April 2002

Taking Behavioralism Too Seriously? The Unwarranted Pessimism of the New Behavioral Analysis of Law

Gregory Mitchell

Follow this and additional works at: <https://scholarship.law.wm.edu/wmlr>

Part of the [Jurisprudence Commons](#)

Repository Citation

Gregory Mitchell, *Taking Behavioralism Too Seriously? The Unwarranted Pessimism of the New Behavioral Analysis of Law*, 43 Wm. & Mary L. Rev. 1907 (2002),
<https://scholarship.law.wm.edu/wmlr/vol43/iss5/4>

TAKING BEHAVIORALISM TOO SERIOUSLY? THE UNWARRANTED PESSIMISM OF THE NEW BEHAVIORAL ANALYSIS OF LAW

GREGORY MITCHELL*

ABSTRACT

Legal scholars increasingly rely on a behavioral analysis of judgment and decision making to explain legal phenomena and argue for legal reforms. The main argument of this new behavioral analysis of the law is twofold: (1) All human cognition is beset by systematic flaws in the way that judgments and decisions are made, and these flaws lead to predictable irrational behaviors and (2) these widespread and systematic nonrational tendencies bring into serious question the assumption of procedural rationality underlying much legal doctrine. This Article examines the psychological research relied on by legal behavioralists to form this argument and demonstrates that this research does not support the bleak and simple portrait of pervasive irrationality painted by these scholars. Careful scrutiny of the psychological research reveals greater adherence to norms of rationality than that implied by the legal behavioralists, and the methodological and interpretive limitations on this psychological research make extrapolation from experimental settings to real world legal settings often inappropriate. Accordingly, this Article argues that legal scholars should exercise greater care and precision in their uses of psychological data to avoid advocating further legal reforms based on flawed understandings of psychological research.

* Assistant Professor of Law and Adjunct Professor of Psychology, Michigan State University; B.A. 1988, University of Arkansas; M.A. (Psychology) 1990, J.D. 1993, and Ph.D. (Psychology) 1994, University of California, Berkeley. The author may be contacted by e-mail at mitch328@law.msu.edu or by regular mail at Michigan State University-Detroit College of Law, 416 Law College Building, East Lansing, Michigan 48824. Many thanks to Phil Tetlock, Barb Mellers, Christine Klein, and Glen Staszewski for helpful comments on an earlier version of this Article.

TABLE OF CONTENTS

INTRODUCTION	1909
I. THE EMERGENCE OF LEGAL DECISION THEORY	1913
A. <i>An Assault on the Rationality Assumption</i>	1913
B. <i>The Need for Greater Scrutiny of Legal Decision Theory</i>	1929
II. BEHAVIORAL DECISION THEORY AND ITS INTERPRETIVE LIMITS	1938
A. <i>Features of Behavioral Decision Theory That Mask Individual and Situational Differences in Rational Behavior and Distort Perceptions of the Prevalence of Irrational Behavior</i>	1945
1. <i>Between-Subjects Designs and Emphasis on the "Average Judge"</i>	1946
2. <i>Emphasis on Null Hypothesis Significance Testing: Statistical Significance ≠ Practical Significance</i>	1954
a. <i>Framing Effects</i>	1961
b. <i>The Hindsight Bias</i>	1963
3. <i>The File Drawer Problem</i>	1966
4. <i>Summary: The Aggregation-Generalization Fallacy</i>	1967
B. <i>Features of Behavioral Decision Theory That Increase the Likelihood of Irrational Behavior in Experimental Settings</i>	1971
1. <i>Experiments Designed to Elicit Non-normative Responses</i>	1971
2. <i>Lack of Feedback and Learning Opportunities</i>	1977
3. <i>Conversational Cues and Demand Characteristics</i> ..	1979
4. <i>Use of Ecologically Suspect Formats</i>	1985
5. <i>Deprivation of Decision Tools</i>	1993
C. <i>Features of Behavioral Decision Theory That Limit the Importance of the Research for the Legal System</i>	1995
1. <i>Real-World Success Versus Normative Coherence</i> ...	1996
2. <i>Individual Decision Making Versus Group Decision Making</i>	2002
3. <i>Framing Research Confounds and Confusions</i>	2005
4. <i>Neglect of Systematic Information-Processing Modes</i>	2011
CONCLUSION	2017

Within the last few years, ... we have seen an outpouring of scholarship addressing the impact of behavioral research over a wide range of legal topics. Indeed, one might predict that the current behavioral movement eventually will have an influence on legal scholarship matched only by its predecessor, the law and economics movement.¹

INTRODUCTION

In a series of prominent, recent articles, Jon Hanson and Douglas Kysar argue that the legal system and legal academia should take more seriously behavioral research on judgment and decision making, for this research purportedly reveals that "human decisionmaking processes are prone to nonrational, yet systematic, tendencies."² Hanson and Kysar see this evidence of systematic, nonrational behavior as having great importance for the law: "Ultimately, any legal concept that relies in some sense on a notion of reasonableness or that is premised on the existence of a reasonable or rational decision maker will need to be reassessed in light of the mounting evidence that a human is a 'reasoning rather than a reasonable animal.'"³ Hanson and Kysar predict, moreover, that this "behavioralism" movement will eventually rival law and economics in degree of influence on legal scholarship.⁴

1. Jon D. Hanson & Douglas A. Kysar, *Taking Behavioralism Seriously: The Problem of Market Manipulation*, 74 N.Y.U. L. REV. 630, 634 (1999) [hereinafter Hanson & Kysar, *Behavioralism I*] (footnote omitted).

2. *Id.* at 633. In addition to their 1999 article in the *New York University Law Review*, Hanson and Kysar have published two more articles on this theme of taking behavioral research on decision making more seriously: Jon D. Hanson & Douglas A. Kysar, *Taking Behavioralism Seriously: Some Evidence of Market Manipulation*, 112 HARV. L. REV. 1420 (1999) [hereinafter Hanson & Kysar, *Behavioralism II*]; Jon D. Hanson & Douglas A. Kysar, *Taking Behavioralism Seriously: A Response to Market Manipulation*, 6 ROGER WILLIAMS U. L. REV. 259 (2000) [hereinafter Hanson & Kysar, *Behavioralism III*]. Robert Prentice includes Hanson and Kysar's works among a short list of "important articles ... that demonstrate the value that behavioral insights bring to legal analysis." Robert A. Prentice, *The Case of the Irrational Auditor: A Behavioral Insight into Securities Fraud Litigation*, 95 NW. U. L. REV. 133, 134-35 (2000) (footnotes omitted).

3. Hanson & Kysar, *Behavioralism I*, *supra* note 1, at 634-35 (footnote omitted).

4. *Id.* No single, dominant name for this new field has yet emerged, but "behavioralism"

The prediction of Hanson and Kysar seems well on its way to coming true.⁵ With rapidly increasing frequency, legal scholars employ a behavioral analysis of human judgment and decision making to critique existing laws and legal theory, especially the law and economics movement and its assumption of rational behavior.⁶

and "behavioral law and economics" appear to be the leading contenders. See Hanson & Kysar, *Behavioralism II*, *supra* note 2, at 1425 ("In [the] field [of behavioralism], cognitive psychologists and behavioral researchers study the decisionmaking processes of individuals, with an eye toward comparing actual behavior with that of rationalistic ideals."); Christine Jolls, *Behavioral Economic Analysis of Redistributive Legal Rules*, 51 VAND. L. REV. 1653, 1653 (1998) ("Behavioral law and economics' . . . seeks to bring together 'behavior' and 'law and economics.'"); Cass R. Sunstein, *Introduction to BEHAVIORAL LAW AND ECONOMICS* 1, 9 (Cass R. Sunstein ed., 2000) (referring to the field as "behavioral law and economics"); see also Russell Korobkin, *Inertia and Preference in Contract Negotiation: The Psychological Power of Default Rules and Form Terms*, 51 VAND. L. REV. 1583, 1585 n.8 (1998) (discussing preference for the label "law and behavioral science" for the field); Russell B. Korobkin & Thomas S. Ulen, *Law and Behavioral Science: Removing the Rationality Assumption from Law and Economics*, 88 CAL. L. REV. 1051, 1057 (2000) (referring to this new movement as "law and behavioral science").

5. Within the last five years alone, five law reviews have conducted symposia addressing the question of how psychology and behavioral economics may inform legal decision making. See Symposium, *Behavioral Economics and Labor Law*, 77 N.Y.U. L. REV. (forthcoming 2002); Symposium, *Law, Psychology, and the Emotions*, 74 CHI-KENT L. REV. 1423 (2000); Symposium, *The Legal Implications of Psychology: Human Behavior, Behavioral Economics, and the Law*, 51 VAND. L. REV. 1495 (1998); Symposium, *New and Critical Approaches to Law and Economics (Part I): Behavioral Economics, Law and Psychology*, 79 OR. L. REV. 1, 23-102 (2000); Symposium, *Rational Actors or Rational Fools? The Implications of Psychology for Products Liability*, 6 ROGER WILLIAMS U. L. REV. 1 (2000). Professor Cass Sunstein recently said that the "most exciting work in law schools right now is to try to bring an understanding of how people actually think in contact with the law." Patricia Cohen, *Judicial Reasoning Is All Too Human*, N.Y. TIMES, June 30, 2001, at B7. The details of this emerging scholarship are discussed in Part I.A, and numerous articles in this new field are discussed throughout the Article.

6. See, e.g., Russell B. Korobkin, *Behavioral Analysis and Legal Form: Rules vs. Standards Revisited*, 79 OR. L. REV. 23, 44 (2000) ("Implicit within economic analysis is a set of predictions about how individuals will respond to law, known generally as 'rational choice theory' (RCT). . . . [A] large body of social science literature demonstrates that [RCT's] predictions are not always accurate, and that deviations from so-called 'rational' behavior are often systematic."); Robert B. Thompson, *Securities Regulation in an Electronic Age: The Impact of Cognitive Psychology*, 75 WASH. U. L.Q. 779, 780 (1997) (reporting comments of Professor Cass Sunstein to the effect that this new movement will likely become as important as the law and economics movement); Jennifer Arlen et al., *Endowment Effects Within Corporate Agency Relationships* 1 (Aug. 2001) (University of Southern California Law School, Olin Working Paper No. 00-2, 2001) ("In recent years, behavioral law and economics (BLE) has

Unfortunately, the facile way in which these scholars summarize and then incorporate psychological research findings into legal theory ignores important limitations on this research. These scholars also evince a “pessimism bias” in their work: They tend to ignore or discount research findings contrary to their view of legal decision makers as afflicted by numerous judgmental biases and decision-making errors, while simultaneously interpreting ambiguous research findings as supportive of their pessimistic view of human rationality. Simply put, the empirical research does not support the dire pronouncements of legal scholars regarding the human capacity for irrational behavior. Just as troubling as the overreaching claims about human cognition that these scholars make is their uncritical acceptance by others.⁷

This Article presents the first detailed analysis of legal scholars’ misinterpretations of behavioral decision theory—the body of

risen to prominence within legal scholarship. Its ascension is due (at least in large part) to its aim of unifying two historically distinct analytic paradigms: economics and psychology. While embracing the conventional economic premise that legal rules embody behavior-altering incentives, behavioral law and economics nevertheless disengages from neoclassical economic theory by explicitly analyzing incentives through a psychological lens.”), available at http://papers.ssrn.com/abstract_id=276110.

7. For example, Professor Fanto cites the works of these scholars for the proposition that the uncertainty and pressure from business competitors and rapid technological change has led

executives to enter into mega-mergers on the basis of the kind of decision-making biases that, according to behavioral and psychological research, are typical of human beings reacting in the face of complexity and uncertainty. People often make decisions from little data, or from data that is exemplary, in the foreground or available, but that is not statistically representative.

James A. Fanto, *Braking the Merger Momentum: Reforming Corporate Law Governing Mega-Mergers*, 49 BUFF. L. REV. 249, 288-89 (2001) (footnotes omitted). See Mark A. Lemley & Lawrence Lessig, *The End of End-to-End: Preserving the Architecture of the Internet in the Broadband Era*, 48 UCLA L. REV. 925, 950 (2001) (stating that “[t]he ‘behavioral law and economics’ literature offers strong reason to believe that at least in certain circumstances, people don’t always do the rational thing. Rather, systematic biases can infect decision making.”) (footnote omitted); see also *infra* note 46.

