors produced by that heuristic (including the conjunction fallacy and neglect of base-rates). Kahneman and Frederick suggest that when these problems are overcome, it is often because of the operations of System II, which works as a kind of supervisor. Hence intelligent people, and those with statistical sophistication, are less likely to err (p 68); for such people, System II is especially active.⁵⁶

For law and policy, the general lesson is simple: Whenever possible, *institutionalize System II*, at least when questions of fact are involved. Frequently the legal system disregards this advice, relying on juries and hence on ordinary intuitions about probability and causation.⁵⁷ The twentieth-century movement toward greater reliance on technical expertise and actual data might well be seen as an implicit recognition of the unreliability of ordinary intuitions. Indeed, there is reason to think that experts themselves are vulnerable to heuristics, and that reliance on actuarial data could improve accuracy.⁵⁸ In the domain of regulation, quantitative risk analysis is the most straightforward way of overcoming the errors that sometimes accompany heuristics. Consider the controversy over regulation of arsenic in drinking water.⁵⁹ The availability and representativeness heuristics ensure that many people will be quite frightened of arsenic, even in extremely low doses. Quantitative risk analysis can work as a kind of System II check on potential errors.

⁵⁶ With respect to intelligence, there is a nice qualification: When the problem is very hard for everyone, intelligent respondents are most likely to err, because they "are more likely to agree on a plausible error than to respond randomly" (p 68).

⁵⁷ See Gigerenzer, *Calculated Risks* at 85–86, 139, 159, 229–46 (cited in note 19) (offering several suggestions for educating doctors, administrators, judges, law students, and others about avoiding errors in risk assessment).

⁵⁸ See Robyn Dawes, et al, *Clinical versus Actuarial Judgment* (p 716) (distinguishing between error-prone clinical evaluation and data-based actuarial judgment).

⁵⁹ See Cass R. Sunstein, *The Arithmetic of Arsenic*, 90 *Georgetown L J* 2255, 2255–60 (2002) (arguing that scientific data only produce a wide "benefit range" of possible arsenic levels within which the EPA must set its standard).

C. Generalizing Representativeness

Kahneman and Frederick suggest that judgment heuristics, understood to involve attribute substitution, operate not only to answer questions about uncertain events, but also in a diverse class of judgments. Kahneman and Frederick go so far as to urge that a modest generalization of the representativeness heuristic helps to explain strikingly similar biases in economic valuations of public goods and retrospective evaluations of past events. In particular, Kahneman and Tversky emphasize the crucial role of *prototypes*, or representative exemplars, in making complex judgments.

How much are people willing to pay to save animals? It turns out that people are highly sensitive to the prototypes involved and highly insensitive to the number of animals at stake. A program that involves members of a popular species will produce a much higher willingness to pay than a program that involves members of a less popular species (p 71). More strikingly, people's willingness to pay does not differ greatly with large variations in the numbers involved; their willingness to pay is about the same to save 2,000, 20,000, or 200,000 birds (p 75). There is a clear parallel here to people's neglect of base-rates in using the representativeness heuristic to make probability judgments. Kahneman and Frederick also show that in evaluating past experiences, such as exposure to unpleasant noises, painful medical procedures, or horrific film clips, people show *duration neglect* (p 77). In one experiment, for example, people's evaluations of horrific movies were largely unaffected by substantial variations in their length (id). In another experiment, people's evaluations of colonoscopies were greatly influenced by the highest level of pain involved and also by the level of pain at the end, but not much by significant variations in the duration of the procedure (from four to sixty-nine minutes). Here too, Kahneman and Frederick urge that the prototype, captured in the Peak Affect and the End Affect, dominates evaluation.

With respect to law and policy, an intriguing implication here is that people's use of prototypes will crowd out variables that, on reflection, have clear importance. There is a serious problem with contingent valuation studies if people's judgments do not attend to the number of animals at stake. And indeed, some of the pathologies in regulatory policy do seem connected with this problem. Evidence suggests, for example, that people "worry more about the proportion of risk reduced than about the number of people helped."⁶⁰ A striking study of this effect finds that people pervasively neglect absolute

⁶⁰ Jonathan Baron, *Thinking and Deciding* 500 (Cambridge 3d ed 2000) (explaining the generally-observable confusion between relative and absolute risks).

numbers, and that this neglect maps onto regulatory policy.⁶¹ In a similar vein, it has been shown that when emotions are involved, people neglect two numbers that should plainly be relevant: the probability of harm and the extent of harm.⁶²

III. EMOTIONS, CONTAGION, AND AFFECT

How are judgments, especially judgments about the likelihood of risk or benefit, influenced by emotions and affect? Several papers explore that question.

A. False Contagions and Phony Cures

Paul Rozin and Carol Nemeroff explore “sympathetic magical thinking,” including the beliefs that some objects have contagious properties, and that causes resemble their effects.⁶³ Many educated Americans will not eat food touched by a sterilized cockroach (p 202). They refuse chocolates that have been shaped into realistic-looking dog feces (id). They are reluctant to use sugar from a bottle labeled “Sodium Cyanide, Poison,” even if they are assured, and believe, that the bottle really contains sugar and never contained cyanide (id)—and indeed even if they themselves placed the label, arbitrarily, on that particular bottle (p 205)! In fact people are reluctant to eat sugar labeled, “Not Sodium Cyanide,” apparently because the very words “Sodium Cyanide” automatically bring up negative associations. People are reluctant to wear a sweater that has been worn for five minutes by a person with AIDS (p 207). In this case, as in other cases involving contagion, people are relatively insensitive to dose. A sweater worn for five minutes by someone with AIDS, and then washed, is not much more undesirable than a sweater used by someone with AIDS for a full year. According to most respondents, a single live AIDS virus that enters the human body is as likely to infect someone with the vi-

⁶¹ See Timothy L. McDaniels, *Comparing Expressed and Revealed Preferences for Risk Reduction: Different Hazards and Question Frames*, 8 *Risk Anal* 593, 602–03 (1988) (finding that both ordinary people and policymakers frame valuations as percentage changes from some fixed rate of deaths).

