

ISSN 1045-6333

HARVARD

JOHN M. OLIN CENTER FOR LAW, ECONOMICS, AND BUSINESS

DEBIASING THROUGH LAW

Christine Jolls
Cass R. Sunstein

Discussion Paper No. 495

09/2004

Harvard Law School
Cambridge, MA 02138

This paper can be downloaded without charge from:

The Harvard John M. Olin Discussion Paper Series:
http://www.law.harvard.edu/programs/olin_center/

JEL Classification Codes K00, K11, K13, K22, K31

Debiasing Through Law

Christine Jolls*
Cass R. Sunstein**

Abstract

Human beings are often boundedly rational. In the face of bounded rationality, the legal system might attempt either to “debias law,” by insulating legal outcomes from the effects of boundedly rational behavior, or instead to “debias through law,” by steering legal actors in more rational directions. Legal analysts have focused most heavily on insulating outcomes from the effects of bounded rationality. In fact, however, a large number of actual and imaginable legal strategies are efforts to engage in debiasing through law – to help people reduce or even eliminate boundedly rational behavior. In important contexts, these efforts promise to avoid the costs and inefficiencies associated with regulatory approaches that take bounded rationality as a given and respond by attempting to insulate outcomes from its effects. This Article offers both a general theory of debiasing through law and a description of how such debiasing does or could work to address central legal questions in a large number of domains, from employment law to consumer safety law to corporate law to property law. Discussion is devoted to the risks of overshooting and manipulation that are sometimes raised when government engages in debiasing through law.

* Professor of Law, Harvard Law School.

** Karl N. Llewellyn Distinguished Service Professor of Jurisprudence, University of Chicago. For helpful comments, we are very grateful to Bruce Ackerman, Ian Ayres, Lucian Bebchuk, Cary Coglianese, Bert Huang, Louis Kaplow, Eric Posner, Jeffrey Rachlinski, Reva Siegel, Peter Siegelman, Adrian Vermeule, and participants in workshops at Boston University Law School, Fordham Law School, Harvard Law School, the John F. Kennedy School of Government at Harvard University, and Yale Law School.

Debiasing Through Law

Christine Jolls and Cass R. Sunstein

© 2004 Christine Jolls and Cass. R. Sunstein. All rights reserved.

A growing body of legal analysis focuses on how human behavior deviates systematically from what would be predicted by the traditional economic assumption of unbounded rationality.¹ To the extent that legal rules are designed on the basis of their anticipated effects on behavior, bounded rationality is obviously relevant to the formulation of legal policy. But an important and little addressed question is precisely *how* it is relevant to the formulation of legal policy. The most obvious possibility is that, given a demonstration of the existence and importance of a particular aspect of bounded rationality, the law should be structured to presume the existence of that particular shortcoming in human behavior.

Consider, for instance, the possibility that boundedly rational consumers believe that potentially risky products are substantially safer than they in fact are.² If this is the case, then the law might respond by adopting

¹ See, e.g., THE LAW AND ECONOMICS OF IRRATIONAL BEHAVIOR (Francesco Parisi & Vernon Smith eds., forthcoming 2004); BEHAVIORAL LAW AND ECONOMICS (Cass R. Sunstein ed., 2000); Symposium, *Empirical Legal Realism: A New Social Scientific Assessment of Law and Human Behavior*, 97 NW. U.L. REV. 1075 (2003); Symposium, *The Legal Implications of Psychology: Human Behavior, Behavioral Economics, and the Law*, 51 VAND. L. REV. 1497 (1998); Christine Jolls, Cass R. Sunstein & Richard Thaler, *A Behavioral Approach to Law and Economics*, 50 STAN. L. REV. 1471 (1998); Mark Kelman, *Behavioral Economics as Part of a Rhetorical Duet: A Response to Jolls, Sunstein, and Thaler*, 50 STAN. L. REV. 1577 (1998); Richard A. Posner, *Rational Choice, Behavioral Economics, and the Law*, 50 STAN. L. REV. 1551 (1998) (responding to Jolls, Sunstein and Thaler article); Russell Korobkin & Thomas Ulen, *Law and Behavioral Science: Removing the Rationality Assumption from Law and Economics*, 88 CALIF. L. REV. 1051 (2000).

² See Jon D. Hanson & Douglas A. Kysar, *Taking Behavioralism Seriously: Some Evidence of Market Manipulation*, 112 HARV. L. REV. 1420, 1511-12 (1999).

heightened standards of manufacturer liability for consumer products.³ Or suppose that the phenomenon of “Monday morning quarter-backing” adversely affects judgments reached by ex post decision makers on matters of corporate law, so that corporations are held liable for bad events even if preventing those events would have been extremely difficult.⁴ If so, then the law could respond, as indeed it has with the “business judgment rule,” by largely vitiating the liability of corporate law actors, who would otherwise be vulnerable to such second-guessing on the part of adjudicators.⁵ More generally, rules and institutions might be designed so that legal outcomes do not fall prey to problems of bounded rationality.

Boundedly rational behavior thus might be taken to justify a strategy of *insulation*, attempting to protect legal outcomes from falling victim to bounded rationality. To date, most treatments of bounded rationality in law have been of this character. Strategies for insulation can be characterized as a method for “debiasing law.”

A quite different possibility – one that has received surprisingly limited attention in law and elsewhere – is that legal policy may respond best to problems of bounded rationality not by insulating legal outcomes from its effects, but instead by operating directly on the boundedly rational behavior and attempting to help people either to reduce or to eliminate it. We describe legal policy in this category as “debiasing through law.” Debiasing through law will often be a less intrusive, more direct, and more democratic response to the problem of bounded rationality.

In fact there exists a substantial, empirically-oriented social science literature on the debiasing of individuals after a demonstration of a given form of bounded rationality.⁶ But empirical findings on these forms of

³ See *id.* at 1560.

⁴ See Jeffrey J. Rachlinski, *A Positive Psychological Theory of Judging in Hindsight*, 65 U. CHI. L. REV. 571, 620-21 (1998).

⁵ See *id.*

⁶ Leading examples include Baruch Fischhoff, *Debiasing*, in *JUDGMENT UNDER UNCERTAINTY: HEURISTICS AND BIASES* 422 (Daniel Kahneman et al. eds., 1982); Lawrence Sanna, Norbert Schwarz & Shavaun L. Stocker, *When Debiasing Backfires: Accessible Content and Accessibility Experiences in Debiasing Hindsight*, 28 J. EXPERIMENTAL PSYCHOL.: LEARNING, MEMORY, & COGNITION 497 (2002); and Neil D. Weinstein & William M. Klein, *Resistance of Personal Risk Perceptions to Debiasing Interventions*, in

debiasing have made only limited appearances in the legal literature,⁷ and equally important, social scientists interested in such forms of debiasing have generally not investigated the possibility of achieving them through law. In many important settings, empirical evidence suggests the substantial potential of these sorts of debiasing strategies, and from a legal policy perspective it is obviously important to ask about the role that law can play in facilitating such debiasing. That is our major focus in this Article.

When debiasing through law has been discussed in the legal literature, the treatment has focused on existing or proposed steps taken in *procedural* rules governing adjudication by judges or juries. A well-known example is the work by Linda Babcock, George Loewenstein and Samuel Issacharoff on the tendency of litigants to evaluate likely outcomes, as well as questions of fairness, in ways that systematically serve their own interests.⁸ Thus, for instance, individuals assigned to the role of the plaintiff and presented with exactly the same information as is presented to individuals assigned to the role of the defendant offer far higher estimates of the likelihood of a plaintiff victory in a lawsuit.⁹ Babcock, Loewenstein and Issacharoff find, however, that in an experimental setting, this self-serving bias may be eradicated by requiring litigants to consider the weaknesses in their case or reasons that the judge might rule against them. In these circumstances, individuals in the plaintiff's and defendant's roles have similar views on likely trial outcomes and fair settlements.¹⁰ This Article, by contrast, gives primary emphasis to a

HEURISTICS AND BIASES: THE PSYCHOLOGY OF INTUITIVE JUDGMENT 313
(Thomas Gilovich et al. eds., 2002). Many other illustrations appear below.
See infra notes 84-85, 93-94, 97-98, 142, 186-88, and 194-95.

⁷ As noted in the text just below, where debiasing of boundedly rational actors has been examined in the legal literature, the focus has been on achieving debiasing through procedural rules governing adjudication by judges or juries. We discuss several examples below.

⁸ Linda Babcock, George Loewenstein & Samuel Issacharoff, *Creating Convergence: Debiasing Biased Litigants*, 22 L. & SOC. INQUIRY 913, 917-18 (1997); Seth Seabury & Eric Talley, Private Information, Self-serving Biases, and Optimal Settlement Mechanisms: Theory and Evidence (unpublished manuscript, on file with authors).

⁹ Babcock, Loewenstein & Issacharoff, *supra* note 8, at 917-18.

¹⁰ *Id.* at 918-19. It is possible that “real” self-serving bias – outside of the lab – is more resilient. *See* Ward Farnsworth, *The Legal Regulation of Self-Serving Bias*, 37 U.C. DAVIS L. REV. 567, 582-85 (2003).

different and broader form of debiasing through law – a category we call “debiasing through substantive law.”

The central idea of debiasing through substantive law is that in some cases it may be desirable to structure the substance of law – not merely the procedures by which the law is applied in an adjudicative setting – with an eye toward debiasing those who suffer from bounded rationality. As a simple example pointing in the general direction we will explore below, consider “cooling-off” periods for consumer decisions.¹¹ Under temporary pressures, consumers might make ill-considered or improvident decisions, in part as a result of bounded rationality. Responding to this concern, the Federal Trade Commission imposes a mandatory cooling-off period for door-to-door sales.¹² Under the Commission’s rule, door-to-door sales must be accompanied by written statements informing buyers of their right to rescind purchases within three days of transactions.¹³ Aware of the possibility that people might err, regulators do not block their choices (a “debiasing law” response), but instead ensure a period for sober reflection. Thus, the law seeks to help people move in more rational directions – in this example through the simple and limited expedient of requiring a meaningful time window before decision making can occur. A central point of this Article is that, well beyond the simple case of cooling-off periods, the possibility of attempting to reduce bounded rationality through substantive law holds previously unrecognized promise for both understanding and improving the legal system.

Part I below offers a general descriptive theory of debiasing through law. We start by developing the basic distinction between debiasing of boundedly rational actors and the provision of incentives – another important instrument for affecting actors’ behavior. We then move to the specific role of the legal system in achieving debiasing of boundedly rational actors. We offer a general typology of strategies for debiasing through law, attempting to clarify and refine the distinction noted above between debiasing through procedural rules and debiasing through substantive law.

¹¹ Colin Camerer, Samuel Issacharoff, George Loewenstein, Ted O’Donoghue & Matthew Rabin, *Regulation for Conservatives: Behavioral Economics and the Case for “Asymmetric Paternalism,”* 151 U. PA. L. REV. 1211, 1241-42 (2003), offers a valuable discussion.

¹² 16 C.F.R. § 429.1(a) (2003).

¹³ *Id.* § 429.1(a).

Part II illustrates the general scope and power of the idea of debiasing through substantive law by describing the role it does or could play in addressing central questions across a range of legal domains, from employment law to consumer safety law to corporate law to property law. Our analysis of debiasing through substantive law contrasts with the almost reflexive focus in the existing legal literature on “debiasing law” solutions to the problem of bounded rationality.

Part III offers a framework for resolving the important but little discussed normative questions raised by debiasing through law. Compared to the usual approach of “debiasing law,” an important advantage of strategies for debiasing through law is that they aim to correct errors while still preserving some opportunity to make choices. Unlike attempts to insulate legal outcomes from boundedly rational judgments, debiasing through law generally preserves a space for citizens to arrange their affairs as they like. Under the Commission’s door-to-door sales rule, for instance, no transaction is blocked. An important corollary is that, unlike “debiasing law” strategies, the approach of debiasing through law will frequently make it possible for government to improve outcomes for individuals who exhibit bounded rationality while leaving unrestricted the choices of those who would not otherwise err. An emerging theme in the legal literature on bounded rationality is that it is preferable, when possible, to develop legal approaches that avoid imposing significant costs on those who do not exhibit boundedly rational behavior¹⁴; below we describe specific strategies for debiasing through law that achieve this goal. In this sense, debiasing through law provides real advantages over “debiasing law” strategies.

Still, at bottom, debiasing through law in either of its two varieties (substantive or procedural) involves the government in a self-conscious process of changing the behavior of at least some people by altering their perceptions of or attitudes toward the reality around them. In some contexts such government action is entirely routine and wholly uncontroversial, as when government provides accurate information to otherwise misinformed individuals. But in some contexts, debiasing through law raises important problems that require discussion. Our normative framework identifies some

¹⁴ See Camerer, Issacharoff, Loewenstein, O’Donoghue & Rabin, *supra* note 11, at 1212; Gregory Mitchell, *Why Law and Economics’ Perfect Rationality Should Not Be Traded for Behavioral Law and Economics’ Equal Incompetence*, 91 GEO. L.J. 67, 132 (2002).

of the main objections to certain kinds of debiasing through law and, more importantly, provides a way of evaluating specific strategies for debiasing through law. We conclude by briefly discussing a corollary of our main analysis of debiasing through law – the idea of debiasing through switches in default rules.

I. A General Theory of Debiasing Through Law

A foundational question in developing a general theory of debiasing through law is how, as a general matter, debiasing of boundedly rational actors should be defined. We offer a definition below, emphasizing the distinction between such debiasing and incentives, another instrument for affecting people’s behavior. After developing a general definition of debiasing of boundedly rational actors, we focus on the law’s actual and potential role in facilitating such debiasing. We elaborate the distinction between debiasing through procedural rules and debiasing through substantive law and offer a general typology of strategies for debiasing through law. Our focus throughout this Part is descriptive and analytic; we postpone normative questions until Part III below.

A. Preliminaries

If debiasing through law is a response to bounded rationality, an obvious first step is to understand the basic idea of bounded rationality. As is now well known in the legal literature and beyond, researchers in psychology and behavioral economics have uncovered a wide range of departures from unboundedly rational behavior. These departures take one of two general forms. First, individuals may exhibit “judgment errors.”¹⁵ Second, human behavior may deviate from the precepts of expected utility theory.¹⁶ We describe these two basic categories in turn.

¹⁵ See generally HEURISTICS AND BIASES: THE PSYCHOLOGY OF INTUITIVE JUDGMENT, *supra* note 6.

¹⁶ See generally CHOICES, VALUES, AND FRAMES (Daniel Kahneman & Amos Tversky eds., 2001).

1. “Judgment Errors”

a. Racial and other group-based unconscious biases. The well-known heuristics-and-biases literature discussed below reveals a wide range of ways people make systematic errors, but perhaps the most basic, everyday conception of a “bias” involves the unwarranted attribution of negative traits to members of racial or other discrete groups. A striking feature of this form of bias in its modern incarnation is that often people are wholly unaware of the negative attributions they make. Even those who sincerely and consistently disclaim racial and other prejudice often show substantial signs of racial or other group-based *unconscious* bias.¹⁷

Scholars from a wide range of fields have identified diverse means of assessing and measuring racial and other group-based unconscious bias,¹⁸ and we do not begin to attempt a comprehensive treatment here. Instead we focus on one leading technique from the modern social psychology literature for measuring racial or other group-based unconscious bias. This technique is the Implicit Attitudes Test (IAT), which has had widespread influence in part because of its ready availability on the Internet and the resulting large number of individuals who have taken the test.¹⁹

In the IAT, respondents are asked to categorize a series of stimuli (words or pictures) into four groups, two of which are demographic categories (such as “black” and “white”), and two of which are the categories “good” and “bad.” Groups are paired, so that a respondent would be asked to press one key on the computer for either “black” or “bad” and a different key for either “white” or “good” (a stereotype-consistent pairing); or would be asked to press one key on the computer for either “black” or “good” and a different key for either “white” or “bad” (a stereotype-consistent pairing). Stimuli are (for example) pictures of black faces, pictures of white faces, “good” words such as joy, love, peace, wonderful, pleasure, glorious,

¹⁷ See, e.g., Samuel L. Gaertner & John P. McLaughlin, *Racial Stereotypes: Associations and Ascriptions of Positive and Negative Characteristics*, 46 SOC. PSYCHOL. Q. 23 (1983); Anthony G. Greenwald & Mazharin R. Banaji, *Implicit Social Cognition: Attitudes, Self-Esteem and Stereotypes*, 102 PSYCHOL. REV. 4 (1995); Charles R. Lawrence III, *The Id, The Ego, and Equal Protection: Reckoning with Unconscious Racism*, 39 STAN. L. REV. 317 (1987).

¹⁸ See sources cited *supra* note 17.

¹⁹ See <http://projectimplicit.net/about.htm> (visited 9/1/04).

laughter, and happy, and “bad” words such as agony, terrible, horrible, nasty, evil, awful, and failure. Unconscious racial bias is defined as faster categorization when the “black” and “bad” categories are paired than when the “black” and “good” categories are paired.²⁰

The results of the IAT are striking. Three-quarters of respondents exhibit faster categorizations with the stereotype-consistent pairing (black-bad and white-good) than with the stereotype-inconsistent pairing (black-good and white-bad), while twelve percent exhibit no difference in speed of categorization. Only sixteen percent of respondents exhibit faster categorizations with the stereotype-inconsistent pairing. The tendency can be found among both whites and African-Americans; but looking at whites alone, the tendency to exhibit faster categorizations with the stereotype-consistent pairing is even more pronounced.²¹

As noted above, even those who explicitly disclaim racial or other group-based biases typically display such biases on the IAT and other tests of unconscious attitudes. Importantly, these individuals also display such biases in more everyday ways; scores on the IAT and similar tests show correlations with nonverbal (although not verbal, and thus presumably more conscious) behavior toward group members.²² The scores also show correlations with third parties’ assessments of the degree of friendliness individuals show toward group members.²³

Although racial and other group-based unconscious bias is not generally grouped with other forms of bounded rationality in the existing academic literature, the fit is closer than has been supposed. Racial or other group-based unconscious bias often results from the way in which the characteristic of race or other group membership operates as a “heuristic” – a mental short-cut that functions well in some settings but leads to systematic

²⁰ The test may be accessed from <https://implicit.harvard.edu/implicit/Instructions> (visited 9/1/04).

²¹ Results are summarized upon completion of the test available at <https://implicit.harvard.edu/implicit/Instructions> (visited 9/1/04).

²² See John F. Dovidio, Kerry Kawakami & Samuel L. Gaertner, *Implicit and Explicit Prejudice and Interracial Interaction*, 82 J. PERSONALITY & SOCIAL PSYCHOL. 62 (2002).