Kyron Huigens, in his review of the bellwether book for this new scholarship, *Behavioral Law and Economics*, *supra* note 4 (collecting essays in which the authors apply psychological research to various legal decision-making topics), accepts the basic contention that “real people do not behave at all in the way that economists have supposed they do.” Kyron Huigens, *Law, Economics, and the Skeleton of Value Fallacy*, 89 CAL. L. REV. 537, 537 (2001) (book review) (footnote omitted). Professor Huigens criticizes this new scholarship not for unwarranted empirical claims but rather for maintaining a consequentialist view of behavior. *Id.* at 568 (“The difficulty is that they still cling to the tools of consequentialism.”).

empirical research on which this new field so heavily relies.⁸ An examination of the empirical evidence on human judgment and decision making demonstrates that, contrary to the contentions of Hanson, Kysar, and the other legal behavioralists, this body of evidence does not prove that experimental subjects—much less real-world legal decision makers—systematically violate norms of rationality when forming judgments and making decisions. This Article discusses three sets of limitations on behavioral decision theory research that the legal behavioralists fail to appreciate, thereby leading them to incorrect and simplistic conclusions about human rationality: (1) features of this research that mask individual and situational differences in rational behavior and artificially heighten the apparent frequency of irrational behavior; (2) features unique to the experimental research setting that intentionally and unintentionally increase the likelihood of finding irrational behavior and distort perceptions of the frequency of irrational behavior outside of the laboratory; and (3) features of this research that diminish its real-world importance and its ability to provide prescriptive guidance in the law. The Article concludes by discussing possible explanations for why legal scholars have taken behavioral research on judgment and decision making too seriously and too unskeptically, leading to the unwarranted pessimism about human rationality found in so much recent legal scholarship.

When the full range of empirical research into judgment and decision making is considered, and when the methodological assumptions and choices of this research are laid bare, it becomes clear that this body of research does not present an unqualified account of pervasive and systematic irrationality. Rather, this research reveals that some people can be made to appear irrational under some circumstances while many people can be made to appear quite rational under other circumstances. A rational review of the evidence on human judgment and decision making should

8. For a detailed description of behavioral decision theory, see *infra* notes 52-68 and accompanying text. In a companion paper, I discuss the growing body of research demonstrating important individual and situational differences in rational thought and behavior—a body of research that further brings into question the strength and coherence of the behavioral critique of law and economics. See Gregory Mitchell, *Why Law and Economics' Perfect Rationality Should Not Be Traded for Behavioral Law and Economics' Equal Incompetence*, 91 GEO. L.J. (forthcoming 2002).

lead one to agnosticism rather than empirical certainty on the matter of the rationality or irrationality of legal decision making.

I. THE EMERGENCE OF LEGAL DECISION THEORY

A. *An Assault on the Rationality Assumption*

A growing number of scholars assert that legal theory and doctrine should be reformed in light of the research of psychologists and behavioral economists demonstrating that the psychological mechanisms involved in human information-processing often lead to biases and errors in judgment and decision making.⁹

9. See, e.g., Hanson & Kysar, *Behavioralism I*, *supra* note 1, at 634 ("Within the last few years, ... we have seen an outpouring of scholarship addressing the impact of behavioral research over a wide range of legal topics."); Hanson & Kysar, *Behavioralism II*, *supra* note 2, at 1425 ("We argue that, because a multitude of nonrational factors influence individual decisionmaking, consumers cannot be expected to engage in efficient product purchasing analyses—regardless whether manufacturers are required to supply product warnings."); Christine Jolls et al., *A Behavioral Approach to Law and Economics*, 50 STAN. L. REV. 1471, 1473 (1998) ("Our goal in this article is to advance an approach to the economic analysis of law that is informed by a more accurate conception of choice, one that reflects a better understanding of human behavior and its wellsprings."); Russell Korobkin, *The Status Quo Bias and Contract Default Rules*, 83 CORNELL L. REV. 608, 675 (1998) ("This Article has made the positive claim that the preference exogeneity assumption, implicit in all law-and-economics theories of efficient contract default rule selection, is probably false, and the normative claim that its falsity has important implications for efficiency theory."); Korobkin & Ulen, *supra* note 4, at 1059 ("This Article provides an early blueprint for research in 'law and behavioral science,' which we hope will help guide the emerging scholarship in this area. ... We conclude with our vision of how this Article might serve as a basis for a substantial research agenda for the next generation of scholars interested in the confluence of law and behavioral science.") (footnote omitted); Donald C. Langevoort, *Ego, Human Behavior, and Law*, 81 VA. L. REV. 853, 853 (1995) [hereinafter Langevoort, *Ego & Law*] ("This Essay will begin to show how understanding the pervasiveness of egocentric biases can alter our thinking about questions of law and lawyering."); Jeffrey J. Rachlinski, *The "New" Law and Psychology: A Reply to Critics, Skeptics, and Cautious Supporters*, 85 CORNELL L. REV. 739, 739 (2000) [hereinafter Rachlinski, *New Law & Psychology*] ("Recently, legal scholars have become interested in new theories of human decision making that researchers in psychology and empirical economics are developing. These new theories promise to predict people's reactions to law more accurately than either law and economics or traditional legal scholarship.") (footnote omitted); Jeffrey J. Rachlinski, *A Positive Psychological Theory of Judging in Hindsight*, 65 U. CHI. L. REV. 571, 624-25 (1998) [hereinafter, Rachlinski, *Hindsight*] ("The story of the hindsight bias in legal judgments has lessons for legal policymakers. It holds a warning about developing a new area of law that demands a judgment in hindsight. Any new area will have to make some accommodation for the bias."); Cass R. Sunstein, *Human Behavior and the Law of Work*, 87 VA. L. REV. 205, 206 (2001) ("The basic purpose of this Article is to bring a better understanding of human behavior, with the

These scholars contend that proponents of an economic analysis of legal behavior should pay particular attention to this body of research, which is often referred to as "behavioral decision theory,"¹⁰ for it purportedly demonstrates that legal decision makers do not conduct themselves according to the norms of rational behavior assumed by economic theory.¹¹

assistance of cognitive psychology and behavioral economics, to bear on the most basic questions in employment law.") (footnotes omitted).

The most recent articles by no means represent the first efforts to educate legal audiences about the potential importance of the psychological literature on human information-processing for substantive and procedural law and the practice of law. *See, e.g.*, Hal R. Arkes, *Principles in Judgment/Decision Making Research Pertinent to Legal Proceedings*, 7 BEHAV. SCI. & L. 429 (1989); Craig R. Callen, *Hearsay and Informal Reasoning*, 47 VAND. L. REV. 43 (1994); David Cohen & Jack L. Knetsch, *Judicial Choice and Disparities Between Measures of Economic Values*, 30 OSGOOD HALL L.J. 737 (1992); Ward Edwards & Detlof von Winterfeldt, *Cognitive Illusions and Their Implications for the Law*, 59 S. CAL. L. REV. 225 (1985); Robert C. Ellickson, *Bringing Culture and Human Frailty to Rational Actors: A Critique of Classical Law and Economics*, 65 CHI.-KENT L. REV. 23 (1989); Jeffrey L. Harrison, *Egoism, Altruism, and Market Illusions: The Limits of Law and Economics*, 33 UCLA L. REV. 1309 (1986); Thomas H. Jackson, *The Fresh-Start Policy in Bankruptcy Law*, 98 HARV. L. REV. 1393 (1985); Karyl A. Kinsey et al., *Framing Justice: Taxpayer Evaluations of Personal Tax Burdens*, 25 LAW & SOC'Y REV. 845 (1991); Michael J. Saks & Robert F. Kidd, *Human Information Processing and Adjudication: Trial by Heuristics*, 15 LAW & SOC'Y REV. 123 (1980-81); Robert E. Scott, *Error and Rationality in Individual Decisionmaking: An Essay on the Relationship Between Cognitive Illusions and the Management of Choices*, 59 S. CAL. L. REV. 329 (1986); Lee E. Teitelbaum et al., *Evaluating the Prejudicial Effect of Evidence: Can Judges Identify the Impact of Improper Evidence on Juries?*, 1983 WIS. L. REV. 1147. *See generally* Phoebe C. Ellsworth & Robert Mauro, *Psychology and Law*, in 2 THE HANDBOOK OF SOCIAL PSYCHOLOGY 693-717 (Daniel T. Gilbert et al. eds., 4th ed. 1998) (reviewing social-psychological research on legal decision making); Paul J. Heald & James E. Heald, *Mindlessness and Law*, 77 VA. L. REV. 1127, 1127-30, 1131-37 (1991) (reviewing the several attacks that had already been made by 1991 on the behavioral assumptions underlying law and economics); Donald C. Langevoort, *Behavioral Theories of Judgment and Decision Making in Legal Scholarship: A Literature Review*, 51 VAND. L. REV. 1499, 1506-19 (1998) [hereinafter Langevoort, *Literature Review*] (reviewing past legal works utilizing research and theories from psychology and behavioral economics); Arlen et al., *supra* note 6, at 3-4 n.4 (citing articles bringing behavioral decision theory insights to corporate law issues).

10. I provide an overview of behavioral decision theory below. *See infra* notes 52-68 and accompanying text. The field of behavioral decision theory may be described briefly as the loose collection of psychologists, behavioral economists, experimental economists, political scientists, and others within the social sciences who concern themselves with developing a descriptive account of human judgment and decision making, particularly in relation to normative accounts of good judgment and decision making.

11. *See, e.g.*, Korobkin & Ulen, *supra* note 4, at 1055 ("There is simply too much credible experimental evidence that individuals frequently act in ways that are incompatible with the assumptions of rational choice theory.") (footnote omitted); Cass R. Sunstein, *Behavioral Analysis of Law*, 64 U. CHI. L. REV. 1175, 1194 (1997) ("Economic analysis of law has

In the view of these "legal decision theorists,"¹² behavioral decision theory establishes conclusively that humans systematically violate principles of rationality when making judgments and decisions about all sorts of matters in all sorts of domains.¹³ Our

proceeded on the basis of inaccurate understandings of decision and choice.").

12. I apply the label "legal decision theorists" to scholars working in this area because the topic of interest is legal decision making in its various forms and because this label evokes the notion of behavioral decision theory, the body of research on which so many of these scholars rely. Most of these legal scholars simply apply findings from behavioral decision research to legal issues and do not conduct their own behavioral studies (with Professors Chris Guthrie, Russell Korobkin, and Jeffrey Rachlinski being notable exceptions). Thus, the work tends to be more theoretical and speculative than empirical in nature. I also use "legal decision theory" generically to refer to the increasing application of behavioral decision theory research to legal topics. Of course, not all legal decision theorists will subscribe to every claim made by every other legal decision theorist, and no single legal decision theory yet exists. To limit confusion and increase the fairness of my remarks, I attempt to provide as much detail as possible on who has made what particular claims, including direct quotation of the statements at issue.

Some minor controversy surrounds the two leading names for this new field, behavioralism and behavioral law and economics. See *supra* note 4. Professors Henderson and Rachlinski assert that "[t]he term, 'behavioralism,' [as used by Jon Hanson and Douglas Kysar] is inappropriate in this context. The psychological research that Professors Hanson and Kysar rely on arises from cognitive, not behavioral, psychology. Behavioral psychology can be closely tied to microeconomic theory, which cognitive psychology can be used to critique." James A. Henderson, Jr. & Jeffrey J. Rachlinski, *Product-Related Risk and Cognitive Biases: The Shortcomings of Enterprise Liability*, 6 ROGER WILLIAMS U. L. REV. 213, 218 n.22 (2000) (citation omitted). That is, Henderson and Rachlinski assert that "behavioralism" may be confused with "behaviorism," which is most often associated with the works of John Watson and B. F. Skinner. But see Hanson & Kysar, *Behavioralism III*, *supra* note 2, at 264 n.9. Responding to Henderson and Rachlinski, Hanson and Kysar state that,

although the use of "behavioralism" is arguably inappropriate in light of the fact that it is cognitive, not behaviorist, psychology that is being integrated into the legal economic methodology, we take some comfort in the fact that a distinction is sometimes made between "behaviorism" (without the 'al'), which is the school of psychology that excludes cognitions from its analysis, and "behavioral science," which is not necessarily so constrained. In any event, the terminology at this point appears to be firmly entrenched in the legal literature.

Id. (citations omitted).

Professor Rostain points out that retention of the term "economics" in the appellation "behavioral law and economics" is problematic, for without a rational behavior assumption "the project's positive merit comes from the validity of the research about the specifics of human decision making and behavior, not from an abstract model of strategic reasoning based on the assumption that people are self-interested maximizers of their preferences." Tanina Rostain, *Educating Homo Economicus: Cautionary Notes on the New Behavioral Law and Economics Movement*, 34 LAW & SOC'Y REV. 973, 982 (2000). In other words, behavioral law and economics omits the quintessential element of microeconomics.