⁶² See Sunstein, 112 *Yale L J* at 70–83 (cited in note 25) (discussing neglect of probability when strong emotions are involved); Christopher K. Hsee and Yuval Rottenstreich, *MUSIC, Pandas, and Muggers: On the Affective Psychology of Value*, working paper (2002) (on file with author) (suggesting that the dualism between valuation by feeling and by calculation produces a non-linear probability weighting of value); Yuval Rottenstreich and Christopher K. Hsee, *Money, Kisses, and Electric Shocks: On the Affective Psychology of Risk*, 12 *Psych Sci* 185, 186–88 (2001) (finding that in three separate studies undergraduates fail adequately to account for the extent and frequency of harm).

⁶³ Paul Rozin and Carol Nemeroff, *Sympathetic Magical Thinking: The Contagion and Similarity “Heuristics”* (p 201) (reviewing laws of contagion, similarity, and opposites in the way individuals think).

rus as 10,000 or even 1,000,000 viruses (p 207). Note in this regard that disgust and fear tend to “travel”; in both experimental and real-world settings, people are especially likely to spread “urban legends” that involve risks of contamination.⁶⁴

In some of these cases, the intuitive fear or revulsion can be easily overridden, as reflection reveals that there is no real hazard. System I gives rise to an immediate sense of alarm or revulsion, but System II will usually provide a corrective (even if System I continues to squawk). But not always. Paul Slovic has found that most people accept a kind of “intuitive toxicology,” showing agreement with the suggestion that “there is no safe level of exposure to a cancer-causing agent” and that “if you are exposed to a carcinogen, then you are likely to get cancer.”⁶⁵ Apparently some intuitions about fear are part of everyday thinking about social risks.

Thomas Gilovich and Kenneth Savitsky use the idea that “like goes with like” to unpack the structure of a wide range of false beliefs, both ancient and modern.⁶⁶ Many primitive beliefs about medicine reflect the belief that the symptoms of a disease are likely to resemble both its cause and its cure. According to ancient Chinese medicine, those with vision problems should eat ground bats, on the theory that bats have especially good vision, which might be transferred to people (p 619). Homeopathy, which remains quite popular, depends in part on the idea that if a substance creates disease symptoms in a healthy person, it will have a healthy effect on someone who currently suffers from those symptoms (p 620). The idea has some valid applications, but often the symptoms of a disease do not resemble its cause or its cure; consider sanitation and antibiotics (p 620).

I speculate that the immense popularity of organic foods owes a great deal to heuristic-driven thinking, above all to the view that there is an association between the natural and the healthy, and between chemical and danger.⁶⁷ To the extent that people trust scientifically du-

⁶⁴ See Heath, Bell, and Sternberg, 81 *J Personality & Soc Psych* at 1032–39 (cited in note 32) (presenting three studies where stories with high disgust factors tended to spread faster than others).

⁶⁵ Nancy Kraus, Torbjörn Malmfors, and Paul Slovic, *Intuitive Toxicology: Expert and Lay Judgments of Chemical Risks*, in Slovic, ed, *The Perception of Risk* 285, 290–91 (cited in note 22).

⁶⁶ Thomas Gilovich and Kenneth Savitsky, *Like Goes with Like: The Role of Representativeness in Erroneous and Pseudo-Scientific Beliefs* (p 617) (reviewing representative heuristic-related errors in medical beliefs, astrology, graphology, and psychoanalysis). Some of these themes are illuminatingly addressed in Thomas Gilovich, *How We Know What Isn't So: The Fallibility of Human Reason in Everyday Life* (Free Press 1995) (positing that the representative heuristic, the clustering illusion, and erroneous perception of random dispersions account for everyday intuitive fallacies).

⁶⁷ For criticism of that association, see James P. Collman, *Naturally Dangerous: Surprising Facts about Food, Health, and the Environment* (University Science 2001) (debunking the belief that natural and organic qualities correspond to safety and wholesomeness).

bious cures and treatments, it is often because they are neglecting base-rates, making selective use of the availability heuristic, and misperceiving the effects of randomness, which inevitably produces apparent patterns.⁶⁸ Here System I is the culprit, but it can be corrected by System II.

B. The Affect Heuristic

In emphasizing the affect heuristic, Kahneman and Tversky refer to the chapter of that title by Paul Slovic and several coauthors.⁶⁹ This chapter is one of the most interesting and suggestive in the volume. It also creates numerous puzzles, many of them involving law and policy.⁷⁰

People often have a rapid, largely affective response to objects and situations, including job applicants, consumer products, animals, cars, and causes of action. A jury might have an immediate negative reaction to a plaintiff in a personal injury case; a judge might have a positive intuitive reaction to an equal protection claim; an employer might instantly like, or dislike, someone who has applied for a job.⁷¹ But what does it mean to say that affect is a “heuristic”? Slovic, et al, urge that our affective responses occur rapidly and automatically, and that people use their feelings as a kind of substitute for a more systematic, all-things-considered judgment.⁷² It is in this sense that attribute substitution, as meant by Kahneman and Frederick, may be at work; affect toward an object substitutes for a more reflective assess-

⁶⁸ See Nassim Nicholas Taleb, *Fooled by Randomness: The Hidden Role of Chance in the Markets and in Life* (Texere 2001) (giving many entertaining examples of how random-generated effects are perceived as inevitable consequences); Gilovich, *How We Know* at 133–45 (cited in note 66) (debunking beliefs in holistic, new age, and alternative medicine).