²³ See *id.*

errors in others.²⁴ (The concept of a heuristic is developed more fully just below.) Indeed, recent research in psychology emphasizes that heuristics typically work through a process of “attribute substitution,” in which people answer a hard question by substituting an easier one.²⁵ For instance, people might resolve a question of probability not by investigating statistics, but by asking whether a relevant incident comes easily to mind.²⁶ The same process operates to produce unconscious bias against racial or other groups, although of course it is not a complete explanation of such bias. And whether the heuristic is generally accurate (as in the case of rational statistical discrimination²⁷) or not, it can produce severe and systematic errors in particular cases, much like the behavior traditionally emphasized in the heuristics-and-biases literature in psychology and behavioral economics.

b. The heuristics-and-biases literature. Heuristics are important across a wide range of contexts, many of relevance to law. Consider a familiar example. Asked how many words in a 2,000-word section of a novel end in “ing,” people give much larger estimates than those asked how many words have “n” as the second-to-last letter in the same material, notwithstanding the obvious fact that more words satisfy the latter criterion than the former.²⁸ According to the “availability heuristic” at work in cases of this sort, the probability of an event is estimated after an assessment of how easily examples of the event can be called to mind. The

²⁴ For a general discussion of heuristics, see Daniel Kahneman & Shane Frederick, *Representativeness Revisited*, in HEURISTICS AND BIASES: THE PSYCHOLOGY OF INTUITIVE JUDGMENT, *supra* note 6, at 49; Peter M. Todd, *Fast and Frugal Heuristics for Environmentally Bounded Minds*, in BOUNDED RATIONALITY: THE ADAPTIVE TOOLBOX 51 (Gerd Gigerenzer & Reinhard Selten eds., 2002).

²⁵ See Kahneman & Frederick, *supra* note 24, at 53.

²⁶ See Norbert Schwartz & Leigh Ann Laughn, *The Availability Heuristic Revisited: Ease of Recall and Content of Recall As Distinct Sources of Information*, in HEURISTICS AND BIASES: THE PSYCHOLOGY OF INTUITIVE JUDGMENT, *supra* note 6, at 103, 103.

²⁷ See generally Kenneth J. Arrow, *The Theory of Discrimination*, in DISCRIMINATION IN LABOR MARKETS 3 (Orley Ashenfelter & Albert Rees eds., 1973); George Akerlof, *The Economics of Caste and of the Rat Race and Other Woeful Tales*, 90 Q.J. ECON. 599 (1976).

²⁸ Amos Tversky & Daniel Kahneman, *Extensional Versus Intuitive Reasoning: The Conjunction Fallacy in Probability Judgment*, 90 PSYCHOL. REV. 293, 295 (1983).

availability heuristic often produces sensible judgments and behavior for people who lack detailed statistical information, but it can lead to significant and severe errors.²⁹ The prospect of errors in some cases does *not* suggest that the behavior in question is “irrational” in the sense of being arbitrary or lacking plausible justification. The point instead is that the behavior, even if sensible in many cases, leads to systematic error in some of them. Bounded rationality is hardly the same as “irrationality.”³⁰

The use of heuristics has also been shown to lead people to misestimate probabilities by committing the “conjunction fallacy” (concluding that characteristics X and Y are more likely to be present than characteristic X alone) – errors produced by the so-called representativeness heuristic.³¹ For instance, after reading a paragraph about a thirty-one year old woman, Linda, who was concerned with issues of social justice and discrimination in college, most people erroneously tend to say that Linda is more likely to be “a bank teller and active in the feminist movement” than to be “a bank teller.”³² Heuristics, then, are not themselves biases, but they can produce biases. Thus “availability bias” might be said to arise when the availability heuristic leads people to make predictable errors in assessing probabilities.

A related set of findings by psychologists and behavioral economists emphasizes not mental short-cuts, but more direct biases that lead to inaccurate judgments. An example is hindsight bias, in which decision makers attach excessively high probabilities to events that ended up occurring; we referred to this bias above in discussing corporate law’s business judgment rule.³³ We also referred above to self-serving bias – in which individuals interpret information in directions that serve their own interests – in illustrating prospects for successful debiasing through procedural rules.³⁴

²⁹ See Jolls, Sunstein & Thaler, *A Behavioral Approach to Law and Economics*, *supra* note 1, at 1477-78.

³⁰ For further discussion, see Jolls, Sunstein & Thaler, *Theories and Tropes*, *supra* note 1, at 1594.

³¹ Amos Tversky & Daniel Kahneman, *Judgments of and by Representativeness*, in JUDGMENT UNDER UNCERTAINTY: HEURISTICS AND BIASES, *supra* note 6, at 84, 92-94.

³² *Id.*

³³ See *supra* notes 4-5 and accompanying text.

³⁴ See *supra* notes 9-10 and accompanying text.

Another bias that has received significant attention in the legal literature – and that we suggest in Part II creates important opportunities for debiasing through substantive law – is optimism bias. Optimism bias refers to the tendency of people to believe that their own probability of facing a bad outcome is lower than it actually is. People typically think that their chances of having an auto accident, contracting a particular disease, or getting fired from a job are significantly lower than the average person’s chances of suffering these misfortunes – although of course this cannot be true for everyone.³⁵ Estimates offered by individuals for their own probabilities range from twenty to eighty percent below the average person’s probability.³⁶

While the “above average” effect is well established, it does not by itself establish that people optimistically underestimate their statistical risk.³⁷ People could believe, for example, that they are less likely than most people to contract cancer, while also having an accurate sense of the probability that they will contract cancer. But substantial evidence suggests that people sometimes exhibit optimism bias in the estimation of actual probabilities, not simply relative risk. For example, professional financial experts consistently overestimate likely earnings, and business school students overestimate their likely starting salary and the number of offers that they will receive.³⁸ People also underestimate their own likelihood of being involved in a serious automobile accident,³⁹ and their frequent failure to buy insurance for floods and earthquakes is consistent with the view that people are excessively optimistic.⁴⁰ It is also noteworthy that these data pointing to optimism bias come from individuals making judgments that they make regularly in their everyday lives, rather than judgments far removed from those they would ordinarily make.⁴¹

³⁵ See Christine Jolls, *Behavioral Economic Analysis of Redistributive Legal Rules*, 51 VAND. L. REV. 1653, 1659-62 (1998) (discussing studies).

³⁶ *Id.* at 1659.

³⁷ See W. KIP VISCUSI, *SMOKE-FILLED ROOMS* 162-66 (2002).

³⁸ See David Armour & Shelley E. Taylor, *When Predictions Fail: The Dilemma of Unrealistic Optimism*, in *HEURISTICS AND BIASES: THE PSYCHOLOGY OF INTUITIVE JUDGMENT*, *supra* note 6, at 334, 334-35.

³⁹ See Jolls, *supra* note 35, at 1660-61.

⁴⁰ See *id.* at 1661.

⁴¹ Cf. Chris Guthrie, *Prospect Theory, Risk Preference, and the Law*, 97 NW. U. L. REV. 1115, 1156-57 (2003) (discussing relevance of study designs in

2. Departures from Expected Utility Theory

Along with the broad category of judgment errors, the idea of bounded rationality includes ways in which actual choices depart from the predictions of expected utility theory – a foundational feature of standard rational-choice analysis.⁴² A leading alternative to expected utility theory is Daniel Kahneman and Amos Tversky’s prospect theory.⁴³ According to this theory, people evaluate outcomes based on the change they represent from an initial reference point, rather than based on the nature of the outcome itself. For example, discovering that one will receive a bonus of \$2500 is often experienced differently if the previous year’s bonus was \$0 than if the previous year’s bonus was \$5000, wholly apart from any tangible financial obligations the individual faces. Prospect theory also posits that people weigh losses more heavily than gains, thus showing “loss aversion.”⁴⁴

Related to loss aversion is the “endowment effect,” according to which an individual’s valuation of an entitlement depends on whether the individual is given initial ownership of that entitlement.⁴⁵ Thus, for example, individuals endowed with university mugs demand substantially more to sell these mugs than unendowed individuals are willing to pay to buy them.⁴⁶ While this feature of bounded rationality has received only modest attention in the existing social science literature on debiasing of boundedly rational actors, we suggest its relevance to debiasing through law below.

Also related to loss aversion are framing effects. Losses matter more than gains, and thus framing outcomes as losses rather than gains will greatly

which subjects were making decisions of a type that they were accustomed to making).

⁴² See generally CHOICES, VALUES AND FRAMES, *supra* note 16.

⁴³ Daniel Kahneman & Amos Tversky, *Prospect Theory: An Analysis of Decision Under Risk*, 47 *ECONOMETRICA* 263 (1979).

⁴⁴ *Id.* at 273.

⁴⁵ See Daniel Kahneman, Jack Knetsch & Richard Thaler, *Experimental Tests of the Endowment Effect and the Coase Theorem*, 98 *J. POL. ECON.* 1325, 1325-26 (1990).

⁴⁶ See *id.* at 1329-42.

affect how people respond. Many studies support this idea⁴⁷ – as does a widely-publicized dispute over the content of a government advertising campaign in the United States. The advertising campaign in question involved the effects of breastfeeding of newborns. In the approach favored by breastfeeding advocates, the advertisements would refer to the risks to the child of leukemia and other diseases from not consuming breast milk – whereas infant formula manufacturers favored an approach that stressed the benefits to the child from breastfeeding.⁴⁸ Showing an intuitive understanding of prospect theory, the infant formula manufacturers fought to have the government emphasize the advantages of breastfeeding rather than the affirmative harms (losses) of not breastfeeding.

B. Debiasing Versus Incentives

What does it mean, in general, to “debias” boundedly rational actors? The general question of what counts as “debiasing” of boundedly rational actors is important because there are many channels by which boundedly rational behavior may be made to “go away” or diminish in degree. Some but not all of these qualify as debiasing of boundedly rational actors under the general definition we develop here.

Consider the following organizing examples:

(1) People are prone to social influences, so much so that many people will ignore the clear evidence of their own senses, and hence provide incorrect answers, if they are confronted with the unanimous views of others.⁴⁹ This kind of “conformity bias,” in which the views of others are used as a kind of heuristic for the proper answer, is significantly reduced when financial incentives are provided. When people stand to gain economically from a correct answer and when they have confidence in their

⁴⁷ For a recent overview, see Barbara A. Mellers, *Pleasure, Utility, and Choice*, in FEELINGS AND EMOTIONS: THE AMSTERDAM SYMPOSIUM 282 (Antony R. Manstead et al. eds., 2004).

⁴⁸ See Melody Petersen, *Breastfeeding Ads Delayed By a Dispute Over Content*, N.Y. TIMES, Dec. 4, 2003, at C1.

⁴⁹ See Solomon Asch, *Opinions and Social Pressure*, in READINGS ABOUT THE SOCIAL ANIMAL 13 (Elliott Aronson ed., 1995).

own judgment, they are far more likely to ignore the crowd, to say what they think, and to answer correctly.⁵⁰

(2) Recall that individuals in the role of litigants have a tendency to see cases in the light most favorable to their own side.⁵¹ But imagine that they are required to consider weaknesses in their side or reasons that the judge might rule against them. In that case, the “self-serving bias” bias they had previously exhibited vanishes.⁵²

(3) We have already mentioned thirty-one year old Linda, who was concerned with issues of social justice and discrimination in college. Recall that most people tend to say that Linda is more likely to be “a bank teller and active in the feminist movement” than to be “a bank teller.”⁵³ This is a familiar example of the conjunction fallacy, produced by the representativeness heuristic. But as Gerd Gigerenzer has demonstrated, people are less likely to commit the conjunction fallacy when asked about frequencies rather than probabilities. If asked, “of 100 people who fit the description” of Linda, how many are bank tellers and how many are bank tellers and active in the feminist movement, the level of conjunction violations drops from 80 percent or more to 20 percent or less.⁵⁴

The first of these three examples is one in which boundedly rational behavior is eliminated by the provision of financial incentives. A broad definition of debiasing of boundedly rational actors might embrace this sort of technique, but we think it is preferable to exclude the underlying form of behavior here from the category of boundedly rational behavior (so that the removal of the behavior by the provision of incentives does not count as “debiasing” of boundedly rational actors in the sense that we use that term). For some purposes, it might be useful to understand incentives as a way of overcoming boundedly rational behavior by increasing the stakes. Baruch

⁵⁰ See Robert Baron, *The Forgotten Variable in Conformity Research: Impact of Task Importance on Social Influence*, 71 J. PERSONALITY & SOCIAL PSYCHOL. 915 (1996).

⁵¹ See *supra* note 9 and accompanying text.

⁵² See *supra* note 10 and accompanying text.

⁵³ Tversky & Kahneman, *supra* note 31, at 92-94.

⁵⁴ GERD GIGERENZER, ADAPTIVE THINKING 250 (2000).

Fischhoff, for instance, describes “rais[ing] stakes” as a possible strategy for debiasing of boundedly rational actors.⁵⁵

But it seems most conservative, and most consistent with existing conventions in analyses of bounded rationality, to limit the category of boundedly rational behavior to that which survives even in the presence of financial or other consequences for exhibiting the behavior.⁵⁶ If an apparent departure from unbounded rationality is eliminated with the provision of financial incentives, then many would conclude that it was not a departure from unbounded rationality at all, but instead a mere result of lazy or careless decision making by an actor who had no reason to be other than lazy or careless. Under our definition, therefore, the technique used in the first example above is not a strategy for debiasing of boundedly rational actors. And the same goes for techniques that eliminate boundedly rational behavior by improving a previously faulty aspect of an experimental design – although here again Fischhoff’s broad conception of debiasing of boundedly rational actors embraces such manipulations.⁵⁷

The second and third examples above are standard cases of debiasing of individuals exhibiting bounded rationality. Subjects are asked to consider arguments or information of a particular sort, and the consideration of such arguments or information reduces or eliminates the boundedly rational behavior they previously exhibited. It is important to emphasize how the technique here differs from incentives. Actors are not asked to repeat the very same task with the very same structure, with the sole difference that they now have greater reason to take care in making their choices (which in the legal domain would correspond to some, although not all, “debiasing law” strategies – for instance, those that increase punishments for certain types of behavior); instead the environment is restructured in a way that alters not actors’ motivation but the actual process by which they perceive the reality around them. Thus, we define debiasing of boundedly rational actors as using

⁵⁵ Baruch Fischhoff, *Heuristics and Biases in Application*, in HEURISTICS AND BIASES: THE PSYCHOLOGY OF INTUITIVE JUDGMENT, *supra* note 6, at 730, 732; Fischhoff, *supra* note 6, at 424 & tbl.1.

⁵⁶ Colin Camerer & Robin Hogarth, *The Effects of Financial Incentives in Experiments: A Review and Capital-Labor-Production Framework*, 19 J. RISK & UNCERTAINTY 7 (1999), make this implicit claim, and offer a great deal of evidence that many cases of boundedly rational behavior are not eliminated by the provision of incentives.

⁵⁷ Fischhoff, *supra* note 6, at 424 & Table 1.

techniques that intervene in and alter the situation that produces the boundedly rational behavior, without operating on the degree of motivation or effort an actor brings to the task.

The example of employment discrimination law helps to drive home our central distinction between provision of incentives and debiasing of boundedly rational actors. In the face of racial or other group-based unconscious bias, simply creating or enhancing financial penalties for discriminatory conduct may be of quite limited effectiveness in combating the unconscious behavior.⁵⁸ Debiasing of the actors in question, by contrast, holds out promise for reducing the degree of racial or other group-based unconscious bias that is exhibited, as we explore at some length in Part II.A below.

Our definition of debiasing of boundedly rational actors connects well with “dual process” approaches of the sort that have received considerable recent attention in the psychology literature.⁵⁹ According to such approaches, people employ two cognitive systems. System I is rapid, intuitive, and error-prone; System II is more deliberative, calculative, slower, and more likely to be error-free. Heuristic-based thinking is rooted in System I; it is subject to override, under certain conditions, by System II.⁶⁰ The two systems need not be seen as occupying different physical spaces; indeed, the systems might even be understood as heuristics,⁶¹ although there is some evidence that different sectors of the brain can be associated with Systems I and II.⁶² But whether the two systems have a physical presence or not, System II can be used to correct the blunders produced by System I; strategies for debiasing of boundedly rational actors may be understood as an effort to activate System II in order to reduce the risk of mistake. As discussed at some length in Parts

⁵⁸ See, e.g., Linda Hamilton Krieger, *The Content of Our Categories: A Cognitive Bias Approach to Discrimination and Equal Employment Opportunity*, 47 STAN. L. REV. 1161, 1164-65 (1995).

⁵⁹ See generally SOCIAL JUDGMENTS (Joseph P. Forgas et al. eds., 2003); SHELLY CHAIKEN & YAACOV TROPE, DUAL-PROCESS THEORIES IN SOCIAL PSYCHOLOGY (1999); Kahneman & Frederick, *supra* note 24.

⁶⁰ See Kahneman & Frederick, *supra* note 24, at 51.

⁶¹ See *id.*

⁶² See JOSEPH LEDOUX, THE EMOTIONAL BRAIN 106-132 (1996); Arne Ohman & Stefan Wiens, *The Concept of An Evolved Fear Module and Cognitive Theories of Anxiety*, in FEELINGS AND EMOTIONS: THE AMSTERDAM SYMPOSIUM, *supra* note 47, at 58.

II and III below, however, some strategies for such debiasing actually attempt to enlist System I, correcting mistakes by invoking heuristics themselves.

C. A Typology of Strategies for Debiasing Through Law

We have defined debiasing of boundedly rational actors as using techniques that alter the situation that produces the boundedly rational behavior, without operating on the degree of motivation or effort an individual brings to the task. Debiasing *through law* is then the use of legal rules to achieve such an effect. (Thus our analysis bears but does not focus on forms of debiasing of boundedly rational actors – such as Robert Rasmussen’s example of credit scoring, thought to check optimism bias by loan officers⁶³ – that might be undertaken by private parties independent of direct legal intervention.) The present section develops more fully some of the general ways the law might facilitate debiasing of boundedly rational actors.

Figure 1 maps the basic terrain of strategies for debiasing through law. The column division marks the line between procedural rules governing the adjudicative process and substantive rules regulating actions taken outside of the adjudicative process. The row division marks the line between debiasing actors in their capacity as participants in the adjudicative process and debiasing actors in their capacity as decision makers outside of the adjudicative process. (Of course, the simple fact the lawyers play a role in the adjudicative process may blunt the effects of litigants’ or even adjudicators’ bounded rationality⁶⁴; nonetheless we believe that strategies for debiasing through procedural rules hold important promise, although they are not the main focus of this Article.) The upper left box in this matrix represents the type of debiasing through law on which the existing legal literature has focused: the rules in question are procedural rules governing the adjudicative process, and the actors targeted are individuals in their capacity as participants in the adjudicative process. One example of debiasing through procedural rules is that from above of litigation rules responding to self-

⁶³ Robert K. Rasmussen, *Behavioral Economics, The Economic Analysis of Bankruptcy Law and the Pricing of Credit*, 51 VAND. L. REV. 1679, 1695 (1998).