13. See, e.g., Hanson & Kysar, *Behavioralism II*, *supra* note 2, at 1425-26 ("One significant and surprising feature of such departures from rationality is that they are consistent and predictable—that is, they are 'neither rational, nor capricious.'") (footnote omitted); Korobkin

preferences change with normatively irrelevant changes in the decision setting (e.g., support for a public policy may depend on whether it is framed in terms of lives to be lost versus lives to be saved if the policy is implemented, even though the net result remains constant); we are too sensitive to anecdotal but unrepresentative data in our probability and causation judgments and not sensitive enough to diagnostic, base rate information or sample size; we are incapable of ignoring information about the occurrence and severity of events when judging after the fact whether a defendant acted negligently (i.e., we fall prey to a certainty-of-hindsight bias); we are overly optimistic about the likelihood of negative events and overly confident in the correctness of our factual beliefs—to name only a few of the cognitive errors and biases to which we are supposedly susceptible.¹⁴ Suffice it to say that legal

& Ulen, *supra* note 4, at 1085 (“Research in the behavioral sciences has demonstrated that individuals are systematically biased in their predictions of the probable results of various events.”); Cass R. Sunstein, *How Law Constructs Preferences*, 86 GEO. L.J. 2637, 2637 (1998) [hereinafter Sunstein, *Preferences*] (“Decisions are affected by cognitive limitations and motivational distortions that press choices in unanticipated directions. . . . [T]he departures from rational choice can be described, used, and sometimes even modeled, at least in broad outline.”); Sunstein, *supra* note 4, at 1 (“People are not always ‘rational’ in the sense that economists suppose. But it does not follow that people’s behavior is unpredictable, systematically irrational, random, rule-free, or elusive to social scientists. On the contrary, the qualifications can be described, used, and sometimes even modeled.”); see also Jennifer Arlen, *Comment: The Future of Behavioral Economic Analysis of Law*, 51 VAND. L. REV. 1765, 1766 (1998) (“Behavioral economic analysis of law scholars argue that people do not behave consistently with rational choice theory, and, moreover, that the deviations from rational behavior are systematic, not random. Most people are likely to exhibit certain biases, they assert, and thus these deviations from rational choice do not cancel each other out.”) (footnotes omitted).

The definition of the modifier “systematic” is implied rather than explicitly stated by the legal decision theorists. Within behavioral decision theory, the term is used to indicate that an error is the result of some purposeful, repetitive process rather than random error or measurement error. Robyn M. Dawes, *Behavioral Decision Making and Judgment*, in 1 THE HANDBOOK OF SOCIAL PSYCHOLOGY, *supra* note 9, at 497 (distinguishing systematic errors from random or unreliable errors). The term “bias” refers to a “source of error which is systematic rather than random.” Jonathan St. B.T. Evans, *Heuristic and Analytic Processes in Reasoning*, 75 BRIT. J. PSYCHOL. 451, 462 (1984). Hence, the phrase “systematic bias” is redundant when used in the judgment and decision-making context, but this phrase may be employed here at times to ensure the intended meaning.

14. “The sheer number of experiments reporting biases is so great that a sizable number of books and long survey papers have been written just to review the evidence.” John Conlisk, *Why Bounded Rationality?*, 34 J. ECON. LIT. 669, 670 (1996). I discuss specific cognitive anomalies as appropriate in the remaining sections of the Article. For summaries of the cognitive biases identified by behavioral decision researchers see SCOTT PLOUS, *THE*

decision theorists portray legally relevant judgment and decision making as seriously flawed.¹⁵ From this empirical assumption legal

PSYCHOLOGY OF JUDGMENT AND DECISION MAKING 109-88 (Philip G. Zimbardo ed., 1993); Dawes, *supra* note 13, at 499-521, 534-41; Hanson & Kysar, *Behavioralism I*, *supra* note 1, at 643-87; Jolls et al., *supra* note 9, 1476-1541; Korobkin & Ulen, *supra* note 4, at 1084-1126; Langevoort, *Literature Review*, *supra* note 9, at 1503-06; Hazel Markus & R.B. Zajonc, *The Cognitive Perspective in Social Psychology*, in 1 THE HANDBOOK OF SOCIAL PSYCHOLOGY 137, 176-97 (Gardner Lindzey & Eliot Aronson eds., 3d ed. 1985). One useful organizing scheme is provided by the quantitative psychologist Hal Arkes, who divides the anomalies into three categories based on their costs and benefits:

(1) "Strategy-based errors occur when subjects use a suboptimal strategy; the extra effort required to use a more sophisticated strategy is a cost that often outweighs the potential benefit of enhanced accuracy." Hal R. Arkes, *Costs and Benefits of Judgment Errors: Implications for Debiasing*, 110 PSYCHOL. BULL. 486, 486-87 (1991).

(2) "Association-based errors are costs of [an] otherwise highly adaptive system of associations within semantic memory. The automaticity of such associations, generally of enormous benefit, becomes a cost when judgmentally irrelevant or counterproductive semantic associations are brought to bear on the decision or judgment." *Id.* In this second category Arkes places such anomalies as the hindsight bias, the confirmation bias, and the overconfidence effect. *Id.*

(3) "Psychophysically based errors result from the nonlinear mapping of physical stimuli onto psychological responses. Such errors represent costs incurred in less frequent stimulus ranges where very high and very low stimulus magnitudes are located. These costs are more than offset by sensitivity gains in the more frequent stimulus ranges located in the central portion of the stimulus spectrum." *Id.* at 487. In this category Arkes places such anomalies as the sunk cost effect, the reflection effect, and the anchoring bias. *Id.*

15. Legal decision theorists address judgment and decision making in a wide variety of legally relevant settings. See, e.g., Chris Guthrie, *Better Settle Than Sorry: The Regret Aversion Theory of Litigation Behavior*, 1999 U. ILL. L. REV. 43, 61-62 (applying psychology to litigant judgment and decision making); Hanson & Kysar, *Behavioralism II*, *supra* note 2 (regarding consumer judgment and decision making); Jolls et al., *supra* note 9, at 1538-41 (regarding criminal judgment and decision making); *id.* at 1541-44 (discussing limits of consumer and bureaucratic decision making and possible implications of such limits for the legal system); Korobkin, *supra* note 9 (applying psychological theory to the question of how default terms affect contracting parties); Donald C. Langevoort, *Taking Myths Seriously: An Essay for Lawyers*, 74 CHI.-KENT L. REV. 1569-97 (2000) (applying psychology to attorney judgment and decision making); Prentice, *supra* note 2 (regarding auditor judgment and decision making); Rachlinski, *Hindsight*, *supra* note 9 (regarding judgments of judges and juries made in hindsight); see also Korobkin & Ulen, *supra* note 4, at 1058 ("By borrowing from psychological and sociocultural theories in addition to economics, the law-and-behavioral-science approach consciously chooses to emphasize its external usefulness in analyzing legal problems rather than either its internal elegance or universal applicability. Its ultimate goal, quite simply, is to understand the incentive effects of law better than modern law and economics is able to do by enlisting more sophisticated understandings of both the ends of those governed by law and the means by which they attempt to achieve their ends."); Sunstein, *supra* note 4, at 2 ("Analysis of law should be linked with what we have been learning about human behavior and choice. After all, the legal system is pervasively in the business of constructing procedures, descriptions, and contexts for choice.").

decision theorists argue that the law should be fashioned to take account of the systematic irrationality of legal actors.¹⁶ Conversely, they argue that proof of systematically irrational behavior makes it unwise to base laws on economic theories that necessarily assume rational behavior.¹⁷

The breadth of legal decision theory's assault on the rationality assumption cannot be overstated: legal decision theorists collectively contend that all judgments and decisions of legal importance—whether made by ordinary citizens or criminals, litigants or lawyers, judges or jurors—involve imperfect psychological processes that consistently cause irrational judgments and choices to be made.¹⁸ The statements of Hanson and Kysar

16. See Rostain, *supra* note 12, at 980 ("Much of this [behavioral law and economics] scholarship is prescriptive."). Professors Korobkin and Ulen offer a series of legal reforms that they claim flow from the premise that humans in general and legal actors in particular fall prey systematically to decision-making heuristics and biases, see Korobkin & Ulen, *supra* note 4, at 1085-1102, 1104-26, as do Professors Jolls, Sunstein, and Thaler, see Jolls et al., *supra* note 9, at 1522-41; Sunstein, *supra* note 4, at 2. Professors Hanson and Kysar likewise state that, "[u]ltimately, any legal concept that relies in some sense on a notion of reasonableness or that is premised on the existence of a reasonable or rational decisionmaker will need to be reassessed in light of the mounting evidence that a human is 'a reasoning rather than a reasonable animal.'" Hanson & Kysar, *Behavioralism I*, *supra* note 1, at 634-35 (footnote omitted).

17. Jolls et al., *supra* note 9, at 1545 ("Traditional law and economics is largely based on the standard assumptions of neoclassical economics. These assumptions are sometimes useful but often false."); Korobkin & Ulen, *supra* note 4, at 1055-56 ("It follows that the analysis of the incentive effects of legal rules based on such implausible behavioral assumptions cannot possibly result in efficacious legal policy, at least not in all circumstances."); Rachlinski, *New Law & Psychology*, *supra* note 9, at 740-41 ("This 'new' law and psychology promises a more accurate description of human choice than the law otherwise has available, which in turn should improve both positive and normative legal analysis."). Furthermore, behavioral decision theory is offered to justify laws that may appear unjustified or inexplicable under a more traditional economic analysis. Jolls et al., *supra* note 9, at 1546 ("Behavioral economics offers other sources of potential explanation—most prominently, perceptions of fairness. We have tried to show that many laws which are seemingly inefficient and do not benefit powerful interest groups may be explained on grounds of judgments about right and wrong."); Edward J. McCaffrey, *Cognitive Theory and Tax*, in *BEHAVIORAL LAW AND ECONOMICS*, *supra* note 4, at 398, 399 ("My central theme is that cognitive biases can help explain major structural features of our existing tax system that are otherwise difficult to understand, and that they must be taken into account in developing any general normative theory of tax.").

18. See Gregory S. Alexander, *A Cognitive Theory of Fiduciary Relationships*, 85 CORNELL L. REV. 767, 773 (2000) ("Legal analysis is particularly vulnerable to schemas and other cognitive biases and heuristics."); Arlen, *supra* note 13, at 1766 ("Behavioral economic analysis of law scholars argue that people do not behave consistently with rational choice theory, and, moreover, that the deviations from rational behavior are systematic, not random. Most people are likely to exhibit certain biases, they assert, and thus these deviations from

rational choice do not cancel each other out.”) (footnotes omitted); Richard Birke, *Reconciling Loss Aversion and Guilty Pleas*, 1999 UTAH L. REV. 205, 209 n.9 (citing personal communication with the late psychologist Amos Tversky for the proposition that all decision makers are loss averse); *id.* at 246 n.132 (citing personal communication with legal decision theorist Jeffrey Rachlinski for the proposition that criminal defendants are subject to the same cognitive biases as everyone else); Melvin Aron Eisenberg, *The Limits of Cognition and the Limits of Contract*, 47 STAN. L. REV. 211, 218 (1995) (noting that “within the last thirty years cognitive psychology has established that real people use certain decisionmaking rules (heuristics) that yield systematic errors, and that other aspects of actors’ cognitive capabilities are also systematically defective”); Larry T. Garvin, *Adequate Assurance of Performance: Of Risk, Duress, and Cognition*, 69 U. COLO. L. REV. 71, 161 (1998) [hereinafter Garvin, *Of Risk, Duress, and Cognition*] (“Cognitive psychology, dispelling the overly optimistic assumptions of expected utility theory, thus shows that the promisee is unlikely to evaluate risk accurately at the time of contracting, and will usually err on the side of underestimating infrequent risk.”); Larry T. Garvin, *Disproportionality and the Law of Consequential Damages: Default Theory and Cognitive Reality*, 59 OHIO ST. L.J. 339, 343 (1998) [hereinafter Garvin, *Cognitive Reality*] (“[W]e carry around a good many departures from rational choice theory. For example, most people tend to undervalue remote risks; overvalue vivid data and undervalue drab data; take risks to avoid loss, but avoid them to protect gains; and many others.”); *id.* at 395-96 (asserting that cognitive biases and errors are present throughout the population and occur systematically); Chris Guthrie et al., *Inside the Judicial Mind*, 86 CORNELL L. REV. 777, 778 (2001) (“Judges, it seems, are human. Like the rest of us, their judgment is affected by cognitive illusions that can produce systematic errors in judgment.”); Stephen D. Hurd, *Introduction*, 51 VAND. L. REV. 1497, 1497 (1998) (“Systematic behavioral deviations from the predictions of standard economic theory are now well-established in the psychological literature, making behavioral research an indispensable analytical tool for legal scholars.”); Jolls et al., *supra* note 9, at 1477 (“Even when the use of mental shortcuts is rational, it can produce predictable mistakes. . . . Actual judgments show systematic departures from models of unbiased forecasts, and actual decisions often violate the axioms of expected utility theory.”); Korobkin & Ulen, *supra* note 4, at 1057 n.19 (“As we show below, the experimental evidence establishes that the deviations are, indeed, systematic and not randomly distributed around a (rational actor) mean.”); *id.* at 1059 n.24 (“We posit in this Article that these anomalies in other social science fields would have been found in empirical legal studies if such studies were to occur.”); *id.* at 1085 (“Research in the behavioral sciences has demonstrated that individuals are systematically biased in their predictions of the probable results of various events.”); Langevoort, *Ego & Law*, *supra* note 9, at 884 (“Manifestations of ego—sometimes amusing, other times frustrating, often intriguing—operate pervasively in settings affected by law. . . . It is very likely . . . that researchers will observe some systematic, nontrivial bias in any settings that are subject to incomplete organizational or market-based checks.”); Rachlinski, *New Law & Psychology*, *supra* note 9, at 744 (“[Behavioral decision theory] certainly suggests that all social institutions, including courts, legislatures, and administrative agencies, will be subject to cognitive biases.”); Cass R. Sunstein, *Economics & Real People*, 3 GREEN BAG 2D 397, 398 (2000) (“It is now well-established, mostly through the pioneering work of psychologists Daniel Kahneman and Amos Tversky, that people make decisions on the basis of heuristic devices, or rules of thumb, that may work well in many cases but that also lead to systematic errors. It is also well established that people suffer from various biases and aversions that can lead to inaccurate perceptions.”); Sunstein, *supra* note 9, at 206 (“Contrary to the conventional wisdom: Workers are especially averse to losses, and not as much concerned with obtaining gains; . . . [and] Workers may well suffer from excessive

illustrate the general view expressed by legal decision theorists regarding the pervasiveness of "cognitive illusions": "These cognitive illusions—sometimes referred to as biases—are not limited to the uneducated or unintelligent, and they are not readily capable of being unlearned. Instead, they affect us all with uncanny consistency and unflappable persistence."¹⁹

A reader of legal decision theory scholarship is hard-pressed to find any clear, much less any theoretically driven, specification of boundary conditions on the contention that legal judgment and decision making are subject to numerous systematic defects.²⁰

optimism . . ."); see also *supra* note 15.