⁶⁹ Paul Slovic, et al, *The Affect Heuristic* (p 397).

⁷⁰ I have elsewhere discussed an earlier and less elaborate version of Slovic’s work on affect, and I draw on that discussion here. See Cass R. Sunstein, *The Laws of Fear*, 115 Harv L Rev 1119 (2002) (reviewing Slovic, ed, *The Perception of Risk* (cited in note 22)).

⁷¹ See Timothy D. Wilson, David B. Centerbar, and Nancy Brekke, *Mental Contamination and the Debiasing Problem* (pp 198–99) (discussing mental contamination in employment decisions, especially as it relates to racial discrimination and Title VII recovery).

⁷² For the view that emotions are a form of cognition, see Martha Nussbaum, *Upheavals of Thought: The Intelligence of Emotions* 19 (Cambridge 2001) (arguing that emotions “involve judgments about important things, judgments in which, appraising an external object as salient for our own well-being, we acknowledge our own neediness and incompleteness before parts of the world that we do not fully control”). For the view that cognition plays a large role in producing emotions, see Jon Elster, *Alchemies of the Mind: Rationality and the Emotions* (Cambridge 1999). It is not clear that Slovic’s work is inconsistent with these views. Slovic does not deny that affect has cognitive antecedents, or even that affect is a form of cognition. It is clear, however, that the affect heuristic can lead to errors, as indeed can many emotional reactions, as Nussbaum agrees. See Nussbaum, *Upheavals of Thought* at 46–48, 51 (discussing in what sense emotional judgments and impressions may be false or mistaken).

ment of the object. Affect is an example of System I in operation—quick but error-prone.

But there is an obvious sense in which it is unhelpful to treat “affect” as an explanation for someone’s attitude toward objects. In some settings, affect *represents*, or *is*, that very attitude, and therefore cannot explain or account for it. (Would it be helpful to explain Tom’s romantic attraction to Anne by saying that Anne produces a favorable affect in Tom?) Slovic, et al, must be urging that sometimes affect works in the same way as availability and representativeness: In many contexts, people’s emotional reactions are substituting for a more careful inquiry into the (factual?) issues at stake.

The simplest way to establish this would be to proceed as Kahneman and Tversky originally did, by showing, for example, that people assess questions of probability by reference to affect, and that this method leads to predictable errors. What is the probability of death from smoking, driving, flying, or eating pesticides? If people’s affect toward these activities matched their probability judgments, producing systematic error, it would certainly be plausible to speak of an affect heuristic. Slovic, et al, do not have data of exactly this sort. But they do have some closely related evidence, suggestive of an affect heuristic in the domain of risk (pp 410–13). When asked to assess the risks and benefits associated with certain items, people tend to say that risky activities contain low benefits, and that beneficial activities contain low risks. It is rare that they will see an activity as *both* highly beneficial and quite dangerous, or as both benefit-free and danger-free. Because risk and benefit are distinct concepts, this finding seems to suggest that “affect” comes first, and helps to “direct” judgments of both risk and benefit.

Two studies fortify this hypothesis (pp 411–12). The first of these tests whether new information about the risks associated with some item alters people’s judgments about the benefits associated with the item—and whether new information about benefits alters people’s judgments about risks. The motivation for this study is simple. If people’s judgments were analytical and calculative, information about the great *benefits* of, say, food preservatives should not produce a judgment that the *risks* are low—just as information about the great *risks* of, say, natural gas should not make people think that the *benefits* are low. Strikingly, however, information about benefits alters judgments about risks, and information about risks alters judgments about benefits. When people learn about the low risks of an item, they are moved to think that the benefits are high—and when they learn about the high benefits of an item, they are moved to think that the risks are low. The conclusion is that people assess products and activities through affect—and that information that improves people’s affective re-

sponse will improve their judgments of all dimensions of those products and activities.

The second study asked people to make decisions under time pressure (p 412). The motivating claim is that the affect heuristic is more efficient than analytic processing in the sense that it permits especially rapid assessments. Under time pressure, Slovic, et al, hypothesize that there would be an unusually strong inverse correlation between judged risk and judged benefit, because affect will be the determinant of assessment, and people will have less time to undertake the kind of analysis that could begin to pull the two apart (p 412). In other words, System I is most important when time is scarce, and in such circumstances, System II will be a less effective monitor. The hypothesis is confirmed: Under time pressure, the inverse correlation is even stronger than without time pressure.

The affect heuristic casts a number of facts in a new light. Background mood, for example, influences decisions and reactions in many domains.⁷³ Consider the remarkable fact that stock prices increase significantly on sunny days, a fact that is hard to explain in terms that do not rely on affect.⁷⁴ Another study suggests that when people are anxious and fearful, they are less likely to engage in systematic processing, and hence System II is especially unreliable.⁷⁵ Note here that there is an evident relationship between social influences and the emotions: If emotions weaken systematic processing, they simultaneously increase susceptibility to the apparent views of others.⁷⁶ Fear itself is likely to make people susceptible to the acceptance of faulty logic and to pressure to conform.⁷⁷

The authors emphasize another point with important implications for risk regulation: When an outcome is accompanied by strong emotions, variations in probability have surprisingly little weight on people's decisions.⁷⁸ What matters are the images associated with the re-

⁷³ See Alice M. Isen, *Positive Affect and Decision Making*, in William M. Goldstein and Robin M. Hogarth, eds, *Research on Judgment and Decision Making: Currents, Connections, and Controversies* 509, 512 (Cambridge 1997) (exploring "the impact of mild positive affect on thinking and motivation").