⁶⁴ See generally Russell Korobkin & Chris Guthrie, *Psychology, Economics, and Settlement: A New Look at the Role of the Lawyer*, 76 TEX. L. REV. 77 (1997).

serving bias.⁶⁵ In a similar vein, anchoring – in which judgments are influenced by arbitrary cues presented to decision makers, such as the dollar amount requested in a legal complaint – has been shown to produce important effects on jury awards; in response, procedural reforms might be adopted to ensure that juries consider other facts or features of the case in addition to the anchor.⁶⁶

Moving counterclockwise, the lower left box in the matrix is marked with an “X” because procedural rules governing the adjudicative process do not have any obvious role in debiasing actors outside of the adjudicative process – although these rules certainly may affect such actors’ behavior in various ways by influencing what would happen in the event of future litigation.⁶⁷ The lower right box in the matrix represents the category of debiasing through law emphasized in Part II below: the rules in question are substantive rules regulating actions taken outside of the adjudicative process, and the actors targeted are decision makers outside of the adjudicative process.

⁶⁵ See *supra* notes 9-10 and accompanying text.

⁶⁶ See Gretchen Chapman & Eric Johnson, *Incorporating the Irrelevant: Anchors in Judgments of Belief and Value*, in *HEURISTICS AND BIASES: THE PSYCHOLOGY OF INTUITIVE JUDGMENT*, *supra* note 6, at 120, 137; Gretchen Chapman & Eric Johnson, *Anchoring, Activation, and the Construction of Values*, 79 *ORGANIZATIONAL BEHAV. & HUM. DECISION PROCESSES* 115, 144 (1999).

⁶⁷ Perhaps there are even feasible mechanisms for debiasing, through the changes in procedural rules governing the adjudicative process, actors in their capacity as decision makers outside of the adjudicative process; but we do not explore the possibility here.

Figure 1: Typology of Strategies for Debiasing Through Law

		Type of Law	
		Procedural rules governing the adjudicative process	Substantive rules regulating actions taken outside of the adjudicative process
Role of Actor	Debiasing actors in their capacity as participants in the adjudicative process	Debiasing through procedural rules	“Hybrid” debiasing
	Debiasing actors in their capacity as decision makers outside of the adjudicative process	X	Debiasing through substantive law

Finally, the upper right corner of the matrix represents a hybrid category that warrants brief discussion, in part to demarcate it from the category of debiasing through law emphasized in Part II. In this hybrid category, it is substantive, rather than procedural, law that is structured to achieve debiasing of boundedly rational actors, but the aspect of bounded rationality that this debiasing effort targets is one that arises within, rather than outside of, the adjudicative process. For example, Ward Farnsworth’s recent work on self-serving bias suggests that such bias on the part of employment discrimination litigants (actors in their capacity as participants in an adjudicative process) might be reduced by restructuring employment discrimination standards (substantive rules regulating action outside of the adjudicative process) to increase the reliance of such standards on objective facts as opposed to subjective or normative judgments.⁶⁸ Farnsworth’s suggestion of debiasing through law operates through reform of substantive rather than procedural rules, but the actions to be debiased are those of litigants within the adjudicative process. In our analysis in Part II, by contrast, both the legal rules through which debiasing of boundedly rational

⁶⁸ Farnsworth, *supra* note 11, at 593-95.

actors occurs and the capacities in which actors are targeted for debiasing through law are distinct from the context of the adjudicative process.

D. The Domain of Analysis

Not all types of bounded rationality respond well to techniques for debiasing of boundedly rational actors. Many manipulations fail to reduce hindsight bias; even worse, some seemingly sensible strategies for eliminating hindsight bias have actually increased it.⁶⁹ To be sure, studies that have required subjects to “rethink the inferences that they have made upon learning [an] outcome” and have then “demonstrated to them that other inferences remained plausible” have shown some success in reducing hindsight bias.⁷⁰ However, in most cases strategies – even fairly aggressive ones – for debiasing boundedly rational actors have enjoyed limited, if any, success in combating hindsight bias.⁷¹

But in other contexts, techniques for debiasing of boundedly rational actors have shown substantial promise. One example is the case of self-serving bias; as noted, having subjects consider the weaknesses in their case or reasons that the judge might rule against them appears effective in eliminating self-serving bias in litigation – a form of debiasing through procedural rules.⁷² A second example is that racial or other group-based unconscious bias may be reduced upon exposure to bias-challenging stimuli. If, for instance, people view pictures of Michael Jordan and Tiger Woods before submitting to testing of unconscious racial bias through the IAT, they will exhibit substantially less unconscious racial bias – both immediately

⁶⁹ See Sanna, Schwarz & Stocker, *supra* note 6.

⁷⁰ Rachlinski, *supra* note 4, at 586-88. Merrie Jo Stallard & Debra L. Worthington, *Reducing the Hindsight Bias Utilizing Attorney Closing Arguments*, 22 J. L. & HUM. BEHAV. 671, 680-81 (1998), is an example of successful debiasing of hindsight-biased behavior.

⁷¹ See Fischhoff, *supra* note 6, at 427-31; Reid Hastie & W. Kip Viscusi, *What Juries Can't Do Well: The Jury's Performance as a Risk Manager*, 40 ARIZ. L. REV. 901, 917 (1998); Kim A. Kamin & Jeffrey J. Rachlinski, *Ex Post ≠ Ex Ante: Determining Liability in Hindsight*, 19 L. & HUM. BEHAV. 89, 97-98 (1995); Jeffrey J. Rachlinski, *Regulating in Foresight Versus Judging Liability in Hindsight: The Case of Tobacco*, 33 GA. L. REV. 813, 824 (1999).

⁷² Babcock, Loewenstein & Issacharoff, *supra* note 8, at 918-19.

after viewing the pictures and twenty-four hours later – than those not exposed to the pictures of Jordan and Woods.⁷³

The basic domain of debiasing through law is cases in which debiasing of boundedly rational actors has shown a strong likelihood of success in the existing social science literature. The illustrations of debiasing through substantive law discussed in Part II all fit this basic pattern.

II. Debiasing Through Substantive Law

In this Part we fill out the general theory described in Part I by developing a set of organizing examples of debiasing through substantive law. We show how this idea can or does address basic choices the legal system must make across a wide range of legal domains.

The discussion is organized by reference to the three distinct subcategories of bounded rationality described in Part I.A above. We begin with judgment errors – racial or other group-based unconscious bias and judgment errors described in the heuristics-and-biases literature – and then consider departures from expected utility theory. In each area we start by describing the social science findings on the effects of debiasing of boundedly rational actors and then turn to the heart of our analysis in this Part – tracing both descriptive and prescriptive implications for debiasing through substantive law.

A. Debiasing Through Substantive Law in the Context of Racial or Other Group-Based Unconscious Bias

Building on the large psychology literature on racial or other group-based unconscious bias noted above, legal commentators writing on employment discrimination law have recently explored the vexing question

⁷³ Nilanjana Dasgupta & Anthony G. Greenwald, *On the Malleability of Automatic Attitudes: Combating Automatic Prejudice With Images of Admired and Disliked Individuals*, 81 J. PERSONALITY & SOC. PSYCHOL. 800, 803-04 (2001).

of how the law should deal with such bias in the workplace.⁷⁴ The most obvious target of employment discrimination prohibitions, such as Title VII of the Civil Rights Act of 1964, is discrimination that is consciously motivational in nature. But a fairly broad consensus now supports the idea that the most significant problem in today's workplace is unconscious bias, understood as group-based devaluation of which the perceiver is not consciously aware.⁷⁵ An employer might well harbor no racial "animus," and sincerely disclaim and reject prejudice, but nonetheless act in response to subtle bias that affects decisions about hiring, firing, and the conditions of employment.

It is possible to argue that racial or other group-based unconscious bias ("unconscious bias" for the remainder of this section) is accurate as a general matter⁷⁶ – just as are many of the heuristics emphasized by the heuristics-and-biases literature. However, under existing law, employment discrimination that grows out of such bias is a clear wrong,⁷⁷ and, accordingly, we will take it as uncontroversial for purposes of our analysis that reducing such bias is an objective of employment discrimination law.

Legal commentators have responded to the problem of unconscious bias with a variety of suggested reforms. These reforms include replacing the current pretext model of disparate treatment proof with a "motivating factor" analysis that looks to whether bias played *some* role in the decision maker's behavior; permitting a cause of action for "nonwillful" discrimination (although perhaps with lesser remedies than those available for "willful" discrimination); and adopting a negligence approach to discrimination.⁷⁸ These proposed reforms, while worthy of consideration, are striking for

⁷⁴ See, e.g., Krieger, *supra* note 58; Lawrence, *supra* note 17; David Benjamin Oppenheimer, *Negligent Discrimination*, 141 U. PA. L. REV. 899 (1993).

⁷⁵ See, e.g., Krieger, *supra* note 58, at 1164; Oppenheimer, *supra* note 74, at 900-17; Susan Sturm, *Second Generation Employment Discrimination: A Structural Approach*, 101 COLUM. L. REV. 458, 460, 468-69 (2001).

⁷⁶ See Amy L. Wax, *Discrimination as Accident*, 74 IND. L.J. 1129, 1142-43 (1999).

⁷⁷ See David A. Strauss, *The Law and Economics of Racial Discrimination in Employment: The Case For Numerical Standards*, 79 GEO. L.J. 1619, 1623 & n.13 (1991).

⁷⁸ See Krieger, *supra* note 58, at 1241-47; Oppenheimer, *supra* note 74, at 967-72.

taking the fact of unconscious bias entirely as a given, rather than as something itself to be examined and possibly reduced or even undone by the law. As Linda Hamilton Krieger, one of the central contributors to the literature, expresses the assumption, “[W]e [don’t] know enough about how to reduce cognition-based judgment errors to enable us to translate such a duty into workable legal rules [because c]ognitive psychologists have told us more about the shortcomings of human social inference cognition than about how the various biases they identify can be reduced or controlled.”⁷⁹

Our analysis here is focused on an alternative possibility – that both existing aspects and possible reforms of Title VII might attempt to strike at the underlying problem of unconscious bias by seeking directly to debias employment decision makers. Notably, we suggest that some important pockets of current employment discrimination law actually attempt to engage in such debiasing. That the bias in question is unconscious does not mean that it is not changeable or manipulable. A substantial body of recent social psychology literature strongly questions “the assumption that automatic processes are inflexible and impervious to the perceiver’s intentions and goals.”⁸⁰ Evidence suggests that “both tacit and expressed social influence reduce[] the expression of automatic prejudices.”⁸¹ Exploring whether the law can or does help to reduce such unconscious bias directly is an especially urgent task given the barriers confronting the reform proposals advocated in the existing literature.⁸²

Two preliminary comments are important to our discussion here. First, our focus is on unconscious bias that produces unlawful discrimination, not on unconscious bias as such. Some people would be troubled by the idea that government does or should attempt to alter people’s unconscious attitudes toward racial or other groups simply because it rejects those attitudes; we are concerned here with the narrower question of how to prevent discriminatory conduct in the workplace.

⁷⁹ Krieger, *supra* note 58, at 1245; *see also* Wax, *supra* note 76, at 1133 (similar).

⁸⁰ Irene V. Blair, *The Malleability of Automatic Stereotypes and Prejudice*, 6 PERSONALITY & SOC. PSYCHOL. REV. 242, 243 (2002).

⁸¹ Brian S. Lowery, Curtis D. Hardin & Stacey Sinclair, *Social Influence Effects on Automatic Racial Prejudice*, 81 J. PERSONALITY & SOC. PSYCHOL. 842, 842 (2001).

⁸² *See, e.g.*, Barbara Reskin, *Imagining Work Without Exclusionary Barriers*, 14 YALE J.L. & FEMINISM 313, 329 (2002).

Second, while our analysis will point to how both current employment discrimination law and possible reforms can facilitate debiasing through substantive law, our claim is not that the law does or could eliminate unconscious bias entirely. Some sources of such bias – such as the very concept of “groupness,” which leads individuals to perceive members of their group as more similar to them and members of different groups as more different from them⁸³ – may not be amenable to debiasing through employment discrimination law. Our claim is only that employment discrimination regimes carry the potential – already realized to some degree – to achieve some important measure of debiasing through substantive law.

1. Debiasing Through Workplace Diversity Rules

- a. Social science evidence. An intriguing set of results in the social science literature demonstrates that the composition of the leadership group in the workplace often shapes the degree of unconscious bias workers exhibit. The studies suggest that role models or authority figures in the individual’s environment have a significant effect on the degree of unconscious bias as measured by the Implicit Attitudes Test (IAT) described in Part I.A above.

One notable study, for instance, showed that participants who were administered an in-person IAT by an African-American experimenter exhibited substantially less unconscious racial bias than subjects who were administered an in-person IAT by a white experimenter.⁸⁴ In other words, subjects’ speed in categorizing black-bad and white-good pairs was closer to their speed in categorizing black-good and white-bad pairs when an African-American experimenter was presiding than when a white experimenter was presiding.

A second study paired white test subjects with African-American experimenters and assigned the pair a task in which (1) the white participant evaluated the African-American experimenter; (2) the white participant needed to cooperate with the African-American experimenter; or (3) the white participant was evaluated by the African-American experimenter. White participants who were told to evaluate the African-American

⁸³ See Krieger, *supra* note 58, at 1186-88.

⁸⁴ Lowery, Hardin & Sinclair, *supra* note 81.

experimenter subsequently exhibited the greatest degree of unconscious racial bias, while white participants who were evaluated by the African-American experimenter demonstrated the least unconscious racial bias (all measured again by speed of categorization of stereotypical versus counter-stereotypical pairs).⁸⁵

b. Implications for employment discrimination law. The results just described point toward an empirical foundation for the commonsensical idea that having a diverse supervisory workforce may well reduce the degree of racial or other group-based bias – including unconscious bias – in the workplace. A straightforward causal path here is that, compared to supervisory figures who are not members of a particular group, supervisory figures from the group will be less likely to harbor unconscious bias against group members with whom they interact.⁸⁶ Consistent with this suggestion, a recent empirical study by labor economists finds that African-American hiring officers hire a significantly greater proportion of African-American applicants than do white hiring officers.⁸⁷ But the more subtle causal path – and the one we wish to add to the scholarly and policy discussion – is that the simple presence of supervisory figures from the group in question may reduce the degree of unconscious bias exhibited in the workplace by non-group members. If someone in authority in my workplace is African-American, the evidence suggests that I will tend to exhibit less unconscious racial bias at the workplace than if everyone in authority there is white. In a sense, the story here is the unconscious-bias counterpart to the old idea that contact with members of other groups should reduce conscious animus.⁸⁸ With respect to the modern situation of unconscious bias, unlike in the social science studies discussed above – where a real worry is that the decline in unconscious bias lasts only for the time of the experiment or a

⁸⁵ Jennifer A. Richeson & Nalini Ambady, *Effects of Situational Power on Automatic Racial Prejudice*, 39 J. EXPERIMENTAL SOC. PSYCHOL. 177 (2003).

⁸⁶ IAN AYRES, *PERVASIVE PREJUDICE? UNCONVENTIONAL EVIDENCE OF RACE AND GENDER DISCRIMINATION* 419-425 (2001), offers an engaging and insightful discussion.

⁸⁷ Michael A. Stoll, Steven Raphael & Harry J. Holzer, *Black Job Applicants and the Hiring Officer's Race*, 57 INDUS. & LAB. REL. REV. 267 (2004).

⁸⁸ For a discussion of the contact hypothesis, and a suggestion that it does not hold in the context of the mentally ill, see Elizabeth Emens, *The Sympathetic Discriminator: Mental Illness and the ADA* (unpublished manuscript, on file with authors).

short period longer – in the workplace the presence of a diverse supervisory workforce would be a long-term feature of the environment.

Thus, a logical path for debiasing through employment discrimination law is debiasing through the achievement of a diverse supervisory workforce. How might that be accomplished? The difficulty here is that precisely the existence of the unconscious bias in the first instance suggests the profound barriers standing in the way of achieving a diverse supervisory workforce. The controversial implication, then, is that debiasing through a diverse supervisory workforce may require special *short-term* efforts to achieve diversity among authority figures in the workplace. Title VII's focus on fair treatment of individuals and its disavowal of any mandatory preferences for particular groups – however sensible these twin measures may be from the perspective of solving the problem of conscious employer animus – are simply not apt when the problem is that the very environment and composition of the workplace structure and affect the degree of unconscious bias employment decision makers exhibit. If the degree of such bias is substantial without diverse authority figures and is noticeably lower with them, then a short-term strategy that emphasizes achieving diversity as a central objective may influence the processes that otherwise create the need for an ongoing, long-lived, and invasive program of government regulation of employment relationships, as many believe we currently have under Title VII.

Of course, to some degree government policy in the employment discrimination area already reflects the sort of debiasing through substantive law that we are describing here. At all levels of government, public officials have chosen to adopt affirmative action plans governing public sector employees.⁸⁹ These initiatives reflect a limited form of the debiasing strategy described here; the plans ensure that there is meaningful diversity across ranks – including supervisory ranks – in the workforce. (Achieving diversity in non-supervisory ranks may be a necessary first step in creating the pool needed to achieve diversity in supervisory ranks.) Many institutions, private and public, are aware of the value of diverse supervisors as a means of reducing discrimination and improving morale; indeed, when the Supreme Court upheld narrowly drawn affirmative action programs in the educational setting, it did so with reference to the arguments of the United States military,

⁸⁹ See, e.g., *Johnson v. Santa Clara Transportation Agency*, 480 U.S. 616 (1987).

stressing points in the same family as those we are offering here.⁹⁰ The military's efforts to ensure diversity in its officer corps have been spurred in part by a desire to ensure unity in the ranks and to promote morale among African-American soldiers.⁹¹

We are not suggesting that private employers, as a matter of employment discrimination law, should be required to take definitive steps to achieve diverse supervisory workforces. But in terms of existing law, we can take the courts' willingness to allow voluntary affirmative action with respect to supervisory workers at private firms and to some extent in public institutions⁹² as responsive (whether or not intentionally) to the underlying impetus for debiasing of unconscious bias through a diverse supervisory workforce.

3. Debiasing Through Regulation of the Workplace Environment

a. Social science evidence. The composition of the leadership group in the workplace is not the only factor that plays a role in determining the degree of unconscious bias in the workplace. Social science research suggests that unconscious bias may also be reduced by the promotion of counter-stereotypes or elimination of negative stereotypes in the physical or sensory surroundings. In one study, for instance, participants who spent five minutes creating a mental image of a strong woman showed markedly reduced levels of unconscious bias against women.⁹³ In another study, exposure to pictures of counter-stereotypical group members altered unconscious racial bias; participants who were exposed to photographs of Martin Luther King, Jr. and Timothy McVeigh demonstrated less unconscious bias against African-Americans than subjects exposed to

⁹⁰ See *Grutter v. Bollinger*, 539 U.S. 306, 334 (2003). The brief from the United States military offered a range of claims about the importance of affirmative action in the military; its emphasis was on the significance of diverse supervisors to recruitment and morale. See Consolidated Brief of Julius W. Becton et al., *Grutter v. Bollinger* (hereinafter "Military Brief").