19. Hanson & Kysar, *Behavioralism I*, *supra* note 1, at 633 (footnotes omitted). The article by Hanson and Kysar from which this quote is taken is replete with very pessimistic accounts of human judgment and decision making of which the following comments are indicative:

[T]hese biases [in human thinking] reveal a deep misunderstanding of the nature of scientific judgment and a series of mental crutches awkwardly constructed to take up the slack. Their cumulative effect is to impair our appreciation of risk and uncertainty so severely that one commentator has termed the condition *probability blindness*. Whether blind or just impaired, people's sense of scientific and probabilistic judgment is inarguably of doubtful acuity.

Id. at 645-46 (footnote omitted); see also *id.* at 662 ("Special attention must be given to the subject of probabilistic judgment, for this is an area where researchers have uncovered a veritable fool's gold mine of nonrational cognitive anomalies.") (footnote omitted); *id.* at 671 (noting that "individuals consistently fail to assess costs and benefits separately, as an analytic approach to decisionmaking would require"); *id.* at 678 ("Both elastic justification and motivated reasoning capture ways in which individuals purport to 'justify' or 'reason through' their beliefs and decisions, when in actuality their initial impulses have biased the process all along.").

Later in this article, Hanson and Kysar contend that, even if cognitive illusions turn out to be rare or unimportant on average, the possibility of firms manipulating consumer biases makes the cognitive illusions important. *Id.* at 692-93 ("We will argue that even if cognitive biases are comparatively rare and insignificant, they assume a special importance in the consumer product context because consumers are rendered susceptible to bias manipulation by marketers and manufacturers, however rare or isolated the consumer bias may be.") (footnote omitted); *id.* at 747 ("Because individuals are subject to a host of nonrational yet systematic cognitive phenomena, any party who has control over a decisionmaking context can influence the perceptions of the decisionmaker."). Hanson and Kysar do not withdraw their broad assumption of uniformity, pervasiveness, and systematicness in cognitive biases that supposedly enables prediction by marketing manipulators (for instance, Hanson and Kysar make no distinctions between types of consumers); they simply argue that even if the frequency or severity of the biases is less than they contend, the biases still present considerable opportunity for exploitation by marketers. See *id.* at 747-48 (offering examples of marketers' exploitation of consumers).

20. Indeed, legal decision theorists rarely even define the central objects of their scholarship, the concepts of judgment and decision making, except by example. See, e.g.,

Hanson & Kysar, *Behavioralism I*, *supra* note 1, at 640-42 (discussing the concept of decision making by way of example without explicit definition); *id.* at 645-46 (discussing "scientific and probabilistic judgments" without an explicit definition); Rachlinski, *Hindsight*, *supra* note 9, at 571 (providing examples of judgments made in hindsight but no formal definition).

For present purposes, "judgment" may be defined as the process of perceiving and cognitively integrating stimuli to form a global evaluation. See Robert P. Abelson & Ariel Levi, *Decision Making and Decision Theory*, in 1 THE HANDBOOK OF SOCIAL PSYCHOLOGY, *supra* note 14, at 231, 235 ("Individuals are constantly faced with the need to make judgments—to combine information from a variety of sources into a global evaluation."); William M. Goldstein & Robin M. Hogarth, *Judgment and Decision Research: Some Historical Context*, in RESEARCH ON JUDGMENT AND DECISION MAKING: CURRENTS, CONNECTIONS, AND CONTROVERSIES 3, 4 (William M. Goldstein & Robin M. Hogarth eds., 1997) (noting that one central question in the study of judgment is "[h]ow do people integrate multiple, probabilistic, potentially conflicting cues to arrive at an *understanding* of the situation, a *judgment*?"); *id.* at 8-9 (describing how research on human judgment "examines people's responses to complex and/or uncertain sets of information" and "emphasizes the psychological appraisal of information").

"Decision" may be defined as the expression of a preferential choice. See Abelson & Levi, *supra*, at 231 ("[W]e confine our definition of *decision* to cases in which individuals act in the service of preferences"); Goldstein & Hogarth, *supra*, at 4-6 (describing decision research as primarily the study of preferential choice or the study of the choice of preferred courses of action). Within behavioral decision theory, "decision" and "choice" are typically treated as synonymous terms: "A *decision* is a choice of action—of what to do or not do." JONATHAN BARON, THINKING AND DECIDING 6 (3d ed. 2000).

While some see the distinction between judgment and decision as artificial, with both representing "a process of making a choice among several alternatives," Abelson & Levi, *supra*, at 235 (citation omitted), others find the distinction useful because it focuses attention on different aspects of the decision-making process: "[R]esearch on judgment tends to focus on what a person understands, believes, or feels, as opposed to what course of action a person will choose." Goldstein & Hogarth, *supra*, at 9. Professors Einhorn and Hogarth note:

[F]rom a psychological viewpoint, it may be more accurate to say that while judgment is generally an aid to choice, it is neither necessary nor sufficient for choice. That is, judgments serve to reduce the uncertainty and conflict in choice by processes of deliberative reasoning and evaluation of evidence. Moreover, taking action engenders its own sources of conflict . . . so that judgment may only take one so far; indeed, at the choice point, judgment can be ignored.

Hillel J. Einhorn & Robin M. Hogarth, *Behavioral Decision Theory: Processes of Judgment and Choice*, 32 ANN. REV. PSYCHOL. 53, 73 (1981).

I see the distinction as helpful because judgments and decisions may be complementary and consistent, but not necessarily so. Thus, a juror forms a judgment about a criminal defendant's guilt and then decides whether to vote for conviction or acquittal in light of this judgment and other influences within the jury room that may overwhelm the juror's personal judgment about guilt. Furthermore, the distinction is important because different biases and errors have been found to be associated with judgments versus decisions.

The process of judgment and decision making involves the perceptual, inferential, computational, and emotional processes that lead to a judgment about some aspect of the world and the expression of a decision in the form of an outward statement or action or internal commitment to a chosen course. See J.St.B.T. Evans et al., *Reasoning, decision making and rationality*, 49 COGNITION 165, 165-66 (1993) (noting distinction drawn by some

Even when legal decision theorists concede the point that surely not all decision making is irrational, they counter with the assertion that specific instances of legal decision making, however, will be particularly prone to irrational tendencies—without clearly explaining why this should be the case.²¹

This claim that cognitive biases and errors operate uniformly and pervasively in the population is crucial to the legal decision theory research program. By arguing pervasiveness, the legal decision theorists assert the general import of their work and maximize their comparative advantage relative to standard rational choice theorists

between reasoning processes and actual choices, but also noting that "[i]n real world situations the distinction between reasoning and decisionmaking is blurred"; see also Philip N. Johnson-Laird & Eldar Shafir, *The Interaction Between Reasoning and Decision Making: An Introduction*, 49 COGNITION 1, 1-2 (1993) (describing "folk psychology" distinctions between reasoning and decision making and noting the "interdependence between reasoning and decision making" and the "mutually recursive" relationship of the two); Arthur Lupia et al., *Beyond Rationality: Reason and the Study of Politics*, in ELEMENTS OF REASON: COGNITION, CHOICE, AND THE BOUNDS OF RATIONALITY 1, 1 (Arthur Lupia et al. eds., 2000) ("[C]hoice is the product of reason, where reason is the human process of seeking, processing, and drawing inferences from information."); Lance J. Rips, *Reasoning*, 41 ANN. REV. PSYCHOL. 321, 321 (1990) ("Problem-solving, decision-making, learning, and social understanding all obviously include reasoning. In fact, when we conceive of reasoning in this very general way, it seems to encompass almost any process of forming or adjusting beliefs and is nearly synonymous with cognition itself."); Karl Halvor Teigen, *Commentary: Decision-Making in Two Worlds*, 65 ORGANIZATIONAL BEHAV. & HUM. DECISION PROCESSES 249, 250 (1996) ("The functions of a decision are to finish the period of vacillation and hesitation and to initiate a course of action.").

An error may occur at some point in the underlying reasoning process or at the point at which a judgment or decision is expressed. That is, an apparently erroneous choice may be the result of faulty means-end reasoning or a mistake in choosing between options when trying to express a preference or belief. Cf. Robert T. Clemen, *Does Decision Analysis Work?* 1 (Aug. 1999) (unpublished manuscript, on file with author) ("Decision analysts are quick to point out the distinction between decision process and decision outcome, and that even the best decision process can be derailed by an unlucky outcome.") (citation omitted).

21. See Korbokin & Ulen, *supra* note 4, at 1085 ("But whether or not the well-documented collection of heuristics and biases are rational adaptations in a global sense, they have the consequence of causing actors to make decisions that violate the predictions of rational choice theory in individual circumstances."). Professors Korobkin and Ulen also state:

Our contention that traditional conceptions of rational choice theory are flawed in important ways does not suggest that we believe people are "irrational." Most of the observed deviations from behavior predicted by rational choice theory are quite sensible and understandable, and many seem quite rational in a "global" sense, although they result in behavior that violates the predictions of rational choice theory on the more "local" level on which legal scholarship generally operates.

Id. at 1144.

who do not assume pervasive information-processing and decisional errors. By arguing that errors are systematic and uniform rather than randomly distributed performance errors—and thus, by implication, that there is predictability in the flawed ways of the legal decision maker—legal decision theorists assert that this erroneous behavior can be modeled and taken into account in the formation and reformation of the law.²²

The claim of uniform and widespread irrational behaviors flows from the legal decision theorists' reading of the psychological literature to say that all judgment and decision making involves a set of basic but flawed psychological processes.²³ Bias in

22. *E.g.*, Hanson & Kysar, *Behavioralism I*, *supra* note 1, at 637 (noting that "the predictability of biases makes them easy to assimilate into economic models"); Jolls et al., *supra* note 9, at 1545 ("Most of these bounds [on rationality] can be and have been made part of formal models."). By arguing pervasiveness and systematicness in cognitive biases, and hence, in judgment and decision errors, the legal decision theorists also attempt to avoid or rebut the claim of economists that cognitive errors will be distributed randomly throughout the population and thus will, on average, be corrected by the market. *See, e.g.*, Garvin, *Cognitive Reality*, *supra* note 18, at 395-96 (arguing that such a claim by economists lacks substance because cognitive errors are systematic and pervasive and thus lead to shifts in the mean result within distributions); *see also* Stephen M. Bainbridge, *Mandatory Disclosure: A Behavioral Analysis*, 68 U. CIN. L. REV. 1023, 1035 n.54 (2000) ("It is the systematic nature of these biases that is critical. Standard economic analysis recognizes that individual decision makers may depart from rationality, but assumes that such departures come out in the wash—they cancel each other out so that the average or equilibrium behavior of large groups will be consistent with rational choice. By asserting that decision makers exhibit systematic biases, behavioral economics denies that claim.") (citation omitted); Garvin, *Of Risk, Duress, and Cognition*, *supra* note 18, at 163 (arguing that, "[w]hereas random errors may wash out [in the market], systematic errors probably will not; accordingly, one can model collective behavior better by using cognitive analysis").