⁷⁴ See David Hirshleifer and Tyler Shumway, *Good Day Sunshine: Stock Returns and the Weather* 19–21, 30, Dice Center Working Paper No 2001-3 (2001), online at <http://papers.ssrn.com/id=265674> (visited Feb 18, 2003) (finding a strong correlation between stock returns and morning sunshine for each of several exchanges in cities across the world).

⁷⁵ See Chaiken and Trope, *Dual-Process Theories* 19–20 (cited in note 52) (noting that "people who are anxious about or vulnerable to a health threat, or otherwise experiencing stress may engage in less careful or less extensive processing of health-relevant information").

⁷⁶ See Robert Baron, *Arousal, Capacity, and Intense Indoctrination*, 4 *Personality & Soc Psych Rev* 238, 244–46 (2000) (finding that intense emotional arousal leads to higher susceptibility to indoctrination).

⁷⁷ See *id* at 244.

⁷⁸ See Cass R. Sunstein, 112 *Yale L J* at 70–82 (cited in note 25) (finding a neglect of probability based on a review of experimental and real-world evidence).

sult. The point has received empirical confirmation in a study of people's willingness to pay to avoid electric shocks, or to be able to kiss favorite movie stars.⁷⁹ In one study, people's willingness to pay to avoid an electric shock varied little, depending on whether its probability was 1 percent or 99 percent!⁸⁰ With respect to hope, those who operate gambling casinos and state lotteries are well-aware of the underlying mechanisms. They play on people's emotions in the particular sense that they conjure up palpable pictures of victory and easy living. With respect to risks, insurance companies and environmental groups do exactly the same.

It follows that if government is seeking to encourage people to avoid large risks, and to worry less over small risks, it might well attempt to appeal to their emotions, perhaps by emphasizing the worst-case scenario. It should be no surprise that some of the most effective efforts to control cigarette smoking appeal to people's emotions, by making them feel that if they smoke, they will be dupes of the tobacco companies or impose harms on innocent third parties.⁸¹ There is also an opportunity here to try to activate System II, by promoting critical scrutiny of reactions that are based on "affective ties" in cases in which people are neglecting serious risks or exaggerating them.

IV. ARE PEOPLE UNREALISTICALLY OPTIMISTIC?

With respect to most of the risks of life, people appear to be unrealistically optimistic.⁸² This claim is closely related to the suggestion, with prominent advocates in economics, that people may attempt to reduce cognitive dissonance by thinking that the risks they face are lower than they are in fact.⁸³ If people systematically understate risks, there is a serious problem for law and policy, and a serious problem

⁷⁹ Rottenstreich and Hsee, 12 *Psych Sci* at 186–88 (cited in note 62) (finding that when strong emotions are present, individuals ignore important variations in probability).

⁸⁰ See *id.* at 188 (also reporting that there was a large spread, on the basis of probability, for the less "affect-rich" loss of \$20, where the median willingness to pay was \$1 for a 1 percent chance of loss and \$18 for a 99 percent chance of loss).

⁸¹ See Lisa K. Goldman and Stanton A. Glantz, *Evaluation of Antismoking Advertising Campaigns*, 279 *JAMA* 772 (1998) (finding that more aggressive anti-smoking advertisements that portray cigarette makers as manipulative and emphasize the dangers of second hand smoke are far more effective than ads emphasizing youth access, romantic rejection, and the known short and long term effects of smoking).

⁸² See Shelley E. Taylor, *Positive Illusions: Creative Self-Deception and the Healthy Mind* 6–11 (Basic 1989) (proposing that a healthy self-conception is biased toward optimism).

⁸³ See George A. Akerlof and William T. Dickens, *The Economic Consequences of Cognitive Dissonance*, in George A. Akerlof, ed., *An Economic Theorist's Book of Tales: Essays That Entertain the Consequences of New Assumptions in Economic Theory* 123–44 (Cambridge 1984) (advancing an economic model of cognitive dissonance with implications for standard economic puzzles such as the salience of noninformational advertising and the popularity of Social Security legislation).

too for those who accept the rational actor model in the social sciences.

A. Evidence

The most well-documented findings of optimism involve relative (as opposed to absolute) risk. About 90 percent of drivers think that they are safer than the average driver and less likely to be involved in a serious accident.⁸⁴ People generally think that they are less likely than other people to be divorced, to have heart disease, to be fired from a job, and much more.⁸⁵ At first glance, a belief in relative immunity from risk seems disturbing, but by itself this finding does not establish that people underestimate the risks that they actually face. Perhaps people have an accurate understanding of their own statistical risks even if they say and believe, wrongly, that other people are more vulnerable than they are.⁸⁶ With respect to absolute risk, the evidence for unrealistic optimism is less clear, as Daniel Armor and Shelley Taylor show in their contribution to this collection. For significant and personally relevant events, including unwanted pregnancy, people show an accurate understanding of their susceptibility (p 335). With respect to some low-probability events, including life-threatening risks such as AIDS, people actually tend to overestimate their own susceptibility, and in that sense seem to show pessimistic bias (id).⁸⁷ One survey finds general overestimates of personal risk levels for such hazards as breast cancer (where women rate their actual risk as 40 percent, with the actual risk being roughly 10 percent); prostate cancer (where men rank their actual risk as 40 percent, with the actual risk again being roughly 10 percent); lung cancer (estimated at 35 percent, compared to an actual risk of under 20 percent); and stroke (estimated at 45 percent, compared to an actual risk of roughly 20 percent).⁸⁸

⁸⁴ See Taylor, *Positive Illusions* at 10–11 (cited in note 82).