⁹¹ See Military Brief, *supra* note 90, at 16-17.

⁹² See *United Steelworkers of America v. Weber*, 443 U.S. 193 (1979) (private firms); *Grutter*, 539 U.S. 306 (public institutions).

⁹³ See Irene V. Blair, Jennifer E. Ma & Alison P. Lenton, *Imagining Stereotypes Away: The Moderation of Implicit Stereotypes Through Mental Imagery*, 81 J. PERSONALITY & SOC. PSYCHOL. 828 (2001).

photographs of O.J. Simpson and John F. Kennedy.⁹⁴ The results of these studies may be a testimonial to the power of the availability heuristic; they also suggest a more general role for the “affect heuristic,” by which decisions are formed by reference to a rapid, intuitive, affective judgment about persons, processes, and activities.⁹⁵ In fact the affect heuristic seems likely to play a role in employment discrimination; and evidence suggests that affect can be altered – providing an opening suggestion of the prospects for debiasing through substantive law in this area.⁹⁶ The general implication is that the *context* in which one individual views another has enormous significance for the nature and degree of unconscious bias.

One fascinating study makes the role of context particularly explicit. In this study, subjects viewed an Asian woman either putting on make-up or using chopsticks (or neither). Those who viewed the woman putting on make-up exhibited substantially more female stereotypes and fewer Asian stereotypes than the control subjects, whereas those who viewed the woman using chopsticks had the opposite response.⁹⁷ A similar experiment found that test subjects had less automatic negativity toward African-Americans after viewing an African-American face superimposed onto a picture of the inside of a church than did subjects who viewed the same face superimposed onto a picture of a street. Test subjects had a similar, but more muted, reactions to white faces placed in the two contexts.⁹⁸

Moving beyond the social science studies, the basic idea that features of the surrounding context have an important effect on the degree of unconscious bias has a resonance with a point familiar to (and highly evocative for) many university students. Students frequently take notice of the portraits of famous scholars or benefactors that adorn classrooms, libraries, offices or other university spaces. In the typical case, the portraits

⁹⁴ Dasgupta & Greenwald, *supra* note 73.

⁹⁵ See Paul Slovic, Melissa Finucane, Ellen Peters & Donald G. MacGregor, *The Affect Heuristic*, in HEURISTICS AND BIASES: THE PSYCHOLOGY OF INTUITIVE JUDGMENT, *supra* note 6, at 397.

⁹⁶ See *id.*

⁹⁷ See C. Neil Macrae, Galen V. Bodenhausen & Alan B. Milne, *The Dissection of Selection in Person Perception: Inhibitory Processes in Social Stereotyping*, 69 J. PERSONALITY & SOC. PSYCHOL. 397 (1995).

⁹⁸ See Bernd Wittenbrink, Charles M. Judd & Bernadette Park, *Spontaneous Prejudice in Context: Variability in Automatically Activated Attitudes*, 81 J. PERSONALITY & SOC. PSYCHOL. 815 (2001).

are predominantly white and male. Many students have a strong experience of these depictions as shaping and reinforcing an environment permeated in subtle ways with various forms of unconscious bias.

Recently at Harvard University, for instance, a study revealed that only a small number of portraits at Harvard depict woman, and most of those hang in Radcliffe buildings; this state of affairs led one student leader to comment, “Although it is a minor detail about our campus, it is really a thing that students internalize.”⁹⁹ “Studying in the Widener Library or eating in Annenberg Dining Hall, Harvard scholars are forever in the company of men – men whose images adorn the University’s walls and the halls.”¹⁰⁰ Similarly, Harvard students who conducted a comprehensive survey of portraits across campus found that only three of 302 portraits were of persons of color, and reported their findings to the Harvard administration with a request for portraits of “persons of African-American, Asian-American, Latino-American and Native American background, who have served Harvard with distinction.”¹⁰¹

The social science research described above cannot resolve these complex controversies, but it does suggest the possibility of significant effects from minimizing stimuli such as the largely white and male Harvard portraits and, more expansively, affirmatively using portraits of important alternative figures, including a larger set of nonwhite and female figures, in their place. Indeed at Harvard a faculty-student committee was recently charged with choosing portraits of racially diverse figures who have played important roles at that institution, and these portraits apparently will be placed at significant sites around the university campus.¹⁰²

A quite similar move is afoot at the United States Capitol. The portrait gallery at the Capitol is “populated almost exclusively by images of white men.”¹⁰³ In response, Senator Christopher Dodd, the senior Democrat on the

⁹⁹ Kate L. Rakoczy, *Men Rule These Walls*, HARVARD CRIMSON, June 5, 2003.

¹⁰⁰ *Id.*

¹⁰¹ Ken Gewertz, *Adding Some Color to Harvard Portraits*, HARVARD UNIVERSITY GAZETTE, May 1, 2003, at 11.

¹⁰² *Id.*

¹⁰³ Sheryl Gay Stolberg, *Face Value at the Capitol: Senator Wants to “Promote Some Diversity” in Congressional Artwork*, N.Y. TIMES, Aug. 13, 2003, at E1.

Senate Rules Committee, recently engineered the move of one of the few portraits of an African-American, Senator Blanche Kelso Bruce, to a spot “just outside the entrance to the visitors’ seats over looking the Senate chamber, in view of the thousands of school children and tourists who pass by each year” – a placement that was “no accident.” Likewise, in 1997 a group of lawmakers and advocates for women persuaded Congress to locate a sculpture of three suffragists – Lucretia Mott, Elizabeth Cady Stanton and Susan B. Anthony – in the grand Rotunda at the Capitol.¹⁰⁴ Republican Senator Olympia J. Snowe commented on the placement: “It really talks about the values of our nation and the premium we place on the role of women in our society. Every time I see that statue, I smile, because I think that’s where they belong.”¹⁰⁵ The effects of such initiatives should not be exaggerated. But the social science research described above suggests that Senators Dodd and Snowe are correct to think that reforms of this kind can have real effects on perceptions of particular groups.

Of course – and a similar point applies to the discussion above of debiasing through a diverse supervisory workforce¹⁰⁶ – it is important that the debiasing here not occur in a heavy-handed or aggressive manner. There would presumably be reasonably broad agreement that Mott, Stanton, and Anthony are sufficiently important figures in our nation’s history to warrant a central location in the Capitol, and indeed that the previous lack of prominent representation of such women may itself not have been a neutral outcome. But an approach that replaced representations of important and well-known white, male figures with representations of obscure people from other groups would be likely to end up backfiring. (Indeed, the recent effort by Senator Dodd with respect to the Capitol portrait gallery led one conservative critic to remark that the new initiative “let political correctness triumph over accurate history.”¹⁰⁷)

¹⁰⁴ *Id.* at E5.

¹⁰⁵ *Id.*

¹⁰⁶ Indeed, an overly aggressive strategy could entrench and increase racial or other group-based unconscious bias rather than reducing it. The influential brief for United States military officers in the *Grutter* litigation took pains to emphasize that voluntary affirmative action programs in the military ensure that all officers are highly qualified. See Military Brief, *supra* note 90, at 29-30.

¹⁰⁷ *Stolberg, supra* note 103, at E1, E5.

b. Implications for employment discrimination law. In the domain of law, the current judicial interpretation of Title VII of the Civil Rights Act of 1964 regulates, to a limited extent, the environmental stimuli that surround employees in the workplace. In the widely cited case of *Robinson v. Jacksonville Shipyards, Inc.*,¹⁰⁸ for instance, the court held that sexually explicit photographs and “pinup” calendars with pictures of nude or partially nude women displayed throughout the work environment could constitute actionable sexual harassment.¹⁰⁹ Similarly, in *Waltman v. International Paper Co.*,¹¹⁰ the court relied on factors including sexually oriented pictures, graffiti, and calendars in the workplace in denying an employer’s motion for summary judgment on a sexual harassment claim.¹¹¹ And of course, sexual jokes that demean women are a mainstay of sexual harassment litigation under Title VII.¹¹²

The social science literature described above on the malleability of unconscious bias provides clear empirical support for the conclusion that these judicial determinations may well result in a form of debiasing through substantive law. Just as the depiction of a Chinese woman using chopsticks leads to a heavy focus on her ethnicity while the depiction of the same woman putting on makeup leads to a heavy focus on her sex, the depiction on a poster in the workplace of (for instance) “a prone nude woman with a golf ball on her breast and a man standing over her, golf club in hand, yelling ‘Fore!’”¹¹³ will significantly affect many employment decision makers’ unconscious views of women. (Indeed, in a study relied upon by the plaintiffs’ expert witness in *Robinson*, men who had viewed a pornographic film just before being interviewed by a woman remembered almost nothing about the interviewer other than her physical characteristics, while men who had watched a regular film before the interview remembered the interview’s

¹⁰⁸ 760 F. Supp. 1486 (M.D. Fla. 1991).

¹⁰⁹ *Id.*

¹¹⁰ 875 F.2d 468 (5th Cir. 1989).

¹¹¹ *Id.*; see also *Andrews v. City of Philadelphia*, 895 F.2d 1469 (3d Cir. 1990); *O’Rourke v. City of Providence*, 235 F.3d 713 (1st Cir. 2001); *Barbetta v. Chemlawn Services*, 669 F. Supp. 569 (W.D.N.Y. 1987).

¹¹² See, e.g., *Robinson v. Jacksonville Shipyards, Inc.*, 760 F. Supp. 1486, 1504-05 (M.D. Fla. 1991).

¹¹³ *Rabidue v. Osceola Refining Co.*, 805 F.2d 611, 624 (6th Cir. 1986) (Keith, J., dissenting), *overruled on other grounds*, *Harris v. Forklift Sys.*, 510 U.S. 17 (1993).

content.¹¹⁴) In the words of the *Robinson* court, “The availability of photographs of nude and partially nude women...may encourage a significant proportion of the male population in the workforce to view...women workers as if those women are sex objects.”¹¹⁵ Note that it is certainly possible for those *consciously* biased against women, as well as those unconsciously biased against them, to be affected by physical and environmental stimuli in the workplace. The social science research suggests that removing such demeaning depictions of women is likely to reduce the degree of unconscious bias against women in the workplace.

It is important to emphasize that in the existing judicial treatment of sexually explicit material in the workplace, a work environment featuring such material does not become legally actionable unless the environment alters women’s “terms, conditions, and privileges of employment,” and thus constitutes employment discrimination.¹¹⁶ We do not suggest that the sexual harassment cases find, or should find, actionable employment discrimination solely because demeaning depictions of women may increase unconscious bias in male supervisors or co-workers, apart from the direct effect of these depictions on women’s conditions of employment. (An obvious example would be demeaning depictions of women that are viewed privately by one or more male employees out of view of female employees¹¹⁷; we do not suggest that such occurrences by themselves can support or help to support a finding of unlawful employment discrimination.) Any effort to punish potential thoughts or attitudes apart from a finding of conduct constituting employment discrimination would raise serious first amendment difficulties. In *Johnson v. County of Los Angeles Fire Department*,¹¹⁸ for instance, the court held unconstitutional on first amendment grounds a fire station policy categorically banning sexually explicit magazines in the workplace, including in private areas such as employees’ lockers. The court said that government

¹¹⁴ Doug McKenzie-Mohr & Mark P. Zanna, *Treating Women as Sexual Objects: Look to the (Gender Schematic) Male Who Has Viewed Pornography*, 16 PERSONALITY & SOC. PSYCHOL. BULL. 296 (1990).

¹¹⁵ *Robinson*, 760 F. Supp. at 1503; *see also Barbeta*, 669 F. Supp. at 573 (stating that the proliferation of sexually explicit material “may be found to create an atmosphere in which women are viewed as men’s sexual playthings...”).

¹¹⁶ *See Harris v. Forklift Sys.*, 510 U.S. 17 (1993).

¹¹⁷ *See Jack M. Balkin, Free Speech and Hostile Environments*, 99 COLUM. L. REV. 2295, 2316-17 (1999)

¹¹⁸ 865 F. Supp. 1430 (C.D. Cal. 1994).

action “may not proscribe the communication of ‘sex-role stereotyping’ simply because [it] disagree[s] with the message.”¹¹⁹

It is equally clear, however, that prohibiting conduct that is found to alter women’s conditions of employment in violation of Title VII does not become a first amendment violation simply because the prohibition may *also* accomplish a form of debiasing through substantive law.¹²⁰ While the first amendment does not ordinarily allow government to regulate speech on the ground that people will be persuaded by it,¹²¹ the usual view, defended most prominently by Jack Balkin and Richard Fallon, is that the first amendment does not forbid the regulation of employment discrimination even when the regulation takes the form of forbidding speech that is discriminatory because of its viewpoint or content.¹²² For instance, the government unquestionably may, consistent with the first amendment, impose civil damages on an employer who says, “you’re fired because you’re a woman,” or who repeatedly tells a female employee that she should be posing for Playboy and directs her attention to Playboy posters in the workplace. In either case the purpose of the law is to prevent an act of employment discrimination, and the acts in both of these cases can legitimately be defined as such. Like laws that forbid battery and assault, or for that matter purely verbal threats, those that ban discrimination not only have the effect of forbidding certain conduct (including words, such as “you’re fired,” that amount to conduct), but also the effect of altering perceptions and attitudes. No plausible interpretation of the first amendment suggests that such laws are unconstitutional for that reason.

Our basic claim here is that some acts of employment discrimination also fuel unconscious bias and hence further employment discrimination. The social science evidence suggests that existing employment discrimination

¹¹⁹ *Id.* at 1441.

¹²⁰ For treatments of the constitutionality of existing sexual harassment doctrine under Title VII, *see* Balkin, *supra* note 117; Richard Fallon, *Sexual Harassment, Content-Neutrality, and the First Amendment Dog That Didn’t Bark*, 1994 Sup. Ct. Rev. 1. As Professor Fallon describes, because the first amendment question was fully briefed in *Harris*, the Court’s upholding of Title VII liability in that case seems to point to a position on the first amendment question. *See* Fallon, *supra*, at 1, 5-7.

¹²¹ *See generally* David Strauss, *Persuasion, Autonomy, and the Freedom of Expression*, 91 COLUM. L. REV. 335 (1991).

¹²² *See* Balkin, *supra* note 117, at 2217-18; Fallon, *supra* note 120, at 21-51.

doctrine, by regulating the physical and environmental stimuli displayed in the workplace, accomplishes a form of debiasing through substantive law.

More controversially, the social science research described above suggests the promise of broader forms of debiasing through employment discrimination law, not limited to the removal of affirmatively demeaning and offensive material as under current Title VII law. We do not mean to recommend imposing affirmative liability under Title VII for maintaining (for instance) portraits that fail to include a diverse set of figures; such behavior cannot reasonably be said to support a finding of employment discrimination. But from the perspective of reducing the unconscious bias in the workplace, it might be desirable to make an employer's positive effort to portray diversity in the physical environment an express factor weighing against employers' vicarious liability under Title VII – in just the way that, under current Title VII law, employers regularly defend against such liability for their agents' discriminatory conduct on the basis of actions such as manuals or training videos disseminated in the workplace.¹²³

Our basic suggestion is that the existing Title VII approach to employers' vicarious liability for their agents' discriminatory acts might be extended beyond the specific context of the discrete mechanisms (manuals, handbooks, videos, internet instructional programs) contemplated by present law – at least if doing so remains consistent with the first amendment (a question we do not attempt to resolve here). If depictions of individuals in the workplace have the significant effects that social science research suggests, and if reducing unconscious bias is a desirable objective of employment discrimination law, then it may make a good deal of sense for that body of law to attempt to take such effects into account as one factor in determining employers' vicarious liability under Title VII. Indeed, even without a change in law, it is possible, as Susan Sturm has recently suggested, that the structures set in motion through existing Title VII doctrines governing employers' vicarious liability will lead employers to take steps to increase the diversity of physical and other environmental stimuli in the workplace.¹²⁴ Our emphasis here is instead on direct means of accomplishing debiasing through employment discrimination law.

¹²³ See *Faragher v. City of Boca Raton*, 524 U.S. 775 (1998); *Kolstad v. American Dental Assn.*, 527 U.S. 526 (1999).

¹²⁴ See generally Sturm, *supra* note 75.

B. Debiasing Through Substantive Law in the Context of Judgment Errors from the Heuristics-and-Biases Literature

Judgment errors from the heuristics-and-biases literature are central to the psychology, behavioral economics, and legal literatures on bounded rationality. In response to these errors, the present section describes a series of applications of debiasing through substantive law, in areas ranging from consumer safety to wrongful discharge from employment to the composition of corporate boards.

1. Debiasing Through Consumer Safety Law

A vast number of federal and state laws regulate the safety of products used by consumers.¹²⁵ A major impetus for these laws is the belief that consumers often do not adequately understand the potential risks of such products. Consumers may not adequately understand such risks because they are imperfectly informed, because they suffer from bounded rationality – most familiarly because of the phenomenon of optimism bias described in Part I.A above – or both.

The traditional law and economics view of the consumer safety context is that the problem (if there is one at all) merely involves imperfect information, and thus is appropriately corrected by provision of additional information.¹²⁶ However, as the earlier discussion of optimism bias suggested and as Jon Hanson and Douglas Kysar, among others, have emphasized, optimism bias will lead many consumers to underestimate their personal risks even if they accurately understand average risks.¹²⁷ To be sure, optimism bias is context-dependent.¹²⁸ But the factors that tend to reduce the extent of the

¹²⁵ See, e.g., 15 U.S.C. §§ 2601-2692 (2000) (Toxic Substances Control Act); 15 U.S.C. §§ 2051-2084 (2000) (Consumer Product Safety Act); Restatement (Third) of Torts: Products Liability (state tort law).

¹²⁶ See, e.g., JOSEPH E. STIGLITZ, *ECONOMICS OF THE PUBLIC SECTOR* 90-91 (1986).

¹²⁷ See, e.g., Hanson & Kysar, *supra* note 2, at 1511-12; Jon D. Hanson & Douglas A. Kysar, *Taking Behavioralism Seriously: The Problem of Market Manipulation*, 74 N.Y.U. L. REV. 632, 729-30 (1999); Dave Slovic, *Do Adolescent Smokers Know the Risks?*, 47 DUKE L.J. 1133, 1137 (1998).