23. *See, e.g.*, Korobkin, *supra* note 6, at 44 ("A variety of heuristics and biases cause individuals to take actions that do not necessarily maximize their expected utility."); Donald C. Langevoort, *Organized Illusions: A Behavioral Theory of Why Corporations Mislead Stock Market Investors (and Cause Other Social Harms)*, in *BEHAVIORAL LAW AND ECONOMICS*, *supra* note 4, at 147 ("A well-documented tendency of people who must operate in noisy informational environments is to adopt heuristic forms of thought. Busy executives process extraordinarily large amounts of information in both making decisions and deciding what matters deserve further time and attention. Such processing must necessarily be simplified, sometimes oversimplified, to make the information manageable, lest the executive be overwhelmed by data and paralyzed by ambiguity."); Jeffrey J. Rachlinski, *Heuristics and Biases in the Courts: Ignorance or Adaptation?*, 79 OR. L. REV. 61, 61 (2000) ("The human brain is extremely efficient, but it is not a computer. The brain has a limited ability to process information but must manage a complex array of stimuli. In response to its natural constraints the brain uses shortcuts that allow it to perform well under most circumstances. Reliance on these shortcuts, however, leaves people susceptible to all manner of illusions: visual, mnemonic, and judgmental.") (footnotes omitted); Sunstein, *supra* note 4, at 3 ("It is

now well established that people make decisions on the basis of heuristic devices, or rules of thumb, that may work well in many cases but that also lead to systematic errors.”).

An alternative explanation for deviations from perfect rationality is that humans are capable only of “bounded rationality” due to limited cognitive capacities in the face of numerous environmental demands. See, e.g., Herbert A. Simon, *Alternative Visions of Rationality*, in *RATIONALITY IN ACTION: CONTEMPORARY APPROACHES* 198 (Paul K. Moser ed., 1990). The cognitive psychologist Steven Sloman compares the bounded rationality view to that of the imperfect-psychological-process, or “natural assessment methods,” view:

Two complementary views of the reason we find systematic error can be distinguished. The older view, made popular by the seminal work of Herbert Simon, is commonly referred to as *bounded rationality*. The idea is that people make errors because they operate with limited cognitive resources. Our short-term memories have limited capacity; we can perform only a limited number of operations at any time; we have limited energy; indeed, we are limited in every way. However, many of the problems that confront us are enormously difficult computationally and sometimes impossible.

....
... In the early 1970s, the work of Amos Tversky and Daniel Kahneman made popular a second view of the determinants of systematic error. They argued that people make errors because they make judgments and decisions using heuristics (rules of thumb) that are quick and easy for people and that usually provide reasonable and adequate answers but fail under particular conditions. The heuristics they posit draw on the strengths of the human cognitive machinery, and therefore they refer to these heuristics as natural assessment methods.

....
The main difference between the two views of error has to do with the rationality of the process people are understood to be using when making judgments. Bounded rationality assumes that people are using a rational inference procedure; they are just limited in their ability to fully exploit it. The natural assessment approach assumes that people are using an arational procedure that approximates rational inference. ... [E]vidence exists suggesting that people think in different ways, using both rational and heuristic procedures. Given this multiplicity of thought, the two views of error are complementary. Bounded rationality explains error in situations in which people are using a rational procedure, and the natural assessment approach explains why people also use procedures other than rational ones.

Steven A. Sloman, *Rational versus Arational Models of Thought*, in *THE NATURE OF COGNITION* 573, 575-76 (Robert J. Sternberg ed., 1999).

Some legal decision theorists, and some psychologists, at times conflate these two accounts of error. Compare Korobkin & Ulen, *supra* note 4, at 1075-76 (characterizing bounded rationality as consisting of intentional trade-offs between costs and benefits of full information-processing using normative decision rules), with *id.* at 1069, 1076 (equating bounded rationality with the unintentional use of cognitive heuristics). See also Korobkin, *supra* note 6, at 44 (equating the heuristic and bias view of error with bounded rationality); Douglas L. Medin & Max H. Bazerman, *Broadening Behavioral Decision Research: Multiple Levels of Cognitive Processing*, 6 *PSYCHONOMIC BULL. & REV.* 533, 534 (1999) (equating Simon's bounded rationality view with the cognitive heuristics view). The two perspectives may, however, lead to different predictions. See Gerd Gigerenzer & Daniel G. Goldstein,

judgment, particularly judgments made under conditions of uncertainty, supposedly arises from over-reliance on "a limited number of heuristic principles which reduce the complex tasks of assessing probabilities and predicting values to simpler judgmental operations."²⁴

The heuristics most often cited in this respect are: *representativeness*, whereby "the likelihood of an event is evaluated by the degree to which it is representative of the major characteristics of the process or population from which it originated"; *availability* which refers to the tendency of subjects to "assess the frequency of a class or the probability of an event by the ease with which instances or occurrences can be brought to mind"; and *anchoring and adjustment*, which refers to the use of an initial value (operating as an "anchor") of subjects' estimations which is adjusted to yield the final answers; the effect of anchoring is proposed to create insufficient adjustment.²⁵

Thus, rather than mentally applying rules of probability theory, causal analysis, or logic to resolve the uncertainties necessary to form a judgment, people are said to rely on these simpler mental operations, or cognitive heuristics, to reach their conclusions. While use of these shortcuts can yield acceptable results,²⁶ over-

Reasoning the Fast and Frugal Way: Models of Bounded Rationality, 103 PSYCHOL. REV. 650, 651 (1996) (distinguishing bounded rationality from the heuristics and biases approach). For instance, a proponent of the bounded rationality view might predict that persons with greater computational capacity (for example, greater working memory capacity) would exhibit greater rationality, whereas a proponent of the heuristic view would likely predict more prevalent operation of automatic heuristics leading to biases across persons and situations.

24. Amos Tversky & Daniel Kahneman, *Judgment Under Uncertainty: Heuristics and Biases*, in JUDGMENT UNDER UNCERTAINTY: HEURISTICS AND BIASES 3 (Daniel Kahneman et al. eds., 1982).

25. Dina Berkeley & Patrick Humphreys, *Structuring Decision Problems and the "Bias Heuristic"*, 50 ACTA PSYCHOLOGICA 201, 235-36 (1982) (citations omitted). The psychologist Max Bazerman classifies eleven of the most common judgmental biases identified by behavioral decision researchers as arising from the operation of the representativeness, availability, and anchoring/adjustment heuristics. See MAX BAZERMAN, JUDGMENT IN MANAGERIAL DECISION MAKING 39-40 tbl. 2.2 (4th ed. 1998).

26. See, e.g., PLOUS, *supra* note 14, at 109 ("Normally, heuristics yield fairly good estimates. The disadvantage of using heuristics, however, is that there are certain instances in which they lead to systematic biases (i.e., deviations from normatively derived answers)."); Arkes, *supra* note 14, at 486 ("My goal is to show that the costs of otherwise beneficial

reliance on or unthinking use of heuristics biases judgment and leads to error, because the heuristic processing device fails to take account of some important consideration in the stimulus environment, overemphasizes some stimuli, attends to irrelevant stimuli, or skips an important computational step necessary to perform in accordance with normative principles.²⁷

Decisional errors, particularly the expression of seemingly inconsistent preferences under conditions of risk or uncertainty where the decision maker's subjective utility or subjective probability estimates are involved, are seen as arising from the "constructive" nature of preference and choice: "[P]references for and beliefs about objects or events of any complexity are often constructed—not merely revealed—in the generation of a response

cognitive adaptations are the consequence of appropriate responses to environmental demands.").

27. Distinctions are sometimes drawn between biases arising from cognitive versus motivational sources. See, e.g., Lee Ross & Craig A. Anderson, *Shortcomings in the Attribution Process: On the Origins and Maintenance of Erroneous Social Assessments*, in JUDGMENT UNDER UNCERTAINTY: HEURISTICS AND BIASES, *supra* note 24, at 129, 134-35. Motivated biases involve an effort to protect one's self-image or existing beliefs, theories, or goals, whereas cognitive biases involve the flawed operation of cognitive processes without any presumed directional or self-serving goal. Although cognitive heuristics giving rise to cognitive biases have received a great deal of attention within behavioral and legal decision theory, biases arising from motivational sources represent an important origin of judgmental error as well. See, e.g., Ziva Kunda, *The Case for Motivated Reasoning*, 108 PSYCHOL. BULL. 480, 493 (1990) ("The position that all self-serving biases are due to purely cognitive processes is . . . no longer tenable. . . . [U]nder the current state of knowledge, the motivational account appears to be more parsimonious and coherent than the purely cognitive one.") (citation omitted); *id.* at 495 ("Although the mechanisms underlying motivated reasoning are not yet fully understood, it is now clear that directional goals do affect reasoning."); Medin & Bazerman, *supra* note 23, at 535 ("In its early years, [behavioral decision research] focused almost exclusively on cognitive errors, or errors that have their roots in how we process information. . . . Starting in the late 1980s, research emerged that suggested errors that we make as a result of motivational biases."); see also Guthrie, *supra* note 15, at 61-62 ("When confronted with difficult litigation decisions, like whether to settle a case or go forward with trial, litigants will likely number-crunch, calculate, and value-maximize . . . but they will also feel a range of actual and prospective emotion that they will incorporate into their decision making. Legal scholars seeking to understand, describe, and perhaps even modify litigation behavior need not abandon their elegant, 'calculating theories,' but they need to couple them with richer theories that take the reality—not just the rationality—of human beings into account."); Langevoort, *Ego & Law*, *supra* note 9, at 856 ("[M]otivated biases in cognitive processing allow individuals unconsciously to distort reality about themselves, substituting a set of typically healthy illusions—a form of self-deception—to protect the ego from external threats that might otherwise diminish the functional sense of self-efficacy.") (footnote omitted).

to a judgment or choice task.”²⁸ Because people do not have a tidy rank ordering of stable preferences in their heads ready for access in any given situation but instead often must form a preference given the information at their disposal in the choice situation, preference formation is susceptible to transitory influences that cause choices to appear contradictory across time and situations. Thus, elements of the decision situation that are normatively irrelevant under rational choice theory (for example, whether a choice is phrased in terms of monetary loss or gain should be irrelevant because decision makers should evaluate choices in terms of overall wealth) may be psychologically significant and lead to differences in choices (for example, one of the most prominent findings in behavioral decision research is that the framing of a choice as a loss or a gain can affect choices).²⁹ In short, “[p]references appear to be remarkably labile, sensitive to the way a choice problem is described or ‘framed’ and to the mode of response used to express the preference.”³⁰ This “notion of constructive preferences means as well that preferences are not necessarily generated by some consistent and invariant algorithm such as expected value calculation. It appears that decision makers have a repertoire of methods for identifying their preferences and developing their beliefs.”³¹

28. John W. Payne et al., *Behavioral Decision Research: A Constructive Processing Perspective*, 43 ANN. REV. PSYCHOL. 87, 89 (1992) (citation omitted); see also Rachlinski, *New Law & Psychology*, *supra* note 9, at 752 (noting that “people do not possess a fixed set of preferences that they seek to satisfy with their choices, as economic models usually assume”). This “constructed preferences” view differs sharply from the standard economic view of preferences, which holds that “preferences are predetermined in any choice situation[] and do not depend on what alternatives are available.” Daniel McFadden, *Economic Choices*, 91 AM. ECON. REV. 351, 356 (2001).

29. See generally Daniel Kahneman & Amos Tversky, *Choices, Values, and Frames*, in CHOICES, VALUES, AND FRAMES 1, 1-16 (Daniel Kahneman & Amos Tversky eds., 2000) (providing an overview of their seminal research into the effects of decision frames on choices); Daniel Kahneman & Amos Tversky, *Prospect Theory: An Analysis of Decision under Risk*, in CHOICES, VALUES, AND FRAMES, *supra*, at 17-43 (describing prospect theory, a descriptive theory of choice under risk where the number of possible outcomes is small); Amos Tversky & Daniel Kahneman, *Advances in Prospect Theory: Cumulative Representation of Uncertainty*, in CHOICES, VALUES, AND FRAMES, *supra*, at 44-65 (describing an extension of prospect theory to risky choices with any number of outcomes and situations involving uncertainty).

30. Paul Slovic, *The Construction of Preference*, 50 AMER. PSYCHOLOGIST 364, 365 (1995) (citations omitted).

31. Payne et al., *supra* note 28, at 89 (citation omitted).

In sum, legal decision theorists share the common goal of greatly qualifying, if not tearing down, the rational behavior assumption that undergirds the economic analysis of law and replacing it with what they contend is a more realistic or more descriptively accurate account of judgment and decision making.³² This "new scholarship in law based on behavioral science"³³ essentially would replace the "economic miser" metaphor of human behavior with a "cognitive miser" metaphor.³⁴ Under this latter view, the economic miser is

32. Professor Jolls notes, for instance, that "[b]ehavioral law and economics is not a critique of law and economics. . . . Its goal is to offer better predictions and prescriptions about law based on improved accounts of how people actually behave." Jolls, *supra* note 4, at 1654; see also Korobkin & Ulen, *supra* note 4, at 1144 ("[W]e do not argue that the edifice of rational choice theory, which underlies so much of legal scholarship, be ripped down. Rather, we suggest that it be revised, paying heed to important flaws in its structure that unduly and unnecessarily limit the development of a more nuanced understanding of how law affects society."); Thomas S. Ulen, *The Growing Pains of Behavioral Law and Economics*, 51 VAND. L. REV. 1747, 1748 (1998) ("Behavioral law and economics does not attempt to undo any of the remarkable accomplishments of law and economics. Rather, it is an attempt to refine.").