⁸⁵ See Neil D. Weinstein, *Unrealistic Optimism about Susceptibility to Health Problems*, 10 *J Behav Med* 481, 486 (1987) (listing results from survey obtaining comparative risk judgments from 297 individuals about a broad range of hazards).

⁸⁶ See W. Kip Viscusi, *Smoke-Filled Rooms: A Postmortem on the Tobacco Deal* 162–66 (Chicago 2002) (using survey data to show that smokers do not ignore risks to themselves as much as underestimate them in relation to other smokers).

⁸⁷ Armor and Taylor doubt this conclusion, suggesting that “these estimates may simply reflect difficulties interpreting and reporting extreme possibilities” (p 335).

⁸⁸ See Humphrey Taylor, *Perceptions of Risks: The Public Overestimates the Risks of Most Major Diseases and Types of Accidents—Breast and Prostate Cancer in Particular* (Jan 27, 1999), online at http://www.harrisinteractive.com/harris_poll/index.asp?PID=44 (visited Feb 18, 2003) (reporting survey data showing that the public overestimates its susceptibility to hazards salient in the media).

But in some domains, people do underestimate their statistical risk. For example, professional financial experts consistently overestimate likely earnings, and business school students overestimate their likely starting salaries and the number of offers that they will receive (pp 334–35). People also underestimate their likelihood of being involved in a serious automobile accident,⁸⁹ and their own failure to buy insurance for floods and earthquakes is at least consistent with the view that people are excessively optimistic.⁹⁰ The evidence of optimistic bias, both relative and absolute, is sufficient to raise questions about informational and regulatory interventions.

B. Debiasing?

Neil Weinstein and William Klein explore a variety of apparently promising strategies to reduce optimistic bias with respect to relative risk.⁹¹ The punchline? None of these strategies worked. One study asked people to generate their own list of personal “factors” that might either increase or decrease their risk of developing a weight problem or a drinking problem. The authors hypothesized that an identification of factors would decrease optimistic bias in many cases; no such effect was observed. Nor was optimistic bias reduced by asking participants to read about major risk factors for certain hazards, to report their standing with respect to these factors, and to offer an overall risk estimate after responding to the list of factors. The authors conclude that “health campaigns emphasizing high-risk targets (such as smoking interventions that show unattractive pictures of smokers) and campaigns conveying information about undesirable actions (as with pamphlets listing factors that raise the risk for a particular health problem) may unwittingly worsen the very biases they are designed to reduce” (p 323).

As the authors note, one intervention has been found to reduce optimistic bias: giving people information about their own standing on risk factors or about their peers’ standing on risk factors. But they observe, sensibly enough, that it is not easy to adapt this information to media campaigns designed to improve human health. This is a valuable paper, adding to the still-emerging literature on the possibility of

⁸⁹ Christine Jolls, *Behavioral Economics Analysis of Redistributive Legal Rules*, 51 Vand L Rev 1653, 1660–61 (1998).

⁹⁰ See *id.* at 1658–62 (discussing individuals’ failure to insure as a function of their over-optimism). Note that the availability heuristic can counteract this problem by leading people to insure against salient risks.

⁹¹ Neil D. Weinstein and William M. Klein, *Resistance of Personal Risk Perceptions to Debiasing Intervention*, (pp 313–22) (discussing four strategies for avoiding over-optimism, including informing individuals of their true risk, requiring comparisons with low rather than high risk groups, and having individuals brainstorm ways to alter their risk factors).

debiasing (or the activation of System II). But because the focus is on the “above average” effect, the findings do not offer clear guidance about campaigns designed to give people a better sense of the statistical reality. It would be valuable to learn more about that topic.

C. Optimistic Fools?

David Armour and Shelley Taylor are concerned with some obvious puzzles: If people are excessively optimistic, why don't they pursue ambitious goals recklessly and blunder? Why don't alert people—psychologists? entrepreneurs?—take systematic advantage of human optimism? This is what Armour and Taylor see as the “dilemma of unrealistic optimism”—the likelihood that if real, this bias would produce extremely serious harmful effects. If optimism were widespread, we should probably see far more recklessness and failure than we generally observe.

The authors resolve the dilemma by giving a more refined sense of the nature of optimistic bias. In their view, people are not indiscriminately or blindly optimistic. Their predictions are usually within reasonable bounds (p 346). People are less likely to be optimistic when the consequences of error are severe (p 339). In addition, optimism decreases if the outcome will be known in the near future (id). Optimism also decreases when people are in a predecisional state of deliberation. When people are choosing among goals, or among possible courses of action, the bias is attenuated, and it increases again only after people have selected goals and begin to implement their plans (p 340). There is also evidence that optimistic bias, when it exists, can be adaptive,⁹² leading to (almost) self-fulfilling policies, increasing the likelihood of success (p 341).