¹²⁸ See Armour & Taylor, *supra* note 38, at 338-41.

bias – deliberation, close temporal proximity between the decision and the outcome of the decision, and severe consequences of error – are likely to be absent in the consumer safety context, with the possible exception of the severity of consequences.¹²⁹

It would not be unreasonable to conclude that optimism bias justifies heightened standards of products liability as an alternative to the provision of additional statistical facts about the product in question. Hanson and Kysar, for instance, argue in favor of enterprise liability on the basis of boundedly rational behavior such as optimism bias.¹³⁰ However, such an approach – seeking to “debias law” – imposes large costs of its own.¹³¹ A still more aggressive “debiasing law” approach, available under existing law in the case of some products, is an across-the-board ban on the product’s use. A number of federal statutes give agencies a choice among disclosure requirements and partial or complete bans.¹³² In response to evidence of inadequate information, optimism bias, and other consumer errors, some regulators might well be tempted to impose a ban even if the statute reflects a preference for disclosure.¹³³

An alternative to these “debiasing law” strategies is to use the law to reduce the occurrence of boundedly rational behavior in the first instance. At the broadest level, strategies for debiasing through consumer safety law provide a sort of middle ground between inaction or the economists’ spare prescription of “more information,” on the one hand, and the aggressive “debiasing law” strategies of heightened products liability standards or outright bans, on the other. Strategies for debiasing through consumer safety law may be far more successful than the mere provision of statistical facts, and also far more protective of consumer prerogatives than the strategy of an across-the-board ban. Our analysis shares a starting point with existing

¹²⁹ Hanson & Kysar, *supra* note 2, at 1511-12, provide further discussion of the role of contextual factors in determining optimism bias in the consumer safety context.

¹³⁰ Hanson & Kysar, *supra* note 2, at 1560.

¹³¹ See, e.g., George L. Priest, *The Current Insurance Crisis and Modern Tort Law*, 96 YALE L.J. 1521, 1560 (1987)

¹³² See sources cited *supra* note 61.

¹³³ See *Corrosion Proof Fittings v. EPA*, 947 F.2d 1201 (5th Cir. 1991) (interpreting Toxic Substances Control Act to require the least restrictive regulatory alternative).

proposals for better “informing” consumers,¹³⁴ but comes to a quite different end point given our empirically-based appreciation of the limits of some forms of information provision, such as those that simply offer general statistical facts.

In our discussion of debiasing through consumer safety law, we will focus on scenarios in which optimism bias is likely to produce an *overall* underestimation by consumers of the risk associated with a given product. In some circumstances, a competing form of bounded rationality could lead consumers to overestimate rather than underestimate the risk associated with a product. For instance, highly available instances of accident or injury can lead to excessive pessimism – a distortion opposite to the one produced by optimism bias.¹³⁵ Unless, however, there is a reason to think that events of accident or injury associated with a given product are highly available – a classic example here is a plane crash¹³⁶ – optimism bias suggests that many consumers will tend toward underestimation of the risks associated with a given product.¹³⁷

How could debiasing through substantive law be accomplished in the consumer safety context?

a. Social science evidence. Straightforward strategies for debiasing in response to optimism bias include suggesting reasons that negative outcomes might occur and considering risk factors related to negative outcomes.¹³⁸ However, such approaches usually fail to reduce optimism bias.¹³⁹ The social science evidence thus suggests that successful strategies for debiasing through substantive law in response to consumer optimism bias will typically require harnessing separate aspects of boundedly rational behavior.¹⁴⁰ Consider two possibilities.

¹³⁴ See generally, e.g., Stiglitz, *supra* note 126, at 90-91.

¹³⁵ See Alan Schwartz & Louis L. Wilde, *Imperfect Information in Markets for Contract Terms: The Examples of Warranties and Security Interests*, 69 Va. L. Rev. 1387, 1437 (1983).

¹³⁶ See Cass R. Sunstein, *Bad Deaths*, 14 J. RISK & UNCERTAINTY 259, 259 (1997).

¹³⁷ E.g., Hanson & Kysar, *supra* note 127, at 729-30.

¹³⁸ See Weinstein & Klein, *supra* note 6, at 322-23.

¹³⁹ See *id.*

¹⁴⁰ See Steven J. Sherman, Robert B. Cialdini, Donna F. Schwartzman, & Kim D. Reynolds, *Imagining Can Heighten or Lower the Perceived*

Debiasing through the availability heuristic. One potential response to the risk that optimistically biased consumers believe “it won’t happen to them” is the availability heuristic described in Part I.A. above. Recall our earlier example of this heuristic: individuals asked how many words in a 2,000-word section of a novel end in “ing” give much larger estimates than individuals asked how many words have “n” as the second-to-last letter.¹⁴¹ As noted earlier, heuristics typically operate through a process of attribute substitution, in which people answer a hard question by substituting an easier one; the basic idea of availability is that individuals attach higher likelihoods to events or outcomes that they have greater ease calling to mind. Use of the availability heuristic often produces a form of judgment error. As with optimism bias, availability can lead to systematic mistakes in the assessment of probabilities. (Thus “availability bias,” in the form of excessively high estimates, and “unavailability bias,” in the form of excessively low estimates, involve complementary errors stemming from the use of this heuristic.) But because making an occurrence “available” can cancel out optimism bias, availability is a promising strategy for debiasing of boundedly rational actors suffering from excessive optimism.¹⁴²

In fact a recent study of smoking behavior finds a phenomenon of exactly this kind. While many smokers appear to be unrealistically optimistic about the risks to them (as opposed to the general population) of smoking, their judgments become more realistic when presented with specific evidence of the health harms that accompany smoking.¹⁴³ In the absence of such evidence, smokers often lack a real sense of harms to quality of life from smoking.¹⁴⁴ Presented with such evidence they have less optimistically biased views about the effects of their behavior.¹⁴⁵

Likelihood of Contracting a Disease: The Mediating Effect of Ease of Imagery, in HEURISTICS AND BIASES: THE PSYCHOLOGY OF INTUITIVE JUDGMENT, supra note 6, at 259.

¹⁴¹ Tversky & Kahneman, *supra* note 28, at 295.

¹⁴² See Sherman, Cialdini, Schwartzman & Reynolds, *supra* note 140, at 259.

¹⁴³ FRANK A. SLOAN, DONALD H. TAYLOR & V. KERRY SMITH, THE SMOKING PUZZLE: INFORMATION, RISK PERCEPTION, AND CHOICE 122-23, 127, 161, 180-81 (2003).

¹⁴⁴ *Id.* at 180-81.

¹⁴⁵ *Id.*

As an example of the basic idea of debiasing through the availability heuristic, consider the finding of Neil Weinstein that many people substantially underestimate their risk of cancer.¹⁴⁶ Imagine that women asked to estimate their risk of breast cancer are also told a poignant and detailed story about a woman their age with similar family and other circumstances who was diagnosed with breast cancer. If so, their estimated probabilities are likely to be higher. (Of course, they may be too much higher or not enough higher – points we discuss at some length in our normative analysis in Part III below). The problem of obesity might be approached in an analogous way. If obesity is in part a product of optimism bias with regard to associated health risks, then debiasing through the availability heuristic would be a natural response.

Debiasing through framing. We have noted the social science evidence showing that many people weigh losses more heavily than gains in evaluating potential outcomes.¹⁴⁷ This evidence suggests that framing the presentation of information to exploit the extra weight attached to losses may counteract bounded rationality in the form of optimism bias.

Consider one well-known illustration of the effects of framing. In a study involving breast cancer risk and breast self-examination, material that describes the positive effects of self-examination – such as a higher chance of discovering a tumor at an earlier stage – is ineffective.¹⁴⁸ By contrast, significant behavioral changes result from material that stresses the negative consequences of failing to undertake self-examination – such as a decreased chance of discovering a tumor when it remains treatable.¹⁴⁹ Thus, if women are optimistically biased about the prospects that they will suffer from breast cancer and hence underestimate the value of engaging in recommended self-examinations, then framing the recommendation to self-examine in terms of losses rather than gains should increase the probability they attach to benefiting from a self-examination.

¹⁴⁶ Neil O. Weinstein, *Unrealistic Optimism About Future Life Events*, 39 J. PERSONALITY & SOC. PSYCHOL. 806, 810 tbl.1 (1980).

¹⁴⁷ See *supra* note 47 and accompanying text.

¹⁴⁸ Beth E. Meyerowitz & Shelly Chaiken, *The Effect of Message Framing on Breast Self-Examination: Attitudes, Intentions, and Behavior*, 52 J. PERSONALITY & SOC. PSYCHOL. 500, 505 (1987).

¹⁴⁹ *Id.*

b. Implications for consumer safety law. The social science findings on both debiasing through the availability heuristic and debiasing through framing point to potentially valuable directions in consumer safety law.

Debiasing through the availability heuristic. In the consumer safety context, debiasing through the availability heuristic would focus on putting at consumers' cognitive disposal the prospect of negative outcomes from use, or at least unsafe use, of a particular product. Specifically, the law could impose a set of requirements on the way that information about the product would be presented to consumers. Firms might be required – on pain of administrative penalties or tort liability – to provide a truthful account of consequences that resulted from a particular harm-producing use of the product, rather than simply providing a generalized warning that fails to harness availability. To enhance the efficacy of this proposed strategy, the law could further require that the real-life story of accident or injury be printed in large type and displayed prominently, so that consumers would be likely to see and read it before using the product. Mandatory warnings could conceivably raise first amendment issues, but so long as there is no political or ideological disagreement with the content of the message such warnings are likely to be constitutional.¹⁵⁰

The evidence suggests that the approach of requiring the specific account as opposed to the generalized warning would help to reduce optimism bias.¹⁵¹ Our point here is similar in spirit to Chris Guthrie's suggestion that legal policy makers bring "vivid information about plaintiff losses in frivolous litigation" to bear in reducing plaintiffs' overestimation of the probability of success in such litigation – an illustration of debiasing through procedural rules.¹⁵² More generally, the guiding idea is that the way information is provided may be just as important as (or more important than) *that* information is provided.

¹⁵⁰ See *Glickman v. Wileman Brothers & Elliott*, 512 U.S. 1145 (1997). The Supreme Court will consider a related set of issues this Term. See *Veneman v. Livestock Marketing Ass'n*, cert. granted 124 S.Ct. 2389 (mem.) (No. 03-1164).

¹⁵¹ See Sherman, Cialdini, Schwartzman & Reynolds, *supra* note 140, at 259.

¹⁵² Chris Guthrie, *Framing Frivolous Litigation: A Psychological Theory*. 67 U. CHI. L. REV. 163, 210 (2000).

It bears noting that an effort to use availability to counteract optimism bias would improve not only the decision making of consumers suffering from optimism bias but also that of consumers suffering from simple information failures. A conspicuous, prominent account of injury from a product may help to correct the estimated probability of harm attached to the product by an optimistically biased consumer. At the same time, it should improve the behavior of imperfectly informed but not necessarily biased consumers.

The effort to exploit availability in the way described here should be modest along two separate dimensions. First, a successful strategy would need to target a limited number of discrete products for which the problem of consumer optimism bias was most important. Consumers would begin to suffer from “information overload” if every time they went to buy any product – from a lawnmower to a candy bar to a fast food hamburger – they were hit with a real-life story of an individual harmed by use or consumption of the product. Their natural response might be to tune out all of the accounts provided by firms, even assuming these accounts were prominently displayed.¹⁵³

Second, the law would need to avoid overreaching in the severity of the featured outcomes. Firms should not be required to provide anecdotes reflecting highly unusual consequences of using their products; only if an outcome occurs with some frequency should the law seek to induce firms to make consumers aware of the prospect. An emphasis on worst-case scenarios might produce excessive responses.¹⁵⁴ If requirements of anecdote-based warnings sweep in extremely unusual or unlikely scenarios, consumers might overreact or lose faith and fail to attach any weight at all to the accounts. Of course there are line-drawing problems here, but the basic point is straightforward.

Note in addition that worst-case scenarios are likely to be much more easily avoided with our suggestion of a legal requirement that firms provide truthful anecdotes about genuine harms than with the alternative strategy – frequently used by government – of public information campaigns concerning risky consumer products. Such campaigns have often resulted in

¹⁵³ See, e.g., W. Kip Viscusi, *Individual Rationality, Hazard Warnings, and the Foundations of Tort Law*, 48 RUTGERS L. REV. 625, 665-66 (1996).

¹⁵⁴ See generally Cass R. Sunstein, *Probability Neglect: Emotions, Worst Cases, and Law*, 112 YALE L.J. 61 (2002).

the use of extremely vivid and salient images, to the point of seriously risking overreaction or even backlash as a result of citizens' perceptions of government "manipulation." In the smoking context, for instance, the European Union has experimented with requirements that a percentage of cigarette packages sold have their fronts covered with vivid pictures of rotting teeth and blackened lungs.¹⁵⁵ And the Canadian Health Ministry has required not only clear warnings ("Cigarettes cause strokes," "Tobacco smoke hurts babies," "Don't poison us," and "Tobacco can make you impotent") but also graphic pictures such as bleeding gums and two lungs with cancerous tumors.¹⁵⁶ Likewise, in the United States a well-known anti-drug advertisement from the 1980s featured a picture of an egg frying in a pan with the voiceover, "This is your brain on drugs."¹⁵⁷

The suggestion of requiring, on pain of administrative sanctions or tort liability, truthful narratives of harm is a more modest and measured response to optimism bias than these approaches, which harness availability by aggressively exploiting highly salient, gripping images and which for this very reason may run an especially high risk of manipulation, overshooting, and other problems. Indeed, in the context of smoking, W. Kip Viscusi has argued at length that individuals do not underestimate, and may well overestimate, the risks from tobacco, where such government approaches have been employed.¹⁵⁸

The idea of requiring firms to provide truthful accounts of harm is not without analogies in current practice. The American Legacy Foundation, a non-profit organization founded out of the 1998 settlement agreements between the United States tobacco industry and state attorneys general, has launched an information campaign employing a close parallel to the strategy outlined here of debiasing through the availability heuristic. The Foundation has publicized parting letters to children and other loved ones from mothers dying of smoking-related diseases; for instance, one letter reads, "Dearest Jon, I am so sorry my smoking will cheat us out of 20 or 30 more years together. Remember the fun we had every year at the lake. I will ALWAYS

¹⁵⁵ See <http://lists.essential.org/pipermail/intl-tobacco/2001q1/000426.html> (visited 9/3/04).

¹⁵⁶ See http://www.abc.net.au/news/science/health/2000/12/item20001224133940_1.htm (visited 9/3/04).

¹⁵⁷ Shaila K. Dawan, *The New Public Service Ad: Just Say "Deal with It,"* N.Y. TIMES, Jan. 11, 2004, §3, at 5.

¹⁵⁸ W. KIP VISCUSI, SMOKING 61-86 (1992).

love and treasure you. Linda.”¹⁵⁹ Our suggested approach reflects much the same spirit.¹⁶⁰

Debiasing through framing. Framing effects also point toward potentially effective methods of debiasing through substantive law in the consumer safety context. Simple requirements that firms “provide information” may be ineffective in this context in part because firms’ interest will be in framing the information in a way that minimizes the risks perceived by consumers. (Recall the shrewd infant formula manufacturers, described in Part I.A, who showed an intuitive appreciation of loss aversion.) By contrast, a legal requirement that firms identify the negative consequences associated with their product or a particular use of their product, rather than the positive consequences associated with an alternative product or with an alternative use of their product, may be an effective means of reducing optimism bias exhibited by consumers. Such a step could make significant progress toward ensuring that consumers have a more accurate understanding of the risks associated with particular products, and could reduce the need for either a complete ban on some of the products in question or other “debiasing law” solutions.¹⁶¹

2. Debiasing Through Wrongful-Discharge Law

Just as both informational failures and bounded rationality in the form of optimism bias often lead to erroneous judgments about the risk associated with consumer products, they can – and often do – lead to erroneous

¹⁵⁹ See <http://women.americanlegacy.org/index.cfm> (visited 9/3/04).

¹⁶⁰ It is possible that, as Douglas Kysar has noted, “debiasing law” strategies such as enterprise liability would give some firms indirect incentives to provide the sorts of truthful accounts we suggest here. See Douglas A. Kysar, *The Expectations of Consumers*, 103 COLUM. L. REV. 1700, 1786 n.364 (2003). Kysar’s suggestion here is similar to Susan Sturm’s idea about employers’ vicarious liability under employment discrimination law. See *supra* note 124 and accompanying text. Once again, though, our emphasis is on more direct strategies for debiasing through substantive law.

¹⁶¹ Again, as discussed further in Part III below, debiasing through harnessing framing effects in the consumer safety context – like its “debiasing law” alternatives – requires calibrating the degree of response to the magnitude and scope of the underlying form of bounded rationality at issue. See *infra* Part III.D.

judgments about the risk of losing a job. With respect to information failures, a substantial body of evidence shows that workers have strikingly inaccurate views about whether firms are legally permitted to terminate their employment absent good cause.¹⁶² Pauline Kim found that in Missouri, for instance, extremely strong majorities of employees – eighty percent or more – believed that the following grounds for discharge, entirely lawful in Missouri, are in fact unlawful: the employer wants to hire someone else to do the same job; the employer mistakenly believes that the employee stole money; or the employer personally dislikes the employee.¹⁶³ Similar results were found in California and New York, notwithstanding substantial variations in the law of the three states.¹⁶⁴ Fundamentally, employees are unaware of the background rule in forty-nine of the fifty United States that employment is “at will” – subject to termination by either party at any time for any reason.¹⁶⁵ But even when – as under some of the employment law reforms discussed in Part IV.A below on debiasing through changes in default rules – employees may have good information from their employee handbooks or manuals about the general permissibility of discharge without good cause, the problem of optimism bias suggests that often they will nonetheless underestimate the probability that they *themselves* will be discharged without good cause.

As in the consumer safety context, a prominent view in the existing legal literature on discharge from employment supports a “debiasing law” response to the problem of optimism bias on the part of employees. Employment law should, on this view, adopt a legal rule requiring, regardless of employee agreement to the contrary, that employers have good cause for terminating their employees.¹⁶⁶ But even many skeptics of at-will employment have been quite reluctant to embrace such a mandatory good cause requirement.¹⁶⁷ As above, our idea of debiasing through substantive law provides a potential middle ground between a legal approach focused

¹⁶² RICHARD B. FREEMAN & JOEL ROGERS, WHAT WORKERS WANT 118-21 (1999); Pauline Kim, *Bargaining With Imperfect Information: A Study of Workers Perceptions of Legal Protection in an At-Will World*, 83 CORNELL L. REV. 105, 133-46 (1997).

¹⁶³ Kim, *supra* note 162, at 133-46.

¹⁶⁴ Pauline Kim, *Norms, Learning, and Law*, 1999 U. Ill. L. Rev. 447, 451.

¹⁶⁵ See Kim, *supra* note 162, at 133-46.

¹⁶⁶ See, e.g., Cornelius J. Peck, *Unjust Discharges From Employment: A Necessary Change in the Law*, 40 OHIO ST. L.J. 1 (1979).

¹⁶⁷ See, e.g., Kim, *supra* note 162, at 154-55.

solely on informing employees of the general rule that employment is at will, on the one hand, and a mandatory good-cause requirement, on the other.

Social science evidence. We have already described the social science evidence that making a negative event available to an individual can counteract the effects of optimism bias.¹⁶⁸ In the wrongful-discharge context, this evidence suggests that individuals are less likely to underestimate the probability that they personally will be discharged from employment without good cause if general statements in employee handbooks and manuals providing that employment is at will are accompanied by a particularized account, from case law or elsewhere, of a recent discharge without good cause.