33. Korobkin & Ulen, *supra* note 4, at 1058 n.24. As indicated above in footnote 9, this "new" scholarship is not particularly new in actuality, but the level of attention presently being paid to the scholarship is new. As suggested by Professor Langevoort, legal decision theory's qualifications to the economic analysis of law are a natural result of behavioral decision theory's qualifications to the economic model of rationality, because legal decision theory is an outgrowth of behavioral decision theory. See Langevoort, *Literature Review*, *supra* note 9, at 1500-02 (discussing the diffusion of behavioral decision theory into legal scholarship); cf. Amartya K. Sen, *Rational Fools: A Critique of the Behavioral Foundations of Economic Theory*, in BEYOND SELF-INTEREST 25-43 (Jane J. Mansbridge ed., 1990) (surveying economists' own long-standing unease with and concern about the foundations of the economic theory of rationality).

34. The social psychologists Shelley Taylor and Susan Fiske coined the "cognitive miser" term to portray human decision making as a compromise between accuracy and efficiency demands necessitated by the considerable time constraints and stimulus complexity confronting decision makers. See, e.g., SUSAN T. FISKE & SHELLEY E. TAYLOR, *SOCIAL COGNITION* 11-12 (1984) ("One name for this is the *cognitive miser* model. The idea is that people are limited in their capacity to process information, so they take shortcuts whenever they can."). The cognitive miser metaphor

is an extension of approaches in cognitive psychology and decision making that regard human information-processing as generally adaptive but nevertheless prone to systematic error. According to this view, social judgments and decisions are guided by a range of heuristics and biases that serve to render the judgment situation manageable but which also introduce error unintentionally.

Russell Spears & S. Alexander Haslam, *Stereotyping and the Burden of Cognitive Load*, in *THE SOCIAL PSYCHOLOGY OF STEREOTYPING AND GROUP LIFE* 172 (Russell Spears et al. eds., 1997) (citations omitted). The cognitive miser metaphor is not meant to be entirely pejorative, because use of heuristic thought devices is often considered adaptive, but one cannot help but be left with something of a negative impression on reading the phrase given the negative connotations associated with the term "miser." As the psychologist Peter Suedfeld quips,

often too myopic, biased, and error-prone to achieve his or her own selfish goals. Reconceiving people as cognitive misers rather than economic misers leads to a very different conception of how people should be treated under the law, with cognitive misers generally due less responsibility and autonomy and in need of greater oversight, control, and assistance.³⁵

B. The Need for Greater Scrutiny of Legal Decision Theory

The conception of the legal actor as predictably irrational urged by legal decision theorists deserves strict scrutiny not simply as a matter of knowledge growth and intellectual exchange but also because the controls on the use of legal decision theory scholarship as persuasive authority are weak (particularly when the work is published in non-peer-reviewed journals), whereas the stakes associated with the use of this scholarship as the basis for judicial, legislative, and administrative decisions may be very high.³⁶ Unlike other fields where formal peer review serves a gate-keeping function,³⁷ there are relatively few peer-reviewed legal

within much of psychology "the rational actor has been demoted to a cognitive klutz." Peter Suedfeld, *Cognitive Managers and Their Critics*, 13 POL. PSYCHOL. 435, 436 (1992).

35. See Jolls et al., *supra* note 9, at 1541 ("In its normative orientation, conventional law and economics is often strongly antipaternalistic. . . . [B]ounded rationality pushes toward a sort of anti-antipaternalism—a skepticism about antipaternalism, but not an affirmative defense of paternalism. We also note . . . that while bounded rationality may increase the need for law (if government's failings are less serious than citizens'), bounded self-interest may reduce it, by creating norms that solve collective action problems even without government intervention.") (emphasis added) (footnotes omitted); Sunstein, *supra* note 9, at 271 ("Contrary to the reflexive antipaternalism of many lawyers and economists, workers are likely to have difficulty in deciding whether to waive their rights, and hence the case for nonwaivable terms, defended in behavioral terms, is far more plausible than it appears when defended in terms of simple redistribution.").

36. Cf. Vladimir J. Konečni & Ebbe B. Ebbesen, *External Validity of Research in Legal Psychology*, 3 LAW & HUM. BEHAV. 39, 40 (1979) ("Research in applied disciplines must be concerned with issues of external validity and generalizability to an unusually high degree. The criteria for what is a good experiment, when a certain methodology appears sound, and which results are to be trusted, must necessarily be different and more stringent when sweeping, costly, and far-reaching changes in public policy and—quite literally in the legal domain—people's futures and lives, depend on inferences from research results.").

37. See Lars Noah, *Sanctifying Scientific Peer Review: Publication as a Proxy for Regulatory Decisionmaking*, 59 U. PITT. L. REV. 677, 693 (1998) ("Peer review, in its broadest sense, represents the scientific community's effort to police itself and to assure a certain minimum level of quality so that scientists and others can rely on the results of reported scientific research. Any claim that would significantly add to or change the body of scientific

journals,³⁸ and the persons accepting and editing articles that present empirical data or that make claims about empirical research often have little expertise in, or specialized knowledge about, the area of inquiry.³⁹ Other than the informal peer review that may occur before submission of an article and may be

knowledge is regarded skeptically until it has been subjected to some form of peer scrutiny.") (footnote omitted); *id.* at 693-709 (discussing the benefits and shortcomings of the peer review process); *id.* at 709 (concluding that "[n]otwithstanding its various shortcomings, editorial peer review serves a valuable purpose in the publication of scientific research"). But see Robert J. MacCoun, *Biases in the Interpretation and Use of Research Results*, 49 ANN. REV. PSYCHOL. 259, 277-78 (1998) (discussing how peer review may be unreliable and may fail to serve as an effective screening device and thus should be used in conjunction with other review and screening devices, such as replication and meta-analysis); Peter J. Spiro, *Globablization, International Law, and the Academy*, 32 N.Y.U. J. INT'L L. & POL. 567, 589 (2000) ("Publication decisions at the law journals are made by student editors; for all the shortcomings of this system (it is ridiculed in other disciplines) it does eliminate the censoring possibilities of the peer-review standard found elsewhere.").

38. See Edward J. Conry & Caryn L. Beck-Dudley, *Meta-Jurisprudence: A Paradigm for Legal Studies*, 33 AM. BUS. L.J. 691, 723 n.169 (1996) ("The *American Business Law Journal* is one of the handful of law reviews that employs peer review, blind refereeing, and double-blind refereeing."); Robert J. Spitzer, *Lost and Found: Researching the Second Amendment*, 76 CHI.-KENT L. REV. 349, 376 (2000) ("The discipline of law is unique among academic disciplines in that its professional journals are governed mostly by student-run law review boards, and with a few exceptions, submissions are not subject to the process of peer review, or even faculty oversight.") (footnotes omitted).

Although a few student-edited legal journals do regularly employ peer review (e.g., *Yale Journal of Health Policy, Law and Ethics*) and some will occasionally seek peer review, it should not be assumed that all faculty-edited legal journals uniformly employ formal peer review, blind or otherwise. For example, Professor Epstein has stated that during his tenure as editor of the *Journal of Legal Studies*, he did not always seek a referee's opinion on submissions. Richard A. Epstein, *Faculty-Edited Law Journals*, 70 CHI.-KENT L. REV. 87, 90 (1994) ("In general therefore, my practice was to take the papers I liked, reject the ones I did not, and to ask for referees' opinions on those on which I could not quite make up my mind.").

39. See Bernard J. Hibbitts, *Last Writes? Reassessing the Law Review in the Age of Cyberspace*, 71 N.Y.U. L. REV. 615, 672 (1996) ("We have already seen that the current law review system operates with minimal quality control in the generally accepted ('peer review') sense of that term: there are still very few faculty-edited law journals, and it is at least questionable whether the second- and third-year editors of the student-run reviews can consistently make accurate qualitative (as opposed to institutionally, reputationally, or stylistically related) judgments about submissions other than those dealing with the most familiar subjects or offering the most familiar (e.g. doctrinal) brands of legal analysis."); James Lindgren, *An Author's Manifesto*, 61 U. CHI. L. REV. 527, 527 (1994) ("[Law review editors] often select articles without knowing the subject, without knowing the scholarly literature, without understanding what the manuscript says, without consulting expert referees, and without doing blind reads."). This concern applies equally to faculty- and student-edited journals that do not seek review of submissions from referees with expertise in psychology or behavioral economics.

inadequate for a variety of reasons,⁴⁰ an adversarial model of competing scholarship akin to the "truth-finding" model applied to American trials is perhaps the best external check on validity currently available.⁴¹

Furthermore, once an article enters the marketplace of legal ideas, it may be used by persons motivated to extend a scholarly thesis beyond its intended scope or consumed by persons who possess little or none of the information or knowledge necessary to

40. Professor Austin rejects informal review of articles by colleagues before submission to law reviews (or "vetting") as an adequate substitute for formal peer review and notes that public vetting may in fact be detrimental:

The commentary exchanged in the standard peer review process is an influential voice in the publication decision and thus is detailed, critical, and seen by both writer and editor. Whether law review vetting is probing, irrelevant garbage, or a perfunctory pat on the head by an old friend, makes no difference since the editor never sees anything—except a list of the names of alleged vetters. Editors can be misled into assuming that public vetting is traditional peer review and consequently the piece has been given careful scrutiny. Perhaps editors should adopt the policy of having authors furnish their vetters' commentary. A long list of vetters—often as many as twenty or thirty—implicitly gives the impression of approval, thus putting more pressure on the review to accept. Vetting attracts more attention than footnote density. Moreover, if a well known scholar is listed as a vetter, what editor wants to reject the article and take a chance on antagonizing someone into not submitting an article in the future?

Arthur D. Austin, *The "Custom of Vetting" as a Substitute for Peer Review*, 32 ARIZ. L. REV. 1, 6-7 (1990) (footnotes omitted).

Nor can we count on internal checks on quality control and researcher bias to ensure sound, fair-minded scholarship. See, e.g., Philip E. Tetlock & Gregory Mitchell, *Liberal and Conservative Approaches to Justice: Conflicting Psychological Portraits*, in PSYCHOLOGICAL PERSPECTIVES ON JUSTICE 234, 234-55 (Barbara A. Mellers & Jonathon Baron eds., 1993) (discussing how ideological tunnel vision and initial assumptions about human behavior may have biased the social-psychological study of lay conceptions of justice). For a discussion of researcher biases that may make self-scrutiny an inadequate control in the scientific process see MacCoun, *supra* note 37, at 267-75.

41. Professor Richard Lempert's close examination of an empirical psychological study published in the *Arizona Law Review* illustrates both the dangers of the publication of empirical research in law reviews and the way in which the present system of legal scholarship should work in the absence of formal peer review. See Richard Lempert, *Juries, Hindsight, and Punitive Damage Awards: Failures of a Social Science Case for Change*, 48 DEPAUL L. REV. 867 (1999) (critiquing Reid Hastie & W. Kip Viscusi, *What Juries Can't Do Well: The Jury's Performance as a Risk Manager*, 40 ARIZ. L. REV. 901 (1998)). Professor Lempert's hope was that his commentary would "substitute in some degree for the peer review that a study like this would ordinarily receive." *Id.* at 869 (footnote omitted). For another example of peer review occurring after publication see Robert J. MacCoun, *The Costs and Benefits of Letting Juries Punish Corporations: Comment on Viscusi*, 52 STAN. L. REV. 1821 (2000) (commenting on W. Kip Viscusi, *Corporate Risk Analysis: A Reckless Act?*, 52 STAN. L. REV. 547 (2000)).

examine the article critically.⁴² Therefore, for the marketplace of ideas to function efficiently, consumers need to be given choices among competing and complementary viewpoints to combat informational asymmetries and motivated misuses of ideas and research.

It is not simply a matter of speculation that legal decision theorists' arguments about the lack of rationality in legal decisions may be used in support of efforts to influence verdicts, restrict the role of juries, alter burdens of proof and persuasion, change substantive legal rights, and set caps on damages.⁴³ Indeed, whereas

42. See David L. Suggs, *The Use of Psychological Research by the Judiciary: Do the Courts Adequately Assess the Validity of the Research?*, 3 LAW & HUM. BEHAV. 135 (1979) (discussing the lack of systematic mechanisms in place to test the internal and external validity of research used by judiciaries and legislatures); see also Robert Rosenthal & Peter David Blanck, *Science and Ethics in Conducting, Analyzing, and Reporting Social Science Research: Implications for Social Scientists, Judges, and Lawyers*, 68 IND. L.J. 1209, 1211 (1993) ("In the context of an adversarial process, the problem is that the poorly designed study is likely to lead to unwarranted and inaccurate conclusions by a fact finder.").