These claims raise real doubts about the view that optimistic bias provides a good reason for paternalistic interventions. To be sure, we know enough about optimistic bias to give serious consideration to informational campaigns to ensure that people will not have an inflated belief in their own immunity. In the context of smoking, statistical knowledge of risks⁹³ might be inadequate if people believe themselves relatively immune.⁹⁴ But in view of the arguments by Armour and Taylor, the idea that paternalism is generally justified by optimistic bias must be regarded as an unproven speculation. If people are not excessively optimistic when the consequences of error are severe, if the bias

⁹² A general treatment is Taylor, *Positive Illusions* (cited in note 82).

⁹³ See Viscusi, *Smoke-Filled Rooms* at 221 (cited in note 86) (proposing as a way to reduce smoking “that we target our informational efforts at providing comparative risk information”).

⁹⁴ See John Z. Ayanian and Paul D. Cleary, *Perceived Risks of Heart Disease and Cancer among Cigarette Smokers*, 281 JAMA 1019, 1020–21 (1999) (finding that most smokers think that their risks are average or below average).

is small or nonexistent when decisions are actually being made, and if people overstate low-probability risks, there is no problem for the law to correct.

V. MORAL HEURISTICS?

The heuristics-and-biases literature was originally focused on issues of probability, and while many of the chapters go beyond that topic, they do not much deal with normative questions—with the role of heuristics in informing judgments about morality and politics.⁹⁵ It is natural to wonder whether the rules of morality also have heuristics (isn't that inevitable?), and whether the normative judgments involved in law and politics are also prone to heuristics, or to rapid System I assessments and to possible System II override.

We can imagine some ambitious claims here. Armed with psychological findings, utilitarians might be tempted to claim that ordinary moral commitments are a set of mental shortcuts that generally work well, but that also produce severe and systematic errors (Is retribution a cognitive error? Is Kantianism?). On one view, much of everyday morality, nominally concerned with fairness, should be seen as a set of heuristics for the real issue, which is how to promote utility. For their part, deontologists could easily turn the tables. Deontologists might well claim that the rules recommended by utilitarians are consistent, much of the time, with what morality requires—but also that utilitarianism, taken seriously, produces mistakes in some cases. These large debates are unlikely to be tractable, simply because utilitarians and deontologists are most unlikely to be convinced by the suggestion that their defining commitments are mere heuristics. But in some cases, we might be able to make progress by entertaining the hypothesis that certain widely accepted rules of morality are heuristics. Consider, for example, the idea that one should “never lie” or “never steal”—good rules of thumb, certainly, but injunctions that badly misfire (sane people think) when the lie, or the theft, is needed to protect the deaths of innocent people.⁹⁶ I turn to several possible “moral heuristics” of relevance to law.⁹⁷

⁹⁵ An exception is the suggestive discussion by Philip Tetlock. See Philip E. Tetlock, *Intuitive Politicians, Theologians, and Prosecutors* (pp 596–98) (urging that many people believe in “taboo tradeoffs,” and that we might see such people not as defective intuitive economists, but as defenders of sacred values).

⁹⁶ Note the rule-utilitarian defense of these ideas: They might misfire in particular cases, but it might be best for people to treat them as firm rules, because a case-by-case inquiry would prove even more errors. If people ask whether the circumstances warrant an exception to the prohibition on lying or stealing, there might well be excessive or self-serving lying and stealing. The strong voice of conscience—calling for adherence to what I am calling moral heuristics even in cases in which they badly misfire—probably serves some valuable social functions. For fallible human beings, a decision to go right to the issue of consequences, without firm moral rules of

A. Pointless Punishment

People's intuitions about punishment seem disconnected with the consequences of punishment, in a way that suggests a moral heuristic is at work. Consider, for example, an intriguing study of people's judgments about penalties in cases involving harms from vaccines and birth control pills.⁹⁸ In one case, subjects were told that the result of a higher penalty would be to make companies try harder to make safer products. In an adjacent case, subjects were told that the consequence of a higher penalty would be to make the company more likely to stop making the product, with the result that less safe products would be on the market. Most subjects, including a group of judges, gave the same penalties in both cases. Can this outcome be defended in principle? Perhaps it can, but it is more sensible to think that people are operating under a heuristic, to the effect that penalties should be a proportional response to the outrageousness of the act, and should not be based on consequential considerations.

If this claim seems too adventurous, consider a similar test of punishment judgments, which asked subjects, including judges and legislators, to choose penalties for dumping hazardous waste.⁹⁹ In one

thumb, would likely produce serious problems.

⁹⁷ For a discussion of moral issues that appeals to intuitions, see F.M. Kamm, *Responsibility and Collaboration*, 28 Phil & Pub Aff 169, 173 (1999) (discussing consequentialism and collaboration with evil). Kamm's treatment is extremely impressive, but it seems to place undue emphasis on moral intuitions about exotic cases of the kind never or rarely encountered in ordinary life. I believe that the relevant intuitions ordinarily work well, but that when they are wrenched out of familiar contexts, in which they make a great deal of sense, their reliability, for purposes of legal and moral analysis, is unclear. Consider the following intuition: *Do not kill an innocent person, even if this is necessary to save others*. In all likelihood, a society does much better if most people have this intuition, if only because judgments about necessity are likely to be unreliable and self-serving. But in a hypothetical case, in which it really is necessary to kill an innocent person to save five others, our intuitions might well turn out to be unclear and contested. And if our intuitions about the hypothetical case turn out to be very firm (do not kill innocent people, ever!), they might not deserve to be so firm, simply because they have been wrenched out of the real world context, which is where they need to be to make sense.