Implications for wrongful discharge law. The basic idea of debiasing through wrongful-discharge law by harnessing availability to counteract employees' optimism bias is directly parallel to the idea of debiasing through the availability heuristic in consumer safety context. Instead of individualized stories of risk addressed to consumers, the debiasing strategy here would take the form of requiring that at-will provisions in employee handbooks or manuals incorporate truthful accounts, from case law or elsewhere, of discharges without good cause. Thus, for instance, instead of an employee handbook provision stating that the employer "reserves the right to terminate any employee at any time 'at will,' with or without cause,"¹⁶⁹ the handbook would continue with (for instance), "A recent illustration in this state of a termination without good cause was the firing of an employee of Boston Graphics because he was a Yankees rather than a Red Sox fan."

Of course, as discussed above and in Part III below, a potential risk is that the concrete accounts may overcompensate for, rather than merely offset, the effects of employees' optimism bias. Such accounts could lead employees to have an exaggerated sense of their vulnerability. While more empirical evidence would be needed to resolve the question fully, any problem here is limited by the fact that employers writing handbooks or manuals – like the manufacturers in the consumer safety context – have every incentive to choose an account that is not highly salient or dramatic.

¹⁶⁸ See *supra* note 140 and accompanying text.

¹⁶⁹ *Dillon v. Champion Jogbra, Inc.*, 819 A.2d 703 (Vt. 2002).

3. Debiasing Through Corporate Law

A basic question in corporate law concerns the optimal breakdown of board composition between so-called “inside” and “outside” directors.¹⁷⁰ Inside directors are those who are primarily employed by or otherwise closely connected with the corporation; outside directors, by contrast, have no such close link to the firm.

A number of arguments support the inclusion of at least some outside directors on the board.¹⁷¹ Of particular relevance for our purposes is Donald Langevoort’s suggestion that the involvement of such directors may help to overcome optimistically biased judgments (“organizational optimism”) on the part of inside directors.¹⁷² A “debiasing law” solution to the problem of board decisions impaired by optimistically biased inside directors would be to remove these decisions from the hands of the biased decision makers. By contrast, an approach of debiasing through substantive law would take the shape of increasing the number of outside directors on the board. Might legal rules mandating some threshold number of outside directors on the board constitute an effective form of debiasing through substantive law?

a. Social science evidence. If outside directors on corporate boards would help to overcome optimistically biased judgments on the part of inside directors, two things would have to be true. First, outside directors would have to be less subject to optimism bias than inside directors. Second, the involvement of outside directors would have to alter the ultimate group judgment reached by the board members. What do we know about each of these empirical propositions?

¹⁷⁰ For a recent summary of the debate, see Donald C. Langevoort, *The Human Nature of Corporate Boards: Law, Norms, and the Unintended Consequences of Independence and Accountability*, 89 GEO. L.J. 797, 797-99 (2001).

¹⁷¹ See Ronald J. Gilson & Reinier Kraakman, *Reinventing the Outside Director: An Agenda For Institutional Investors*, 43 STAN. L. REV. 863, 873 (1991); Robert C. Pozen, *Institutional Investors: The Reluctant Activists*, HARV. BUS. REV., Jan.-Feb. 1994, at 140.

¹⁷² See Langevoort, *supra* note 170, at 803, 809. Langevoort does not, however, ultimately join those pressing for further increases in outside directors.

The degree of optimism bias exhibited by inside versus outside directors has not been rigorously explored; but the corporate law literature suggests two tentative reasons for believing that outside directors will show a lesser degree of bias. The first is that the selection of outside directors is less likely to be heavily influenced by whether candidates have highly optimistic views of the firms' prospects (in contrast to the case of top executives' selection¹⁷³). The second reason is that outside directors' self-conception and esteem are less closely bundled up with the firm's fortunes.¹⁷⁴

But will some minimum number of outsider directors improve the collective judgment reached by board members? A large body of empirical evidence shows that the probability of erroneous decisions often increases when deliberations are undertaken by like-minded people; those who agree with one another typically end up at a more extreme point in line with their predeliberation tendencies.¹⁷⁵ In the context of corporate boards, the prediction is that optimistic members will lead one another in the direction of further optimism and excessive risk-taking. As a result, boards might well end up more optimistic than the median board member before deliberation began. The mandated inclusion of outside directors then might well serve to check deliberative processes that fuel unrealistically optimistic decisions.¹⁷⁶

b. Implications for corporate law. One requirement of the recently-enacted Sarbanes-Oxley Act is that boards use outside directors to perform all auditing functions – so that a threshold number of outside directors must be named to the board.¹⁷⁷ This requirement can be understood as a form of debiasing through substantive law; the presence of the outside directors responds to the risk of optimism bias on the part of boards stacked with inside directors.

Corporate law governing the structure of the legal-organizational form of the board of directors may be a reasonable way to reduce the degree of optimism bias exhibited by inside directors on boards. Of course, it is

¹⁷³ *See id.* at 809.

¹⁷⁴ *Id.* at 803.

¹⁷⁵ *See generally* CASS R. SUNSTEIN, WHY SOCIETIES NEED DISSENT (2003).

¹⁷⁶ *See id.* Because a solitary actor may find it hard to resist pressure from the remainder of the group, *see supra* note 49 and accompanying text, it might be important to ensure the presence of more than one outside director on the board.

¹⁷⁷ Sarbanes-Oxley Act of 2002, Pub. L. No. 107-204, § 301.

possible that market pressures will impose meaningful constraints on optimism bias from inside directors or the boards they populate.¹⁷⁸ But at the same time, other forces may increase the degree of optimism bias such actors exhibit; these include the process by which managerial executives are selected and the link between these individuals' optimistic judgments and their self-conception and esteem.¹⁷⁹ Of course, wholly apart from the requirements of Sarbanes-Oxley, many boards do contain some outside directors, and this may represent a self-conscious effort by firms interested in (among other things) combating the problem of optimism bias on the part of inside directors. A legal requirement such as Sarbanes-Oxley, however, is likely to facilitate such debiasing on a broader scale, although at a cost of requiring outside directors on all covered boards notwithstanding substantial firm- and industry- specific variation in ideal board structure.¹⁸⁰ We return to this last point in Part III below.

C. Debiasing Through Substantive Law in the Context of Departures from Expected Utility Theory

As noted in Part I.A, the social science literature has focused heavily on strategies for debiasing in response to judgment errors, as distinguished from the other central category of boundedly rational behavior, departures from expected utility theory. Indeed, social scientists have paid little attention to debiasing in response to departures from expected utility theory. The reason may be that such departures are not unambiguous “errors,” and thus it is controversial to say (for example) that the endowment effect, or loss aversion, is a kind of mistake that requires correction. Perhaps for the same reason, “debiasing law” strategies – familiar in the contexts discussed above – are not prominent where departures from expected utility theory are in play. Both because of the normative controversies surrounding departures from

¹⁷⁸ Avishalom Tor, *The Fable of Entry: Bounded Rationality, Market Discipline and Legal Policy*, 101 MICH. L. REV. 482 (2002), provides analysis of the effects and limits of market pressures as a constraining force in the context of firm entry into new industry.

¹⁷⁹ Langevoort, *supra* note 170, at 809; Donald C. Langevoort, *Organized Illusions: A Behavioral Theory of Why Corporations Mislead Stock Market Investors (And Cause Other Social Harms)*, 146 U. PA. L. REV. 101, 140 (1997).

¹⁸⁰ On the importance of such variation, see Langevoort, *supra* note 170, at 815.

expected utility theory and because the relevant social science literature is not well-developed, our analysis in this section is tentative and speculative.

To see the normative issues surrounding departures from expected utility theory most clearly, consider the endowment effect, which says that individuals' willingness to accept – the amount at which they would sell an entitlement – differs from their willingness to pay – the amount they would pay to purchase the entitlement. Is the endowment effect an “error,” or are there valid reasons for the difference in measures of value?

A substantial literature in law addresses these questions.¹⁸¹ Among other things, this literature has identified situations in which, when willingness to accept and willingness to pay differ, the latter is the normatively preferred measure of value. Imagine, for instance, a case in which the endowment effect leads individuals to be excessively attached to an entitlement that ideally should trade fairly freely. The excessive attachment may be a product of bargaining considerations; perhaps people are not willing to trade certain goods because they are acting strategically.¹⁸² The excessive attachment may also stem from the desire to avoid regretting a bad decision to part with an entitlement.¹⁸³ It is also possible that a high willingness to accept reflects learning about the good in questions or a judgment that it is morally questionable to sell certain goods (for instance, environmental amenities) at “any” price,¹⁸⁴ although the latter view may be hard to defend if the relevant amounts can be used to provide or to protect other environmental amenities. In situations in which willingness to pay is in fact the normatively preferred measure of value, eliminating the endowment effect through a lowering of entitlement holders' willingness to accept is a desirable step. Might the approach of debiasing through substantive law help? We describe two illustrations below.

¹⁸¹ See, e.g., Russell Korobkin, Note, *Policymaking and the Offer/Asking Price Gap: Toward a Theory of Efficient Entitlement Allocation*, 46 STAN. L. REV. 663 (1994); Edward J. McCaffery, Daniel J. Kahneman & Matthew L. Spitzer, *Framing the Jury: Cognitive Perspectives on Pain and Suffering Awards*, 81 VA. L. REV. 1341, 1353-54, 1364-73 (1995).

¹⁸² See Russell Korobkin, *The Endowment Effect and Legal Analysis*, 97 Nw. U. L. Rev. 1227, 1243-44 (2003).

¹⁸³ See *id.*

¹⁸⁴ See Kahneman, Knetsch & Thaler, *supra* note 45, at 1327 tbl.1.

1. Debiasing Through the Structure of Property Rights

A fundamental question of property law is whether legal entitlements should be protected by “property rules” or “liability rules.”¹⁸⁵ Under a property rule, entitlement holders are not required to part with their entitlements unless they voluntarily agree (typically in a bargained-for exchange) to do so. Under a liability rule, by contrast, entitlement holders may be forced to give up their entitlements as long as they are paid an agreed-upon amount in damages. When lowering entitlement holders’ willingness to accept to the level of their willingness to pay (in cases which they differ) is desirable, might the choice between property and liability rules play a role?

a. Social science evidence. A preliminary empirical study by Jeffrey Rachlinski and Forest Jourden points to a possible relationship between the divergence between willingness to accept and willingness to pay and the way in which the entitlement being valued is protected from violation.¹⁸⁶ Rachlinski and Jordan’s study finds a marked reduction in the endowment effect, and hence the disparity between willingness to accept and willingness to pay, when liability rules rather than property rules protect the entitlement in question.¹⁸⁷ In the standard endowment effect pattern, willingness to accept is well above willingness to pay when the entitlement is protected by a property rule. But when it is protected by a liability rule, willingness to accept falls to the level of willingness to pay. Rachlinski and Jourden offer an explanation of their results by suggesting that “a right that is protected by a damages remedy might convey less of a sense of ownership than does a right that is protected by an injunctive remedy.”¹⁸⁸ Such incomplete ownership may prevent a perfection of the emotional attachment that is harbinger of the endowment effect.

b. Implications for Liability Rules. As Ian Ayres has suggested, Rachlinski and Jordan’s empirical findings imply that in domains in which lawmakers want to facilitate transactions, liability rules may be

¹⁸⁵ The classic treatment is Guido Calabresi & A. Douglas Melamed, *Property Rules, Liability Rules, and Inalienability: One View of the Cathedral*, 85 HARV. L. REV. 1089 (1972).

¹⁸⁶ Jeffrey J. Rachlinski & Forest Jourden, *Remedies and the Psychology of Ownership*, 51 VAND. L. REV. 1541 (1998).

¹⁸⁷ *Id.*, at 1566.

¹⁸⁸ *Id.* at 1560.

preferable to property rules¹⁸⁹ – although a complete analysis of the choice between property and liability rules involves many additional considerations.¹⁹⁰ Choosing liability rules over property rules may thus be regarded as a form of debiasing through substantive law. Liability rules, under Rachlinski and Jordan’s findings, eliminate the endowment effect by moving individuals’ willingness to accept down to the level of their willingness to pay.

It remains to be seen whether the device of choosing liability rules over property rules would generally have this salutary effect. Of course people are often unaware of how, exactly, their entitlements are protected; if the legal system uses liability rules rather than property rules, many people will not be aware of it. Note also that in Rachlinski and Jordan’s study, the entitlements involved environmental amenities. In that distinctive context, the occurrence of the endowment effect under a property rule may have been “motivated by subjects’ belief that it is improper to sell an environmental resource that one can protect,” while this belief was not triggered under a liability rule “because the law permitted the destruction of the resource for a price.”¹⁹¹ Absent the societal commitment to environmental amenities, for which people often demand a great deal (and on occasion refuse to sell at any price at all),¹⁹² it remains possible that the choice between property and liability rules would not have the same impact on willingness to accept versus willingness to pay.¹⁹³ Further empirical work could help to shed light on this question.

¹⁸⁹ Ian Ayres, *Protecting Property with Puts*, 32 VAL. U. L. REV. 793, 811-12 (1998).

¹⁹⁰ See generally Louis Kaplow & Steven Shavell, *Property Rules Versus Liability Rules: An Economic Analysis*, 109 HARV. L. REV. 713 (1996)

¹⁹¹ Korobkin, *supra* note 182, at 1285.

¹⁹² See Kahneman, Knetsch & Thaler, *supra* note 45, at 1327 tbl.1.

¹⁹³ Daphna Lewinsohn-Zamir, *The Choice Between Property Rules and Liability Rules Revisited: Critical Observations from Behavioral Studies*, 80 TEX. L. REV. 219, 250-57 (2001), argues, but without presenting any direct empirical evidence, that property rules may be preferable to liability rules for reducing the endowment effect in some contexts.

2. Debiasing Through Agency Law

In the context of merchants, as opposed to individuals acting in their private capacities, a strong endowment effect is clearly undesirable, simply because it will make merchants reluctant to sell their goods – the very service they are performing for society. Might debiasing of boundedly rational actors be possible in this domain, and if so is there a role for law to play?

a. Social science evidence. An important empirical paper shows how debiasing of boundedly rational actors exhibiting the endowment effect can occur in sales transactions.¹⁹⁴ Jennifer Arlen, Matthew Spitzer, and Eric Talley find that actors exhibit the endowment effect when they are acting in their ordinary individual capacities, but not when they are acting in the role of corporate managers in a business agency context. Individuals who are instructed that they are acting as agents for the corporation that employs them “manifest[] virtually no endowment effect whatsoever.”¹⁹⁵

What explains these findings? The two most likely explanations involve the business context and the manager’s agency relationship with the firm. Someone working in business – even if acting not as the agent of a corporation but as a sole proprietor of the business – will probably exhibit less of an endowment effect than someone acting in an ordinary individual capacity. One might think, for instance, that transactions conducted by a shoe store – whether or not the store is run by a manager – would not exhibit a large endowment effect. As suggested above, such a store would be unlikely to stay in business long if a strong endowment effect led it – whether acting through a manager or a proprietor – to price shoes at an amount well above people’s willingness to pay.

It seems likely, however, that the agency relationship further dampens the tendency toward exhibiting an endowment effect. This is so because, as suggested earlier, the endowment effect is often linked to a desire by entitlement holders to avoid regretting a bad decision to engage in a transaction (sale of the entitlement).¹⁹⁶ But agents are less likely than ordinary individuals to experience regret because their personal stake in the

¹⁹⁴ Jennifer Arlen, Matthew Spitzer & Eric Talley, *Endowment Effects Within Corporate Agency Relationships*, 31 J. LEGAL STUD. 1 (2002).

¹⁹⁵ *Id.* at 5.

¹⁹⁶ For recent discussions of the general phenomenon of regret, see Camerer et al., *supra* note 11, at 1224-25; Korobkin, *supra* note 182, at 1254-55.

outcomes that occur is lower.¹⁹⁷ Thus, it is reasonable to conjecture that the agency relationship itself helps to account for the finding of an absence of the endowment effect in a business agency relationship – although of course a definitive empirical study separately testing the two explanations could help put the conclusion on firmer ground.

b. Implications for agency law. An implication of this analysis is that basic rules of agency law, in structuring the relationship between corporations and their managers, may achieve a form of debiasing through substantive law for managerial actors. Acting in the context of the agency relationship specified by the law, individuals may not display the endowment effect that has proven robust in other settings.

Of course, in theory the obligations associated with a business agency relationship could be specified by contract rather than by the law of agency. Voluntary arrangements, without the assistance of law, could ensure that much of business is done by agents acting on behalf of firms under a specified set of duties. But, as Frank Easterbrook and Daniel Fischel have suggested in the related context of “standard form” fiduciary duties, such arrangements would usually be inferior to what emerges from a well-functioning system of law.¹⁹⁸ The complexity and nuance of the requirements to be imposed point to the important value of legal rules. Easterbrook and Fischel’s specific focus is the standard form set of fiduciary duties provided by corporate law, but their claim that such standard forms allow for the “elastic contours” that a business relationship requires – and that make ex ante private contracting so difficult¹⁹⁹ – is readily applicable here.

A related point is that the provision by agency law of a legal form has the important effect – by comparison with a default rule specifying no duties of agents to their firms, so that everything would be left to private contracting – of economizing on the substantial transaction costs that would arise with private contracting. Such costs might induce parties not to adopt agency relationships that (among other desirable consequences) reduce endowment effects. Our suggestion here is simply that the basic rules of agency law provide an illustration of how existing rules help to promote debiasing through substantive law.

¹⁹⁷ Korobkin, *supra* note 182, at 1255.

¹⁹⁸ Frank H. Easterbrook & Daniel R. Fischel, *Corporate Control Transactions*, 91 YALE L.J. 698, 702 (1982).

¹⁹⁹ *Id.*

III. Normative Issues In Debiasing Through Law

Debiasing through law may provide a more direct and effective response to problems of bounded rationality than the more typical approach of “debiasing law,” which seeks to insulate legal outcomes from boundedly rational behavior that itself is taken as a given. But as the applications developed in Part II began to reveal, debiasing through law also raises important normative questions of its own. At the most basic level, government, in using strategies for debiasing through law, is often deliberately and self-consciously engaged in altering people’s perceptions of or attitudes toward the world around them. A critical question raised by many versions of debiasing through law is whether and when the government is appropriately involved in this task. This Part sets forth a basic framework within which to evaluate debiasing through law.

A. Two Threshold Questions

1. The Problem of the Second Best

A normative point that is common to approaches of “debiasing law” and “debiasing through law” is that, wholly apart from any legal intervention, a given form of boundedly rational behavior may always be offset by another aspect of bounded rationality that tends in the opposite direction. Simply put, some errors can counteract others. In such cases efforts either to insulate legal outcomes from the effects of the first aspect of bounded rationality (“debiasing law”) or to engage in debiasing through law in response to this aspect of bounded rationality might actually make things worse rather than better – a clear application of the general theory of the second best.²⁰⁰

Whether a given aspect of bounded rationality is in fact likely to be in an offsetting relationship with some other feature of bounded rationality will obviously depend on the particular context.²⁰¹ In our applications in Part II, we focused on situations in which there was no readily apparent counterforce

²⁰⁰ See Gregory Besharov, *Second-Best Considerations in Connecting Cognitive Biases* (unpublished manuscript, on file with authors).