Professor Lempert notes, for example, that Exxon Corporation, which has sponsored experimental research that supposedly reveals biases in civil juries when setting punitive damages and which, of course, was subject to a very large punitive damages verdict as a result of its Alaskan oil spill, may reasonably be expected to "seek to disseminate the authors' apparent findings and their conclusions in judicial and political arenas where they will help to make the case against jury awards of punitive damages." Lempert, *supra* note 41, at 869; see also *id.* at 871 n.16 ("[I]t appears that Exxon is making a concerted effort to build a social science case for reducing or taking away the jury's discretion in awarding punitive damages and that the Hastie and Viscusi study is a part of this effort. . . . Indeed, Exxon has recently cited the above research in its appeal of the \$5.3 billion Exxon Valdez award.") (citation omitted). For further discussion of Exxon's funding and use of punitive damages research see Neil Vidmar, *Juries Don't Make Legal Decisions! And Other Problems: A Critique of Hastie et al. on Punitive Damages*, 23 LAW & HUM. BEHAV. 705, 713 (1999); see also Elizabeth Amon, *Exxon bankrolls critics of punitives*, NAT'L L.J., May 17, 1999, at A1, A6. For a reply to the just-referenced Vidmar article, see Reid Hastie et al., *Reply to Vidmar*, 23 LAW & HUM. BEHAV. 715 (1999).

43. See, e.g., Victor Gold, *Covert Advocacy: Reflections on the Use of Psychological Persuasion Techniques in the Courtroom*, 65 N.C. L. REV. 481, 481 (1987) ("For a price, professional psychologists are available to advise lawyers on all aspects of trial advocacy including what to say, where to stand, how to select a jury, when to object, and what clothes to wear.") (footnote omitted); J. Alexander Tanford & Sarah Tanford, *Better Trials Through Science: A Defense of Psychologist-Lawyer Collaboration*, 66 N.C. L. REV. 741, 742 (1988) (responding to Professor Gold's article and seeking to justify the use of psychological research on alleged juror bias by claiming that "[t]his collaborative process of identifying existing barriers to rational decision making and devising strategies to reduce their impact has improved the chances that juries will understand and consider each litigant's case without bias"); see also, e.g., Richard Lempert, *Why Do Juries Get a Bum Rap? Reflections on the Work of Valerie Hans*, 48 DEPAUL L. REV. 453, 458-59 (1998) ("[T]he Exxon Corporation, following

legal decision theorists generally express caution about their endeavor and note the preliminary nature of much of their work,⁴⁴ in some areas they seem already to have persuaded themselves of the correctness of their views and thus have endorsed specific legal reforms based on their conclusions.⁴⁵ Some courts have already

the \$5 billion punitive damage verdict against it in the Exxon Valdez oil spill case invested substantial sums in a series of studies designed to show that juries cannot competently and fairly set punitive damages. One would not expect them to be less shy about investing in judicial elections if the demise of the jury system meant more was at stake.") (footnote omitted). See generally *PSYCHOLOGY AND SOCIAL ADVOCACY* (Peter Suedfeld & Philip E. Tetlock eds., 1991) (collecting essays debating the relevance and propriety of the use of psychological research in various public policy arenas).

Although L. J. Cohen's statement that "[t]he epistemological legitimacy of our liberal-democratic institutions may . . . be seen to stand or fall with belief in the inherent rationality of ordinary people," may sound alarmist to some, his broader observation that "[q]uite a lot is at stake here" at least seems true, particularly when the rationality debate enters the legal arena as it does in legal decision theory. L. Jonathan Cohen, *Can Irrationality Be Discussed Accurately?*, 7 *BEHAV. & BRAIN SCI.* 736, 737 (1984); see also Sally Lloyd-Bostock, *The Benefits of Legal Psychology: Possibilities, Practice and Dilemmas*, 79 *BRIT. J. PSYCHOL.* 417, 436 (1988) ("Like it or not, and however they characterize their work, psychologists who work on legal topics enter a sphere of societal and personal values. As psychologists increase their influence on law, legal decisions and legal policy, ethical questions loom larger. Research and practice in legal psychology frequently raises questions of legal ethics, as well as ethical dilemmas within psychology.").

44. Legal decision theorists generally express caution in the introductions and conclusions of their articles, often with respect to policy arguments, but seem to abandon this caution at other points in the same articles and particularly with respect to empirical claims, making at times very broad, unqualified statements about the scope of research findings. Compare Korobkin & Ulen, *supra* note 4, at 1144 ("We present none of these policy implications as definitive solutions to legal problems. Rather, they are offered as stakes in the ground that, we hope will serve as starting points for a new generation of legal scholarship . . ."), with *id.* at 1057 n.19 ("As we show below, the experimental evidence establishes that the deviations are, indeed, systematic and not randomly distributed around a (rational actor) mean."), and *id.* at 1086 ("The pervasiveness of the representativeness heuristic can help justify a set of rules of evidence law frustrating to rational choice theory."). See Sunstein, *supra* note 9, at 206-07 ("[T]raditional understandings of employee behavior and employment law make many blunders, because they are based on an inadequate sense of workers' actual values and behavior. . . . In short, workers are like most people. They behave like homo sapiens, not like homo economicus.") (footnote omitted); *id.* at 273 ("Much of my argument here has been for a degree of agnosticism with respect to positive issues (what people do) and normative issues (what the law should do). Much more work is necessary on both sorts of issues."). Perhaps, as the psychologist Phoebe Ellsworth suggests, such exaggeration or exuberance is par for the course in empirical research generally; see Phoebe C. Ellsworth, *Sticks and Stones*, 23 *LAW & HUM. BEHAV.* 719, 720 (1999) ("Every empirical study has flaws. Almost every author exaggerates the significance of the problem studied, overstates the implications of the results, or both."). But of course, compliance with a bad professional norm is not a good justification for the practice.

45. Perhaps most common have been arguments for legal reforms to address the hindsight

bias to which jurors are assumed to fall prey when assessing tort liability. See Philip G. Peters, Jr., *Hindsight Bias and Tort Liability: Avoiding Premature Conclusions*, 31 ARIZ. ST. L.J. 1277, 1289-92 (1999) (surveying various reform proposals advanced by legal decision theorists and others to counteract an alleged hindsight bias in legal judgment); see also Arlen, *supra* note 13, at 1768 n.11 (citing "[e]fforts to use behavioral analysis as the basis of normative policy conclusions"). The hindsight bias "refers to the tendency to overestimate the degree to which one would have been able to predict the outcome of an event or the answer to a question after one has received feedback about one's prediction or after the event has occurred." Elizabeth Creyer & William T. Ross, Jr., *Hindsight Bias and Inferences in Choice: The Mediating Effect of Cognitive Effort*, 55 ORGANIZATIONAL BEHAV. & HUM. DECISION PROCESSES 61, 61 (1993) (citation omitted); see Ralph Hertwig et al., *The Reiteration Effect in Hindsight Bias*, 104 PSYCHOL. REV. 194, 195 (1997) (describing "hindsight bias" as occurring with respect to memory judgments and the related "knew-it-all-along effect" as occurring with respect to hypothetical judgments of pre-feedback confidence); *id.* at 201 (distinguishing further the hindsight bias from the knew-it-all-along effect).

While a few studies report hindsight bias in mock jurors who express individual judgments, e.g., Kim A. Kamin & Jeffrey J. Rachlinski, *Ex Post ≠ Ex Ante: Determining Liability in Hindsight*, 19 LAW & HUM. BEHAV. 89, 99-101 (1995) (noting that though the effect may have been exacerbated by the manner in which the experiment was performed, researchers did find evidence suggesting a hindsight bias in some of mock jurors' judgments), there has been little or no demonstration that mock juries (i.e., jurors making group decisions) evince the bias. *Id.* at 100 ("[A]t present the effects of group deliberation on the hindsight bias are unknown"). There are a number of reasons to believe that individual juror judgments may not be identical to jury judgments. See, e.g., Norbert L. Kerr et al., *Bias in Judgment: Comparing Individuals and Groups*, 103 PSYCHOL. REV. 687, 690 (1996) ("It is not the case . . . that groups tend generally to be more biased than individuals for one type of bias, but less biased for another."); *id.* at 713 ("The central question of this paper has been, 'Which is more likely to make a biased judgment, individuals or groups?' Our overview of the relatively small and diverse empirical literature suggested that there was no simple empirical answer to this question. Even when we restrict our attention to particular bias phenomena (e.g., framing effects, preference reversals), there was frequently little consistency in the direction (i.e., sign) and magnitude of observed relative bias, *RB*."); John M. Levine et al., *Social Foundations of Cognition*, 44 ANN. REV. PSYCHOL. 585, 600-03 (1993) (discussing how group processes may cause "joint cognitions" to differ from individual cognitions); see also John C. Brigham & Adina W. Wasserman, *The Impact of Race, Racial Attitudes, and Gender on Reactions to the Criminal Trial of O. J. Simpson*, 29 J. APPLIED SOC. PSYCHOL. 1333 (1999) (illustrating that the details of a specific case and the race of the observers may be very important to the operation of the hindsight bias). In fact, though the hindsight bias has rarely been examined in group settings, one reported study found some attenuation of the hindsight bias in group judgments as compared to individual judgments under certain circumstances, see Dagmar Stahlberg et al., *We Knew It All Along: Hindsight Bias in Groups*, 63 ORGANIZATIONAL BEHAV. & HUM. DECISION PROCESSES 46, 56 (1995) (reporting that groups were as susceptible to hindsight bias as individuals in some circumstances but that group deliberation attenuated the effect in other circumstances), but one could argue that a group judgment may show an even greater hindsight bias than individual judgment in some other circumstances, cf. David Schkade et al., *Deliberating About Dollars: The Severity Shift*, 100 COLUM. L. REV. 1139, 1139 (2000) (finding a "severity shift" in mock juries' punitive damage awards as compared to an averaging of individual jurors' awards). But see James H. Davis et al., *Effects of Group Size and Procedural Influence on Consensual Judgments of Quantity: The Example of Damage*

endorsed claims made by legal decision theorists.⁴⁶ Despite the serious implications of the legal decision theorists' contentions, their work has received little careful scrutiny and even less criticism.⁴⁷

Awards and Mock Civil Juries, 73 J. PERS. & SOC. PSYCHOL. 703, 714 (1997) (finding that deliberating "groups awarded less than preferred by parallel individuals working alone, although this difference was only marginally significant").

In the end, whether group decisionmaking will show greater or less sensitivity to a supposed bias or error found in individual jurors will depend on the facts of the case, the manner in which the case is presented, the specific composition of the jury, and a host of other possible situational influences. See Norbert L. Kerr et al., *Bias in Jurors vs Bias in Juries: New Evidence from the SDS Perspective*, 80 ORGANIZATIONAL BEHAV. & HUM. DECISION PROCESSES 70, 82 (1999) ("For the jury, our findings confirm what is evident in the empirical record—that there is probably no general answer to the question 'which is more biased, jurors or juries?' The answer to this question must be 'it depends.'" (citation omitted); Peters, *supra*, at 1303 ("[T]he jury's group deliberations have a debiasing potential that warrants serious study"). As of now, we just do not know whether the hindsight bias shown by some individual mock jurors translates consistently into a hindsight bias in the decisions of mock and/or real juries on ultimate questions of liability and damages.

46. See Henderson & Rachlinski, *supra* note 12, at 218 & n.21 (citing cases where courts seem to be relying on certain legal decision theory arguments in the products liability domain).

47. Only a few voices have been raised in an effort to encourage greater care and precision in this new psychological approach to the law. Professors Peters and Rostain discuss serious concerns about the new scholarship that are complementary to the concerns raised here. See Peters, *supra* note 45 (discussing reasons why it would be premature to institute legal reforms designed to redress the hindsight bias to which jurors allegedly fall prey); Rostain, *supra* note 12 (discussing the modest success with which social science can predict behavior and why this new behavioral movement is normatively incomplete). A few other critics raise important issues of concern regarding this new scholarship but in less detail. See Arlen, *supra* note 13, at 1768-69 (expressing concern about the predictive power of legal decision theory because it lacks a general model of behavior, about the translation of behavioral research to complex legal environments, and about the lack of clear policy implications to be drawn from legal decision theory because all legal actors are presumed to be subject to biases and errors); Daniel A. Farber, *Toward a New Legal Realism*, 68 U. CHI. L. REV. 279, 296-303 (discussing limits on behavioral law and economics); Robert A. Hillman, *The Limits of Behavioral Decision Theory in Legal Analysis: The Case of Liquidated Damages*, 85 CORNELL L. REV. 717, 729-37 (1998) (questioning the external validity of behavioral decision research, noting lack of adequate attention to possible context effects and individual differences, noting that the errors and biases described by legal decision theorists may lead to conflicting policy prescriptions or none at all because of the supposed ubiquity of errors and biases in all legal actors, and noting a lack of normative guidance from legal decision theory); Samuel Issacharoff, *Can There Be a Behavioral Law and Economics?*, 51 VAND. L. REV. 1729, 1734 (1998) (supporting behavioral research but questioning whether many of its findings are generalizable, of "sufficient magnitude as to systematically undermine predictions of behavior" from rational choice theory, and "capable of being operationalized to condition the behavior of all persons subject to specific legal regulation"); Mark Kelman, *Behavioral Economics as Part of a Rhetorical Duet: A Response to Jolls, Sunstein, and Thaler*, 50 STAN. L. REV. 1577, 1579 (1998) (characterizing behavioral law and economics as an interpretive

If the policy prescriptions drawn from legal decision theory are based on faulty assumptions, bad research, or incomplete understandings of behavior, then unintended results may ensue following implementation of the suggested reforms and the intellectual integrity of the field may suffer.⁴⁸ To avoid such outcomes, we should subject legal decision theory to much closer scrutiny, and the legal decision theorists should recognize the need for greater caution and precision in the drawing of descriptive and prescriptive conclusions from empirical research on judgment and decision making.