I wonder whether some legal and philosophical analysis, based on exotic moral dilemmas, might not be replicating the early work of Kahneman and Tversky: uncovering situations in which intuitions, normally quite sensible, turn out to misfire. The irony is that Kahneman and Tversky meant to devise cases that would demonstrate the misfiring; some philosophers, including Kamm, devise cases with the thought that the intuitions are reliable and should form the building blocks for sound moral judgments. An understanding of how heuristics work suggests reason to doubt the reliability of those intuitions, even when they are very firm. Much work remains to be done on this complicated topic; my discussion in this Part is intended as a tentative start. For a more detailed discussion, see Cass R. Sunstein, *Moral Heuristics* (forthcoming).

⁹⁸ See Jonathan Baron and Ilana Ritov, *Intuitions about Penalties and Compensation in the Context of Tort Law*, 7 J Risk & Uncertainty 17, 17 (1993) (reporting that people use overgeneralized rules to determine penalties, ignoring their deterrent effects on future behavior).

⁹⁹ See Jonathan Baron, et al, *Attitudes Toward Managing Hazardous Waste*, 13 Risk Anal 183, 183 (1993) (reporting that in a study of CEOs, economists, environmentalists, judges, lawmakers, and hazardous waste policy experts, retributive penalties were preferred over welfare-

case, the penalty would make companies try harder to avoid waste. In another, the penalty would lead companies to cease making a beneficial product. Most people did not penalize companies differently in the two cases. Perhaps most strikingly, people preferred to require companies to clean up their own waste, even if the waste did not threaten anyone, instead of spending the same amount to clean up far more dangerous waste produced by another, now-defunct company. It is reasonable to believe that in thinking about punishment, people use a simple heuristic, the now-familiar *outrage heuristic*. This heuristic produces reasonable results in most circumstances, but in some cases, it seems to lead to systematic errors.

B. Aversion to Cost-Benefit Analysis

An automobile company is deciding whether to take certain safety precautions for its cars. In deciding whether to do so, it conducts a cost-benefit analysis, in which it concludes that certain precautions are not justified—because, say, they would cost \$100 million and save only four lives, and because the company has a “ceiling” of \$10 million per life saved. How will ordinary people react to this decision? The answer is that they will not react favorably.¹⁰⁰ They tend to punish companies that base their decisions on cost-benefit analysis, even if a high valuation is placed on human life. By contrast, they do not much punish companies that are willing to impose a “risk” on people.¹⁰¹ What underlies these moral judgments?

It is possible that when people disapprove of trading money for risks, they are generalizing from a set of moral principles that are generally sound, and even useful, but that work poorly in some cases. Consider the following moral principle: *Do not knowingly cause a human death*. People disapprove of companies that fail to improve safety when they are fully aware that deaths will result—whereas people do not disapprove of those who fail to improve safety while appearing not to know, for certain, that deaths will ensue. Companies that fail to do cost-benefit analysis, but that are aware that a “risk” exists, do not make clear, to themselves or to jurors, that they caused deaths with full knowledge that this was what they were going to do. People disapprove, above all, of companies that cause death know-

maximizing penalties for knowing polluters).

¹⁰⁰ See Viscusi, 52 *Stan L. Rev.* at 586–90 (cited in note 41) (concluding that jurors hold a variety of biases that hinder appropriate application of cost-benefit analysis, including possibly a bias against the use of cost-benefit analysis itself).

¹⁰¹ See *id.* See also Philip E. Tetlock, *Coping With Tradeoffs: Psychological Constraints and Political Implications*, in Arthur Lupia, Mathew D. McCubbins, and Samuel L. Popkin, eds., *Elements of Reason: Cognition, Choice, and the Bounds of Rationality* 239, 252–57 (Cambridge 2000) (summarizing different “taboo tradeoff” behaviors and their political implications).

ingly. I suggest, then, that a genuine heuristic is at work, one that imposes moral condemnation on those who knowingly engage in acts that will result in human deaths.

The problem is that it is not always unacceptable to cause death knowingly, at least if the deaths are relatively few and an unintended byproduct of generally desirable activity. Much of what is done, by both industry and government, is likely to result in one or more deaths. Of course it would make sense, in many of these domains, to take extra steps to reduce risks. But that proposition does not support the implausible claim that we should disapprove, from the moral point of view, of any action taken when deaths are foreseeable.

I believe that it is impossible to vindicate, in principle, the widespread social antipathy to cost-benefit balancing. But to adapt a claim about the representativeness heuristic by Stephen Jay Gould (p 68), “a little homunculus in my head continues to jump up and down, shouting at me” that corporate cost-benefit analysis, trading dollars for a known number of deaths, is morally unacceptable. The voice of the homunculus, I am suggesting, is not the result of conscience, but instead of a crude but quite tenacious moral heuristic.

C. Acts and Omissions

There has been much discussion of whether and why the distinction between acts and omissions might matter for law and policy. 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. 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 constitutional law reflects judgments to this effect.¹⁰² But what is the morally relevant difference?

It is worth considering the possibility that the action-omission distinction operates as a heuristic for a more complex and difficult assessment of the moral issues at stake. From the moral point of view, harmful acts are generally worse than harmful omissions, in terms of both the state of mind of the wrongdoer and the likely consequences of the wrong. But harmful acts are not *always* worse than harmful omissions. 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, driven by a

¹⁰² See *Washington v Glucksberg*, 521 US 702, 724–25 (1997) (upholding the state of Washington's law that prohibits aiding a suicide).

heuristic that opposes actions more than omissions, strongly suggest that there must be such a difference. In such cases, we might hypothesize that moral intuitions reflect an overgeneralization of principles that usually make sense—but that fail to make sense in the particular case.¹⁰³ In other words, moral intuitions reflect System I; they need to be corrected by System II. I believe that the persistent acceptance of withdrawal of life-saving equipment, alongside persistent doubts about euthanasia, is a demonstration of the point.