²⁰¹ See, e.g., Jolls, Sunstein & Thaler, *A Behavioral Approach to Law and Economics*, *supra* note 1, at 1524 (discussing the partially offsetting relationship between hindsight bias and optimism bias in the tort law context).

to the aspect of bounded rationality that argued for debiasing through law. Obviously, if an offsetting relationship with another feature of bounded rationality exists, a legal response – whether “debiasing law” or debiasing through law – may well be unwarranted.

2. Attitudes and Values

We will emphasize below that many forms of debiasing through law involve efforts to alter people’s perceptions in response to evidence of what can uncontroversially be viewed as factual errors. But some of the applications in Part II – such as the opening application of debiasing through employment discrimination law – cannot always be put in this category. In some cases, debiasing through employment discrimination law may work by correcting mistaken factual judgments about group members. In others, however, it may work by altering what would more aptly be called “attitudes” or “values” than strictly factual judgments. And the same point holds even more clearly for debiasing through law in response to departures from expected utility theory, discussed in Part II.C. above.

But there should not be any general prohibition on governmental measures that have the effect of altering attitudes or values. Criminal and civil law, frequently designed to forbid force and fraud, have the effect of discouraging people from believing that it is appropriate to engage in force and fraud. In fact shaping people’s beliefs is a large part of the basic point of prohibitions on force and fraud, and hence it would be odd to suggest that government must, with respect to these forms of wrongdoing, remain “neutral” about people’s values. The same can (and has) been said for the basic institutions of a market economy, including freedom of contract and private property. As Albert Hirschman has demonstrated, freedom of contract and its accompanying institutions have important value-shaping purposes and effects. If people see one another as trading partners, they are more likely to cooperate with one another, as captured in the notion of “doux commerce.”²⁰² Hirschman’s account of capitalist institutions shares much with our concept of debiasing through law, as (under Hirschman’s account) socially destructive passions, involving differences of religion and ethnicity, become less important and less damaging under conditions in which people follow their material interests instead.

²⁰² See generally ALBERT HIRSCHMAN, *THE PASSIONS AND THE INTERESTS* (1977).

The case of debiasing through employment discrimination law discussed in Part II.A above is a clear modern analogue to the sort of dynamic just described. Insofar as employment discrimination law forbids racist, sexist, and other group-based discriminatory behavior, it is intended to reduce the extent and effects of certain attitudes and values. Debiasing, in its literal form, lies at the heart of prohibitions on discrimination on the basis of race, sex, and other traits. Of course government targets actions, not what is inside people's heads. It would be possible to insist that the relevant prohibitions are directed against behavior rather than attitudes and values; the law does not direct itself against attitudes and values as such. The point is correct. But as we have emphasized, prohibitions on action affect, and are often intended to affect, attitudes and values as well. If we neglect the value-shaping character of employment discrimination law, we will misconceive its purpose and effect.

When government announces that it will punish employers for their employees' discriminatory behavior, it may encourage, as Susan Sturm has argued, various steps including those that both subtly and significantly shape employees' attitudes²⁰³; but no one suggests that employment discrimination law therefore is inappropriately "manipulative" or otherwise normatively objectionable. As we described in Part II.A.2, both current employment discrimination doctrine governing hostile work environments and our suggested elaboration of current doctrine (governing the determinants of employers' vicarious liability under Title VII) are illustrations of debiasing through law, but they do not for that reason raise any especially novel or vexing normative issues. While debiasing through law may influence people's attitudes and values, that effect does not make this approach normatively distinct from other familiar and widely accepted legal strategies.

B. Correcting Factual Errors

Many forms of debiasing through law involve not alteration of attitudes and values but instead efforts to correct people's erroneous judgments about what would uncontroversially be viewed as questions of fact. Examples include the applications from Part II.B above of debiasing through consumer safety law and wrongful discharge law, where the law seeks to respond to mistaken perceptions of, respectively, the probability of

²⁰³ See generally Sturm, *supra* note 75.

harm from a consumer product and the probability of being discharged from employment without good cause.

With respect to the motivation for legal intervention, debiasing through law in these settings is indistinguishable from a vast array of existing government initiatives. In countless domains, the government either discloses information on its own or requires disclosure by those providing goods or services in response to erroneous perceptions people would otherwise hold.²⁰⁴ When people are committing a clear factual error, there is broad agreement that government may legitimately concern itself with correcting the error. It is hard to think of a plausible objection to this ground for government intervention.

It is important, however, to distinguish among the means government uses to correct factual errors. In some cases, it is possible to correct errors simply by meeting falsehood with fact. Consider the recently-publicized example of government informational campaigns addressing substance abuse by college and high school students. According to a survey by the Harvard School of Public Health, about forty-four percent of college students engaged in binge drinking, defined as five drinks or more in a row for men and four drinks or more in a row for women, in the two-week period preceding the survey.²⁰⁵ Part of the explanation for such behavior appears to be that most students believe that alcohol abuse is far more pervasive than it actually is; because incidents of alcohol abuse are easy to recall, students' perceptions are inflated through the operation of the availability heuristic.²⁰⁶ The resulting bias in estimation is likely to have significant effects for the simple reason that college students are generally influenced by their beliefs about what other college students do, and thus alcohol abuse is likely to increase if students have an inflated perception of other students' drinking behavior.²⁰⁷

²⁰⁴ See, e.g., W. KIP VISCUSI & WESLEY A. MAGAT WITH JOEL HUBER ET AL., *LEARNING ABOUT RISK: CONSUMER AND WORKER RESPONSE TO HAZARD INFORMATION* (1987); Bradley Karkkainen, *Information As Environmental Regulation*, 89 GEO. L.J. 257 (2001).

²⁰⁵ <http://www.hsph.harvard.edu/cas/rpt2000/CAS2000rpt.shtml> (visited 9/2/04).

²⁰⁶ See H. Wesley Perkins, *Introduction*, in *THE SOCIAL NORMS APPROACH TO PREVENTING SCHOOL AND COLLEGE AGE SUBSTANCE ABUSE 8-9* (H. Wesley Perkins ed., 2003) (telling an availability story, although not explicitly using the term).

²⁰⁷ *Id.* at 8-9.

In these circumstances, a natural government response is to emphasize the statistical reality; and this approach has in fact shown substantial success in reducing substance abuse among college and high school students. Montana, for instance, has run an advertisement stating, “Most (81 percent) of Montana college students have four, fewer, or no alcoholic drinks each week.”²⁰⁸ Montana also applies the same basic approach to cigarette smoking, asserting in its advertisements that “Most (70 percent) of Montana teens are tobacco free.”²⁰⁹ The data show that the smoking campaign has produced statistically significant improvements in the accuracy of social perceptions and also statistically significant decreases in smoking.²¹⁰ The point also extends beyond substance abuse to tax fraud. A

²⁰⁸ See Jeffrey W. Linkenbach, *The Montana Model: Development and Overview of A Seven-Step Process for Implementing Macro-Level Social Norms Campaigns*, in THE SOCIAL NORMS APPROACH TO PREVENTING SCHOOL AND COLLEGE AGE SUBSTANCE ABUSE, *supra* note 206, at 182, 195.

²⁰⁹ Jeffrey Linkenbach & H. Wesley Perkins, *MOST of Us Are Tobacco Free: An Eight-Month Social Norms Campaign Reducing Youth Initiation of Smoking in Montana*, in THE SOCIAL NORMS APPROACH TO PREVENTING SCHOOL AND COLLEGE AGE SUBSTANCE ABUSE, *supra* note 206, at 224, 230-33.

²¹⁰ See *id.* Much of the evidence of reduced substance abuse from these studies of the effects of government informational campaigns comes from individuals’ self-reports. The difficulty with such self-reports is that they may move in response to the government information campaign even if underlying behavior is unchanged. One of the studies in this area, however, examines not self-reports of social perceptions and substance abuse but actual arrests for liquor law violations. The study found a forty-six percent drop in arrests, from eighty-four to forty-four per year, after the government information campaign was instituted. See H. Wesley Perkins & David W. Craig, *The Hobart and William Smith Colleges Experiments: A Synergistic Social Norms Approach Using Print, Electronic Media, and Curriculum Infusion to Reduce Collegiate Problem Drinking*, in THE SOCIAL NORMS APPROACH TO PREVENTING SCHOOL AND COLLEGE AGE SUBSTANCE ABUSE, *supra* note 206, at 35, 61. Thus, there is reason to think that the informational campaigns produce actual effects on college and high school students’ behavior.

The apparent success of these government informational strategies is particularly notable given the lack of results associated with other, more familiar strategies to decrease the frequency of substance abuse among

study in Minnesota found that when taxpayers were informed of the high level of tax compliance in the state, they became substantially less likely to cheat – while no effect on compliance levels came from telling them that “your income tax dollars are spent on services that we Minnesotans depend on. Over 30 percent of state taxes go[es] to support education. Another 18 percent is spent on health care and support for the elderly and needy.”²¹¹

The initiatives undertaken by both Montana and Minnesota are straightforward illustrations of government action to replace falsehood with fact. Approaches of this sort should not be controversial. The normative inquiry here does not require investigation of controversial questions about the risk of “manipulation,” government “mind control,” or individual autonomy. No one, for instance, questions consumer credit laws requiring lenders to disclose particular facts such as the total interest payment over the life of the loan, above and beyond the simple interest rate,²¹² on the ground that such requirements might “manipulate.” The real questions are whether the effort at error correction is effective and, if so, whether it is cost-justified.

In many cases, however, government cannot effectively correct errors with strategies that straightforwardly meet falsehood with fact. In the examples of debiasing through consumer safety law, through wrongful-discharge law, and through the law of corporate boards, the government action involves harnessing departures from unbounded rationality to correct errors. This additional complexity raises important and distinctive issues, to which we now turn.

college and high school students. Educational efforts emphasizing the health risks, for instance, have tended to produce small or no effects, as young people often dismiss their own odds of facing serious harm. *Cf.* Armour & Taylor, *supra* note 38, at 334. And attempts to encourage attendance at “alternative social events” on campus have (not too surprisingly) been similarly ineffective. *See* Linkenbach, *supra* note 208, at 196.

²¹¹ STEPHEN COLEMAN, THE MINNESOTA INCOME TAX COMPLIANCE EXPERIMENT STATE TAX RESULTS 5-6, 18-19 (1996), available at <http://www.state.mn.us/ebranch/mdor/reports/compliance/pdf>.

²¹² Fair Credit Reporting Act, 15 U.S.C. §§1681-1681u (2000).

C. Heterogeneous Actors

As recent literature has appropriately emphasized, not all individuals are likely to be boundedly rational, at least not to the same degree.²¹³ With the straightforward strategy of meeting falsehood with fact, heterogeneity among actors is not a major concern. Those who did not err in the first place should not be misled by a message emphasizing fact over falsehood.²¹⁴ But in some cases, a strategy of debiasing through law could introduce new distortions through its effect on those who did not previously err.

Consider the strategic employment of the availability heuristic in response to optimism bias. In this case, the legal intervention might distort the behavior of individuals who did not show bounded rationality in the first place. For those who previously had an accurate understanding of the situation, such strategies for debiasing through law could produce a kind of unrealistic pessimism. In such cases, it is no longer possible to say that, even if the legal intervention does not provide much help, it is unlikely to cause much harm. The same problem arises with respect to debiasing through corporate law in the form of the Sarbanes-Oxley Act. If, among the set of heterogeneous corporate boards, some do not exhibit optimistically biased decision making, then the legally-mandated presence of outsiders on the board could introduce distortions for boards whose behavior was previously undistorted.

In some cases, however, actors who do not suffer from a particular form of bounded rationality, such as optimism bias, will also be free of other forms of bounded rationality, such as reliance on the availability heuristic. If those who are immune from optimism bias also tend to be immune from availability bias, then the strategies described above for debiasing through the availability heuristic should not affect those who did not previously suffer from optimism bias. In such cases, strategies for debiasing through law should not affect those who did not err prior to the legal intervention.

²¹³ See Mitchell, *supra* note 14, at 83-119.

²¹⁴ See Camerer, Issacharoff, Loewenstein, O'Donoghue & Rabin, *supra* note 11, at 1232-35. Thus, for a college student who already believes that most Montana students do not have more than four drinks per week, the informational campaign described above should have no effect – unless it is to correct an *underestimate* of the frequency of college drinking; and a premise of our discussion at this point is that correction of factual errors, whatever their direction, is the government's objective.

When the absence of one form of bounded rationality correlates in this way with the absence of others, strategies for debiasing through law fit with a broader emerging theme in the legal literature on bounded rationality: *Adopt approaches that will correct errors, but without imposing significant costs on those who are unlikely to err.* Colin Camerer, Samuel Issacharoff, George Loewenstein, Ted O’Donoghue, and Matthew Rabin, for instance, have argued on behalf of “asymmetrical paternalism,” that is, a strategy that counteracts errors that reduce welfare, but that does not significantly affect people who did not previously err.²¹⁵ Strategies of this kind are desirable because some people, intuitively or reflectively alert to the risk of bias, can be expected to correct their own errors.²¹⁶ Recall our earlier discussion of two cognitive systems, the heuristic-driven System I and the more deliberative and calculative System II.²¹⁷ People often use their own System II to correct the operation of System I.²¹⁸

To the extent that debiasing through law has no significant effects on those whose behavior is not in need of correction, the approach greatly contrasts with the alternative approach of “debiasing law.” Suppose, for instance, that one responds to optimism bias on the part of employees not by trying to reduce biased judgments but by insulating legal outcomes from the effects of such judgments through a mandatory rule requiring good cause for all discharges from employment.²¹⁹ If so, then one will have altered some legal outcomes that were not in any need of reform at all. The potential contrast with strategies for debiasing through law is clear.²²⁰

In all of our examples of debiasing through law, the government intervention is unlikely to be entirely cost-free for those who did not previously show bounded rationality. But even the provision of statistically

²¹⁵ See *id.* at 1212.

²¹⁶ See, e.g., Jeffrey J. Rachlinski, *The Uncertain Psychological Case for Paternalism*, 97 NW. U. L. REV. 1165, 1211-19 (2003).

²¹⁷ See *supra* notes 59-62 and accompanying text.

²¹⁸ See Kahneman & Frederick, *supra* note 24, at 51.

²¹⁹ See *supra* note 166 and accompanying text.

²²⁰ As Jeffrey Rachlinski has written (in the course of discussing forms of debiasing different from the approach emphasized in this Article), “[G]overnments can adopt measures that restructure decisions as a less intrusive alternative to paternalistic restrictions on choice.” Rachlinski, *supra* note 216, at 1224.

accurate information by government – a wholly uncontroversial strategy in many contexts – will impose costs on those who did not err prior to the intervention simply because of the burden of processing the information. If the intervention produces important benefits for those who are prone to error, then it may be desirable even if it imposes modest costs on others.

The question is the aggregate effect. When there are no significant effects on those not prone to error, intervention is unobjectionable. But even outside of those situations, debiasing through law may be desirable if it improves accuracy for the large majority of people, even if it also decreases accuracy for a few. Here, as in other contexts, the only option is to weigh the effects of the different possible strategies. Of course efforts to debias people through law should be undertaken, whenever possible, in ways that do not produce confusion or misperception, as we discuss more fully in the next subsection.

D. Overshooting and Autonomy

Strategies for debiasing through law that harness other departures from bounded rationality raise two additional concerns that require separate analysis. The first is the risk of overshooting. If truthful narratives are used, people who previously showed optimism bias might be led to exaggerate the risks of consumer products or of discharge from employment without good cause. The effort to debias through law would then be producing biases and errors of its own.

Experimentation would be required to calibrate the degree to which availability or another form of bounded rationality would need to be brought to bear in engaging in debiasing through law – just as, in a conventional “debiasing law” approach, experimentation is necessary to determine the appropriate level or scope of the legal response. The problem of the scope of a legal corrective is ubiquitous in the law, not specific to strategies for debiasing through law.

A second and more fundamental concern involves individual autonomy. In some cases of debiasing through law, government seems to be correcting bounded rationality by exploiting it, in a way that might give rise to fears of manipulation. In the applications discussed in Part II, this occurs most obviously with respect to harnessing availability and framing in

response to optimism bias. Is this a legitimate form of government action? Under what circumstances?

If heuristics and biases are pervasive, then an informed government is likely to have little trouble in manipulating people in its preferred directions. The problem here is that government should respect its citizens, as emphasized, for instance, by the publicity condition in John Rawls's *A Theory of Justice*²²¹; government should not engage in acts that it could not defend in public to those who are subject to those acts. If a public defense could not be made, the acts are an insult to the autonomy of citizens. The publicity condition raises questions about any governmental effort to enlist bounded rationality in the government's preferred directions; and the concern about manipulation is broader still.

If manipulation itself is the focus, then the initial response to the autonomy concern is that, from a behaviorally informed perspective, the worry about government manipulation arises even with the widely accepted approach under which the government corrects simple information failures among citizens. As framing effects as well as other departures from unbounded rationality reveal, there is usually no neutral way to present information. Whenever the government is presenting even accurate information, it is making choices about presentation, choices that will affect how citizens perceive the reality around them. (Public employees who have been subject to retirement options will easily recognize the point.²²²) Thus, it is far too simple, and behaviorally naive, to draw a sharp line between acceptable "provision of information" and unacceptable "mind control." Unless the concern with government manipulation is strong enough to suggest that the government should never provide information to its citizens (an implausible suggestion), there must be some willingness to tolerate the prospect of government influence over citizens' perceptions of reality and the attendant risk of government manipulation.

Thus, for example, if smokers discount the risks that accompany smoking, in part because of optimism bias, it is not obvious that government violates their autonomy by giving a more accurate sense of those risks, even if the best way of giving that accurate sense is through concrete accounts of

²²¹ JOHN RAWLS, *A THEORY OF JUSTICE* 133 (1971).