Consider in this regard the dispute over two well-known problems in moral philosophy.¹⁰⁴ 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.

But what is the difference between the two cases? A great deal of philosophical work has been done on this question, often suggesting that our intuitions can be defended in principle. Let me suggest a simpler answer. As a matter of principle, there is no difference between the two cases. People's different reactions are based on moral heuristics that condemn the throwing of the stranger but support the throwing of the switch. In the footbridge case, the heuristic says, *Do not lay hands on people in order to cause their deaths*. Such heuristics generally point in the right direction. But they misfire in drawing a distinction between the two cases. In this sense, the action-omission distinction leads to systematic errors.

Is there anything to be said to those who believe that their moral judgments, distinguishing the trolley and footbridge problems, are entirely deliberative, and reflect no heuristic at all? Consider an intriguing experiment, designed to see how the human brain responds to the two problems. The authors do not attempt to answer the moral questions in principle, but they find “that there are systematic variations in

¹⁰³ See Jonathan Baron, *Nonconsequentialist Decisions*, 17 *Beh and Brain Sci* 1, 1 (1994) (“I suggest that nonconsequentialist principles arise from overgeneralizing rules that are consistent with consequentialism in a limited set of cases.”).

¹⁰⁴ See Joshua D. Greene, et al, *An fMRI Investigation of Emotional Engagement in Moral Judgment*, 293 *Sci* 2105, 2105–06 (2001) (relying on a study of the two familiar moral dilemmas to argue that moral dilemmas require emotional thought processes to varying degrees, influencing moral judgments).

the engagement of emotions in moral judgment,”¹⁰⁵ and that brain areas associated with emotion are far more active in contemplating the footbridge problem than in contemplating the trolley problem.¹⁰⁶ As in the case of fear, where an identifiable region of the brain makes helpfully immediate but not entirely reliable judgments,¹⁰⁷ and where other, also identifiable regions can supply correctives, so too, perhaps, in the context of morality and law.

D. Betrayals

A betrayal of trust is likely to produce a great deal of outrage. If a babysitter neglects a child, or if a security guard steals from his employer, people will be angrier than if the identical acts were performed by someone in whom trust has not been reposed. So far, perhaps, so good. And it should not be surprising that people will favor greater punishment for betrayals than for otherwise identical crimes.¹⁰⁸ Perhaps the disparity can be justified on the ground that the betrayal of trust is an independent harm, one that warrants greater deterrence and retribution—a point that draws strength from the fact that trust, once lost, is not easily regained. But consider a finding that is harder to explain: People are especially averse to risks of death that come from products designed to promote safety, so much so that people have been found to prefer a greater chance of dying, as a result of accidents from a crash, to a significantly lower chance of dying in a crash as a result of a malfunctioning air bag.¹⁰⁹ Indeed, “most people are willing to double their chance of dying to avoid incurring a very small chance of dying via betrayal.”¹¹⁰

What explains this seemingly bizarre and self-destructive preference? I suggest that a heuristic is at work: *Punish, and never reward, betrayals of trust*. The heuristic generally works well. But it misfires in some cases, as when those who deploy it end up increasing the risks they themselves face. An air bag is not a security guard or a babysitter, endangering those whom they have been hired to protect. It is a prod-

¹⁰⁵ Id at 2106.

¹⁰⁶ Id (presenting evidence that different areas of the brain are affected for the two classic moral problems).

¹⁰⁷ See LeDoux, *The Emotional Brain* at 157–69 (cited in note 17) (discussing the amygdala as a region of the brain responsible for perceiving fear).

¹⁰⁸ See Jonathan J. Koehler and Andrew D. Gershoff, *Betrayal Aversion: When Agents of Protection Become Agents of Harm*, *Org Beh & Human Dec Processes* (forthcoming 2003) (relying on five empirical studies to find that acts of betrayal elicited stronger desired punishments than other bad acts).

¹⁰⁹ Id at 40 (finding that “when faced with a choice among pairs of safety devices . . . most people preferred inferior options (in terms of risk exposure) to options that included a slim (0.01%) risk of betrayal”).

¹¹⁰ Id at 33–34.

uct, to be chosen if it decreases aggregate risks. If an air bag makes people safer on balance, it should be used, even if in a tiny percentage of cases it will create a risk that would not otherwise exist. To reject air bags on grounds of betrayal is irrational but understandable—the sort of mistake to which heuristics often lead human beings. The distinctive feature of the anti-betrayal heuristic is that it involves moral and legal judgments rather than judgments of fact.

These are speculative remarks on some complex subjects. But if heuristics play a role in factual judgments, and sometimes lead people to make systematic errors, there is reason to believe that heuristics also help produce normative judgments, both moral and legal, and sometimes produce errors there as well. If this is harder to demonstrate, it is largely because we are able to agree about what constitutes error in the domain of facts, and often less able to agree about what constitutes error in the domain of values. I believe that *Heuristics and Biases: The Psychology of Intuitive Judgment* will illuminate problems of law and policy for many years to come. And we should not be surprised if the ideas of attribute substitution, and of the correction of rapid, intuitive assessments by more reflective processes, have analogues in moral and legal intuitions as well.