²²² See Shlomo Benartzi & Richard H. Thaler, *Risk Aversion or Myopia? Choices in Repeated Gambles and Retirement Investments*, 45 *MGMT. SCI.* 364, 375-77 (1999).

suffering. And it is far from clear in such a case that the government could not publicly defend its strategy to citizens as required by the publicity condition; recall in this connection the American Legacy Foundation letters campaign described above.²²³

This is not to say, of course, that no form of debiasing through law would be objectionable on autonomy grounds. Some forms might resemble systems of propaganda in clear violation of the publicity condition. Recall the discussion above of government informational campaigns in response to substance abuse by college and high school students. Some of those who have tried to correct common misimpressions about the frequency of alcohol or cigarette consumption among these populations have gone far beyond anything plausibly described as a presentation of the statistical reality, or even the use of a discrete heuristic, in pursuit of their efforts. Some attempts have involved large-scale public advertising campaigns – complete with campus posters, use of public computer terminals that displayed relevant messages whenever they remained idle for ten minutes, and curricular as well as classroom intervention – and have employed self-consciously one-sided use of information.²²⁴ As the leading advocate of this sort of approach has revealingly and somewhat chillingly put the point: “If one measure of an actual norm is not as positive as we might like, we should consider . . . what other measures might also be available that give a different picture.”²²⁵

Under an approach of this sort, there is a real risk that the one-sidedness and aggressiveness of the government’s effort will be exposed. If this happens, public trust will unquestionably be reduced. And if trust is reduced, government strategies are much less likely to succeed. These instrumental concerns are aggravated by strong moral ones: At least when minors are not involved, the law should treat citizens with respect, and extreme marketing strategies (going well beyond what we have suggested in discussing strategies for debiasing through law) violate that principle. Compare imaginable efforts to control sex discrimination through detailed requirements for public portraits of women or through controlling people’s use of sexually explicit pictures in their homes. At the very least, such efforts might be counterproductive, although it is important to emphasize that the area of discrimination is distinctive, and preference-shaping may be more

²²³ See *supra* note 139 and accompanying text.

²²⁴ See Perkins & Craig, *supra* note 210, at 40-51.

²²⁵ Perkins, *supra* note 206, at 10.

acceptable in this context than in others. But even here it is not unreasonable to fear manipulation.

With respect to autonomy, no general conclusion is likely to make sense; the nature and force of moral objections to public “manipulation” will generally depend on the setting. However, the objection from autonomy seems weakest when government is responding to an identifiable bias that unquestionably qualifies as an “error” or “mistake” and is using methods that do not distort the facts.

E. “Behavioral Bureaucrats”

Nothing said thus far denies the important fact that legal policymakers and administrators, including those who seek to engage in debiasing through law, will often suffer from both inadequate information and bounded rationality. No less than ordinary people, bureaucrats themselves use heuristics and are subject to predictable biases²²⁶; they are also susceptible to the influence of powerful private groups with stakes in the outcome. The combination of informational failures, cognitive biases, and interest-group power can lead government in extremely unfortunate directions.²²⁷ In this light we do not make the naive and implausible suggestion that in the real world, strategies for debiasing through law will always be well-motivated and well-designed. (Nor will their “debiasing law” counterpart strategies.) Our claim is only that if people make mistakes as a result of bounded rationality, and therefore reduce their own welfare, debiasing through law is often a promising response – one that it would be foolish to eliminate from the government’s repertoire.

²²⁶ For a general discussion of “behavioral bureaucrats,” see Jolls, Sunstein & Thaler, *A Behavioral Approach to Law and Economics*, *supra* note 1, at 1543-44.

²²⁷ See, e.g., Roger G. Noll & James E. Krier, *Some Implications of Cognitive Psychology for Risk Regulation*, 19 J. LEGAL STUD. 747 (1990); Timur Kuran & Cass R. Sunstein, *Availability Cascades and Risk Regulation*, 51 STAN. L. REV. 683 (1999).

IV. Debiasing Through Changes in Default Rules

Until this point we have analyzed strategies for debiasing through law in response to generally recognized judgment errors and departures from expected utility theory. An intriguing cousin of such strategies, developed in an exploratory way in this brief Part, is switching legal default rules in cases in which individuals exhibit systematically and pervasively biased judgments about the content of the law. Because this sort of debiasing strategy raises distinctive issues, we outline it here as a way of introducing a novel kind of debiasing through law.

A. A “Legal Fairness” Heuristic?

Biased judgments about the content of legal rules have not been considered in the literature on judgment errors summarized in Part I.A above. But such biased legal judgments similarly lead to systematic errors of the sort potentially conducive – because of their predictability – to effective legal responses.

Consider the question of termination from employment analyzed in Part II.B.2 above. As we noted there, empirical evidence shows that employees systematically overestimate the degree of protection the legal system affords them.²²⁸ In broad terms, these worker misperceptions reflect a problem of inadequate information; but because the errors consistently go in a single direction – an exaggerated sense, by workers, of their legal rights – it may be useful to conceptualize the behavior in question as a form of bias. People often reduce cognitive dissonance by drawing their beliefs about how things should be into line with their views about how things are.²²⁹ This phenomenon suggests a role for what might be termed the “legal fairness” heuristic, in accordance with which judgments about legal rights are influenced by judgments about what the law should be. If workers’ beliefs about what the law is tend to reflect their beliefs about what the law ought to be, then they will sometimes systematically misperceive the law, in a way that reflects their normative judgments.

²²⁸ See *supra* notes 162-65 and accompanying text.

²²⁹ See generally George Akerlof & William T. Dickens, *The Economic Consequences of Cognitive Dissonance*, in GEORGE AKERLOF, AN ECONOMIC THEORIST’S BOOK OF TALES 123 (1984).

The mechanisms here are not entirely clear. For our purposes, what is most important is that individuals' beliefs with respect to their legal rights will show large-scale errors in a predictable direction if a legal rule is in conflict with widespread moral intuitions. Again, there is important overlap here with problems of imperfect information, but that category sweeps much more broadly than the specific phenomenon we are describing here.

B. Switching the Default Rule

A modest response to the situation described above, in which employees misperceive the permissible grounds for termination, would involve a simple shift in the default rule governing discharge from employment. Samuel Issacharoff suggested such a reform some years ago.²³⁰ This shift would have the effect of counteracting employees' misperceptions of their legal rights; employees would have protection against discharge without good cause unless they expressly agreed otherwise. In fact a mild but unmistakable movement in this direction can be found in the judicial decisions taking ambiguous employer statements as creating legal protection against discharge without good cause.²³¹

Thus, for instance, the Wyoming Supreme Court was faced with a case in which an employee handbook described detailed procedures the employer would follow in the event of employee performance problems and stated, on the first page, that the handbook "was not an employment contract."²³² The employee who was threatened with termination without the specified procedures had signed an application form containing the following disclaimer: "READ CAREFULLY BEFORE SIGNING. I agree that any offer of employment, and acceptance thereof, does not constitute a binding contract of any length, and that such employment is terminable at the will of either party, subject to appropriate state and/or federal laws."²³³ The court held that the attempt to disclaim legal limits on termination was ineffective

²³⁰ Samuel Issacharoff, *Contracting For Employment*, 74 TEX. L. REV. 1783 (1996).

²³¹ For a general discussion, see Comment, *Judicial Interpretation of Employee Handbooks: The Creation of a Common Law Information Eliciting Penalty Default Rule*, U. CHI. L. REV. (forthcoming 2004).

²³² *McDonald v. Mobil Coal Producing, Inc.*, 789 P.2d 866 (Wyo. 1990), *rev'd*, 820 P.2d 986 (Wyo. 1991).

²³³ *McDonald v. Mobil Coal Producing, Inc.*, 820 P. 2d 986, 987 (Wyo.1991).

notwithstanding this express statement. This holding, otherwise surprising, makes sense in light of the tendency of employees to take ambiguous employer provisions to fortify their belief that they have a right to be discharged only for good cause, the opposite of the actual regime of at-will employment.

The Model Employment Termination Act (META)²³⁴ takes the basic idea much further. It creates a right to protection from discharge in the absence of good cause, but allows employers and employees to waive the right, on the basis of an agreement by the employer to provide a severance agreement in the event of a discharge not based on good cause.²³⁵ Montana, alone among the fifty states, has adopted a statute that, like META, requires good cause for discharge – but Montana’s requirement is essentially not waivable, and hence that state goes beyond a change in the default rule.²³⁶

Switching the default rule to accord with the assumptions produced by “legal fairness” heuristics is parallel to the approach of debiasing through law emphasized throughout this Article. To be sure, under strategies for debiasing through law, the legal regime attempts to counteract people’s departures from unbounded rationality by altering their judgments to bring them in line with reality. A parallel approach, in the case of biased judgments about the law itself, would be to take steps to alter those biased judgments. Switching the default rule is different. It takes the alternative form of bringing legal reality into line with individuals’ judgments – in the employment context by switching from an at-will to a good-cause default. But the difference should not be overstated. Because only default rules are involved, the effect of the shift may be to change workers’ perceptions – just as in the approach emphasized in the balance of this Article – simply because employers will often respond by unambiguously stating that employment is at will.

In fact, a major effect of decisions that move toward a change in the default rule governing discharge from employment may have been to encourage employers to tell employees, in the plainest terms, that they may be discharged for any reason or for no reason at all.²³⁷ The insistence that

²³⁴ Model Employment Termination Act (1991).

²³⁵ *Id.* § 4(c).

²³⁶ *See* Montana Wrongful Discharge from Employment Act, Mont. Code Ann. §§ 39-2-901 to 914 (2002).

²³⁷ *See, e.g., Reid v. Sears, Roebuck & Co.*, 790 F.2d 453 (6th Cir. 1986).

such disclaimers be exceedingly clear is a response to the fact that, as note above, ambiguity is likely to lead employees to misunderstand most disclaimers.

We do not contend that a finding of employee misperceptions about the content of wrongful-discharge law is sufficient to justify a switch in the default rule governing discharge from employment, or even to support all of the judicial decisions described above. Perhaps arbitrary discharges are rare; if so, then there might be no large problem for the legal system to solve. Conversely, it is possible that a change in the default rule is too tepid a response to the at-will regime. But the approach of switching the default rule charts a middle ground between maintaining the status quo and more aggressive steps such as a mandatory rule requiring good cause for termination of employment.

Switching the default rule in response to biased judgments about the content of the law may also be useful outside the context of wrongful discharge law. Suppose, for instance, that consumers also have an exaggerated understanding of their legal rights. If so, then switching the default rule might be used as a way to protect consumers by altering their perceptions of law.

Conclusion

The central goal of this Article has been to draw attention to the broad importance of debiasing through law. The social science literature has devoted a great deal of effort to the study of debiasing of boundedly rational actors, but with little effort to theorize generally about such debiasing or to see how law and legal institutions might accomplish it. Those interested in bounded rationality and law have argued mostly that legal institutions should be insulated from the effects of boundedly rational behavior, and in some cases that debiasing of boundedly rational actors should be pursued through changes in procedural rules. In our view, debiasing through law – especially debiasing through substantive law – is a distinctive and sometimes far preferable alternative to the strategy of insulating legal outcomes from the effects of bounded rationality. Such debiasing – distinct from the more familiar approach of attempting to control behavior through incentives – often promises to be both more successful and less invasive than the more standard alternatives.

If, for instance, employers often show unconscious bias, steps might be taken – indeed, some important ones already have been – to reduce this bias. Sometimes consumers or employers are too optimistic; the availability heuristic and reframing might be enlisted as correctives. Debiasing through the structure of corporate, property and agency law are further illustrations.

From the normative point of view, we have emphasized that a form of “debiasing” generally accompanies many efforts to forbid particular conduct, even in the most uncontroversial domains of civil and criminal law. In that sense, law is pervasively in the business of “debiasing.” Moreover, many strategies for debiasing through law belong in the same category as more familiar efforts to respond to information failures by providing additional facts. Indeed, many forms of debiasing through law may be seen as a distinctive kind of informational regulation. In many cases, the major questions are standard: whether such efforts are effective and whether their benefits justify their costs.

Nothing in our analysis is inconsistent with the claim that in some contexts unfettered markets are the best response to bounded rationality. Such markets might reduce the effects of bounded rationality by raising the stakes, as noted above²³⁸; it is also possible that the costs of boundedly rational behavior are, in some contexts, lower than the costs of any effort to counteract it. We also do not disagree with the now-familiar suggestion that in the face of bounded rationality, aggressive regulation – some form of “debiasing law” – might sometimes be justified.²³⁹ Instead our aim in this Article has been to chart the possibility of a middle course, one that asks legal institutions not to ignore people, but instead to reduce their errors. In some contexts, debiasing through law is likely to be effective, cost-justified, and minimally intrusive. We believe that numerous areas of the law reveal an appreciation of these points and hence an implicit behavioral rationality, using legal strategies as a mechanism for debiasing of boundedly rational actors. Our principal goal has been to understand those strategies in these terms and to explore the possibility of building on them to do far more.

²³⁸ But for an entertaining demonstration of the persistence of bounded rationality amidst high stakes, see MICHAEL LEWIS, *MONEYBALL* (2003).

²³⁹ Jolls, Sunstein, & Thaler, *A Behavioral Approach to Law and Economics*, *supra* note 1, provide various examples.

Appendix: Debiasing of Boundedly Rational Actors – In General and Through Law

This appendix summarizes some leading empirical research on debiasing of boundedly rational actors and its implications for debiasing through law.

	Evidence of unsuccessful debiasing?	Evidence of successful debiasing?	Illustration of debiasing through procedural rules?	Illustration of debiasing through substantive law?
<i>JUDGMENT BIASES</i>				
1. Racial or other group-based unconscious bias		Authority figure or depictions of admired individuals from group. Sources: Brian S. Lowery, Curtis D. Hardin & Stacey Sinclair, <i>Social Influence Effects on Automatic Racial Prejudice</i> , 81 J. PERSONALITY & SOC. PSYCHOL. 842 (2001); Nilanjana Dasgupta & Anthony G. Greenwald, <i>On the Malleability of Automatic Attitudes: Combating Automatic Prejudice With Images of Admired and Disliked Individuals</i> , 81 J. PERSONALITY & SOC. PSYCHOL. 800 (2001).		Title VII law governing workplace composition and physical environment (Part II.A).

	Evidence of unsuccessful debiasing?	Evidence of successful debiasing?	Illustration of debiasing through procedural rules?	Illustration of debiasing through substantive law?
2. Hindsight bias	<p>Almost all approaches Source: Baruch Fischhoff, <i>Debiasing, in JUDGMENT UNDER UNCERTAINTY: HEURISTICS AND BIASES</i> 422 (Daniel Kahneman et al. eds., 1982)</p>	<p>Telling people not to be “Monday morning quarterbacks.” Source: Merrie Jo Stallard & Debra L. Worthington, <i>Reducing the Hindsight Bias Utilizing Attorney Closing Arguments</i>, 22 J. L. & HUM. BEHAV. 671, 680-81 (1998).</p>	<p>Requiring judges to instruct jurors not to be “Monday morning quarterbacks.” Source: Philip G. Peters, Jr., <i>Hindsight Bias and Tort Liability: Avoiding Premature Conclusions</i>, 31 ARIZ. ST. L.J. 1277, 1286 (1999).</p>	
3. Self-serving bias	<p>Alerting people to existence of the bias Source: Linda Babcock, George Loewenstein & Samuel Issacharoff, <i>Creating Convergence: Debiasing Biased Litigants</i>, 22 L. & SOC. INQUIRY 913 (1997).</p>	<p>Asking people to make other side’s argument. Source: Linda Babcock, George Loewenstein & Samuel Issacharoff, <i>Creating Convergence: Debiasing Biased Litigants</i>, 22 L. & SOC. INQUIRY 913 (1997).</p>	<p>Requiring litigants to make other side’s argument. Source: Linda Babcock, George Loewenstein & Samuel Issacharoff, <i>Creating Convergence: Debiasing Biased Litigants</i>, 22 L. & SOC. INQUIRY 913, 922-23 (1997).</p>	

	Evidence of unsuccessful debiasing?	Evidence of successful debiasing?	Illustration of debiasing through procedural rules?	Illustration of debiasing through substantive law?
4. Anchoring		<p>Prompting people to consider reasons or features different from the anchor. Source: Gretchen Chapman & Eric J. Johnson, <i>Anchoring, Activation, and the Construction of Values</i>, 79 <i>ORG. BEHAV. & HUM. DECISION PROCESSES</i> 115 (1999).</p>	<p>Requiring juries to consider case features other than the dollar award requested in the plaintiff's complaint. Source: Gregory Mitchell, <i>Why Law and Economics' Perfect Rationality Should Not Be Traded for Behavioral Law and Economics' Equal Incompetence</i>, 91 <i>GEO. L.J.</i> 67, 133 & n.207 (2002).</p>	

	Evidence of unsuccessful debiasing?	Evidence of successful debiasing?	Illustration of debiasing through procedural rules?	Illustration of debiasing through substantive law?
5. Optimism bias	<p>Almost all approaches that directly target optimism bias. Source: Neil D. Weinstein & William Klein, <i>Resistance of Personal Risk Perceptions to Debiasing Intervention</i>, in HEURIS-TICS AND BIASES: THE PSYCHOLOGY OF INTUITIVE JUDGMENT 313 (Thomas Gilovich et al. eds. 2003).</p>	<p>Giving people a vivid sense of the relevant harms Source: FRANK A. SLOAN, DONALD H. TAYLOR & V. KERRY SMITH, THE SMOKING PUZZLE: INFORMATION, RISK PERCEPTION, AND CHOICE 122-23, 127, 161 (2003).</p>		<p>Consumer safety law governing warnings (Part II.B.1); wrongful discharge law governing employee manuals and handbooks (Part II.B.2) corporate law governing board composition (Part II.B.3).</p>

	Evidence of unsuccessful debiasing?	Evidence of successful debiasing?	Illustration of debiasing through procedural rules?	Illustration of debiasing through substantive law?
6. Availability heuristic and associated biases		<p>Giving people accurate information Source: Jeffrey Linkenbach & H. Wesley Perkins, <i>MOST of Us Are Tobacco Free: An Eight-Month Social Norms Campaign Reducing Youth Initiation of Smoking in Montana</i>, in THE SOCIAL NORMS APPROACH TO PREVENTING SCHOOL AND COLLEGE AGE SUBSTANCE ABUSE 224 (H. Wesley Perkins ed., 2003).</p>		<p>A familiar strategy is government informational campaigns that “meet falsehood with fact” (Part III.B)</p>

	Evidence of unsuccessful debiasing?	Evidence of successful debiasing?	Illustration of debiasing through procedural rules?	Illustration of debiasing through substantive law?
DEPARTURES FROM EXPECTED UTILITY THEORY				
Endowment effect	Almost all approaches. Source: Daniel Kahneman, Jack Knetsch & Richard Thaler, <i>Experimental Tests of the Endowment Effect and the Coase Theorem</i> , 98 J. POL. ECON. 1325 (1990).	Protecting entitlements by liability rule rather than property rule, having agents make decisions. Sources: Jeffrey J. Rachlinski & Forest Jourden, <i>Remedies and the Psychology of Ownership</i> , 51 VAND. L. REV. 1541 (1998); Jennifer Arlen, Matthew Spitzer & Eric Talley, <i>Endowment Effects Within Corporate Agency Relationships</i> , 31 J. LEGAL STUD. 1 (2002).		Property law governing remedy (Part II.C.1); agency law (Part II.C.2).
OTHER				
Legal fairness heuristic (believing that law fits judgments about what is fair)	Experience in relevant legal setting. Source: Pauline Kim, <i>Norms, Learning, and Law</i> , 1999 U. ILL. L. REV. 447.			Shifting default rules so that individuals have accurate perceptions of legal rules (Part IV).