In the next part, I describe behavioral decision theory and limitations on its usefulness for the law. In so doing, I survey a great deal of psychological research that brings into question the claims of the legal decision theorists regarding the consistent fallibility of judgment and decision making in experimental settings and qualifies the generalizations that can be safely drawn from this research. Before proceeding, however, two points should be made.

First, my criticisms of legal decision theory should not be seen as an argument that human decision makers predominantly act rationally and only occasionally make computational errors. In fact, the legal decision theorists, in their rush to portray decision making as a uniform phenomenon, overlook some important personal and situational factors that can actually exacerbate the effects of certain cognitive biases.⁴⁹ Nor is this Article another attempt to rewrite rational choice theory to account for the findings of behavioral

trope rather than a falsifiable theory and arguing that behavioral law and economics offers an approach to data that is ambiguous in its normative implications); Robert E. Scott, *The Limits of Behavioral Theories of Law and Social Norms*, 86 VA. L. REV. 1603, 1639-46 (2000) (cautioning against the neglect of contextual determinants of behavior); Thomas S. Ulen, *The Growing Pains of Behavioral Law and Economics*, 51 VAND. L. REV. 1747, 1754, 1756 (1998) (supporting legal decision theory but suggesting that greater attention needs to be paid to distinguishing the effects of environmental and structural problems from cognitive limitations and the "ability of legislative action and ex ante administrative agency regulation to correct problems of cognitive limitations"). Judge Posner has provided the most explicit defense of law and economics in the face of the attack by behavioralism advocates, arguing that this new behavioral movement does not present a new or serious challenge to the economic analysis of legal behavior. See Richard A. Posner, *Rational Choice, Behavioral Economics, and the Law*, 50 STAN. L. REV. 1551 (1998).

48. See Rostain, *supra* note 12, at 1002 ("Proponents of behavioral law and economics ... risk overselling their approach. A proposal's failure to produce the anticipated results can result in the loss of credibility of law and social science-based approaches more generally.").

49. See Mitchell, *supra* note 8.

decision theory.⁵⁰ Rather, this Article simply provides reasons why legal decision theorists should refrain from broad statements about the manner in which all legal actors process information, make judgments, and reach decisions and why others should be skeptical of such broad claims.

Second, my criticisms also should not be understood as a rejection of the psychological analysis of law. This Article arises from a strong belief in the utility of psychological and other empirical research for legal analysis and from an equally strong belief that overreaching in the uses of empirical research serves to undermine the long-term credibility of the endeavor and may help to unsettle the legal status quo without adequate justification but with serious consequences for legal players.⁵¹ Just as the goal of the legal decision theorist is primarily to qualify rather than overthrow the rational choice model, much of what I say here is meant to qualify legal decision

50. Jolls, Sunstein, and Thaler characterize Judge Posner's reaction to the behavioral law and economics movement, or the legal decision theory movement as I call it, as such an attempt. See Christine Jolls et al., *Theories and Tropes: A Reply to Posner and Kelman*, 50 STAN. L. REV. 1593, 1595 (1998) (replying to Posner, *supra* note 47). Professors Jolls, Sunstein, and Thaler note:

Throughout Posner's commentary, he goes through the following ritual. He discusses one of the phenomena we identify as problematic for economic theory; he offers a modification or elaboration on the standard theory that could, in principle, be consistent with this phenomenon; and then he declares victory. Posner seems to think that the fact that it is possible to tell a rational choice story consistent with the data is sufficient to establish that this explanation is the correct one.

Id.; see also Colin F. Camerer, *Behavioral Economics and Nonrational Organizational Decision Making*, in DEBATING RATIONALITY: NONRATIONAL ASPECTS OF ORGANIZATIONAL DECISION MAKING 53, 55-56 (Jennifer J. Halpern & Robert N. Stern eds., 1998) (discussing further examples of efforts by positive economists to account for findings from behavioral economics).

51. I tend to share the cautious view of the psychologists Konečni and Ebbesen toward legal psychology:

A more cautious point of view, and one that we favor, is that erroneous information obtained by scientific methods (and therefore having an aura of truth) is more harmful than no information at all, especially when issues as sensitive as legal ones are being dealt with, and people's futures are quite literally at stake.

Konečni & Ebbesen, *supra* note 36, at 68. The counterpoint view is that existing laws based on faulty descriptions of behavior or behavioral assumptions should not be given presumptive validity. The unavoidable problem, of course, is deciding what descriptions and assumptions are faulty, what descriptions and assumptions should serve as replacements, what weight social scientific studies on the relevant issues should be given, and what standards of proof should be required to alter the status quo.

theory rather than reject it and is meant to point out areas in need of further investigation and consideration. The legal decision theorists' assumption of pervasive failures of rational thought is overbroad, just as the economic theorists' assumption of pervasive rationality is overbroad. The descriptive truth lies somewhere in between and is likely found in a closer examination of how features of the person and situation interact to cause different responses to different decision tasks.

II. BEHAVIORAL DECISION THEORY AND ITS INTERPRETIVE LIMITS

The empirical study of judgment and decision making is often referred to as "behavioral decision theory."⁵² One of the leading behavioral decision theorists, Robyn Dawes, describes the field in the following way:

Basically, behavioral decision making is the field that studies how people make decisions. Because all types of people are making all sorts of decisions all the time, the field is potentially very broad. What has characterized the field both historically and theoretically is the comparison of actual decision making with certain principles of rationality in decision making—for example, that increasing the number of options available to a decision maker should not increase the probability that a particular option from the more restricted set is chosen, or that the way in which identical choices are described ("framed") should not affect choice. When actual decisions violate such principles *systematically* (not just as a result of unreliability or "error"), this deviation is termed an *anomaly*—if the people who violate these principles simultaneously accept them as ones that they believe *should* govern their decision making.⁵³

52. For recent overviews of behavioral studies of judgment and decision making, see Dawes, *supra* note 13; Goldstein & Hogarth, *supra* note 20; R. Hastie, *Problems for Judgment and Decision Making*, 52 ANN. REV. PSYCHOL. 653 (2001); B. A. Mellers et al., *Judgment and Decision Making*, 49 ANN. REV. PSYCHOL. 447 (1998).

53. Dawes, *supra* note 13, at 497. For a more detailed classification of the various approaches to the study of judgment and decision making, see Abelson & Levi, *supra* note 20, at 236 (drawing a distinction between "behavioral decision theory" and "psychological decision theory," with both approaches examining departures from normative decision theory but with the latter providing a cognitive account of the departures). For another discussion of the meaning and origins of the phrase "behavioral decision theory," see Rachlinski, *New Law & Psychology*, *supra* note 9, at 740-41.

As this description indicates, the defining feature of behavioral decision theory is the comparison of actual judgment and decision-making behavior with normative models of rational behavior.

The norms that serve as the criteria for rational behavior in behavioral decision theory emphasize the internal coherence and logical consistency of decisions and judgments.⁵⁴ The norms used to evaluate preferential choice derive from game theory and models of expected utility (the primary source of "rational choice theory" within economics⁵⁵) and accordingly emphasize internal consistency

Some object to the label "behavioral decision theory" on the grounds that this body of research does not represent a unified theory of behavior but rather only a collection of results and mini-theories. See, e.g., Posner, *supra* note 47, at 1552. But see Jeanne L. Schroeder, *Rationality in Law and Economics Scholarship*, 79 OR. L. REV. 147, 156 (2000) (suggesting that Herbert Simon's bounded rationality theory presents a unified descriptive theory of economic behavior and decision making). Although this objection is valid to some extent (and is accepted by proponents of behavioral decision theory as a valid but not fatal criticism, see Korobkin & Ulen, *supra* note 4, at 1071-73), I consider the "theory" label acceptable for two reasons: (1) this body of research shares important assumptions about, and methods for the study of, judgment and decision making; and (2) the term "behavioral decision theory" has a long tradition within psychology that renders its meaning readily recognizable within the field. With respect to the latter point, consider that the *Annual Review of Psychology* has been publishing reviews of judgment and decision-making research and theory under the label of "behavioral decision theory" since 1961. See Gordon M. Becker & Charles G. McClintock, *Value: Behavioral Decision Theory*, 18 ANN. REV. PSYCHOL. 239 (1967); Ward Edwards, *Behavioral Decision Theory*, 12 ANN. REV. PSYCHOL. 473 (1961); Einhorn & Hogarth, *supra* note 20; Paul Slovic et al., *Behavioral Decision Theory*, 28 ANN. REV. PSYCHOL. 1 (1977). An alternative name sometimes used for the field, which omits the "theory" label, is simply behavioral decision making. See, e.g., Gideon Keren, *Perspectives of Behavioral Decision Making: Some Critical Notes*, 65 ORGANIZATIONAL BEHAV. & HUM. DECISION PROCESSES 169 (1996).

54. See Mellers et al., *supra* note 52, at 449 ("In the rational choice framework, rationality is expressed as internal coherence and logical consistency within a system of beliefs and preferences."). Numerous definitions of rational behavior exist within the social sciences. For a sampling of the diverse definitions see Lupia et al., *supra* note 20, at 3-8; see also Jack Feldman & Michael K. Lindell, *On Rationality*, in ORGANIZATION AND DECISION THEORY 83, 83-91 (Ira Horowitz ed., 1990) (discussing the treatment of rationality within judgment and decision making research); Korobkin & Ulen, *supra* note 4, at 1060-66 (discussing "thin" and "thick" conceptions of rationality that vary along a continuum of specificity in their predictions about human behavior); Mark C. Suchman, *On Beyond Interest: Rational, Normative and Cognitive Perspectives in the Social Scientific Study of Law*, 1997 WIS. L. REV. 475, 475-84 (contrasting instrumental, moral, and cognitive accounts of decision making). For the distinction between decisions and judgments, see sources cited *supra* note 20.

55. See Thomas S. Ulen, *Rational Choice Theory in Law and Economics*, in 1 ENCYCLOPEDIA OF LAW AND ECONOMICS 790, 791-93 (Boudewijn Bouckaert & Gerrit De Geest eds., 2000) (discussing formal and informal definitions of "rational choice theory," with the

of revealed preferences and maximization of utility; the norms used to evaluate judgments derive from systems of logic, probability theory and statistical theory, principles of causal analysis, and sometimes conventions such as legal rules, with the emphasis being on following prescribed rules in the consideration of evidence and the drawing of inferences or conclusions.⁵⁶ The success of a course of behavior in achieving any goal other than fidelity to these normative principles is typically irrelevant in behavioral decision research.⁵⁷

formal definition being drawn from expected utility theory axioms); see also Korobkin & Ulen, *supra* note 4, at 1062-64 (discussing the expected utility version of rational choice theory).

56. See, e.g., RICHARD NISBETT & LEE ROSS, *HUMAN INFERENCE: STRATEGIES AND SHORTCOMINGS OF SOCIAL JUDGMENT* 8 (1980) ("The formal inferential rule system followed by professional scientists is the standard against which the layperson is compared throughout this book."); Daniel Kahneman & Amos Tversky, *On the Study of Statistical Intuitions*, 11 *COGNITION* 123, 124 (1982) ("The presence of an error of judgment is demonstrated by comparing people's responses either to an established fact (e.g., that the two lines are equal in length) or to an accepted rule of arithmetic, logic or statistics."); Kerr et al., *supra* note 45, at 688 ("The concept of biased judgment assumes that one can specify a non-biased standard of judgment against which actual human judgments can be compared. The basis of that standard, the *normative model* of judgment, may be some formal logical system (e.g., syllogistic logic, probability theory, game theory, rational choice models). However, the normative model may also be based on convention. Good examples of the latter types of standards are the common law rules of evidence that proscribe jurors' use of certain information (e.g., a defendant's race or gender or physical appearance)."); Gordon F. Pitz & Natalie J. Sachs, *Judgment and Decision: Theory and Application*, 35 *ANN. REV. PSYCHOL.* 139, 140 (1984) ("[T]here may exist no criterion for determining whether a single choice or judgment is correct, since the response is based in part on personal opinions or preferences. It is possible, however, to impose a mathematical or logical structure on the task that defines the *consistency* of a set of responses. The prescriptions for consistent behavior are generally derived from formal probability theory and from Expected Utility (EU) theory, a prescriptive model of choice founded on axioms proposed by von Neumann & Morgenstern.") (citations and footnote omitted).

One may argue that treating even principles derived from statistics or logic as norms of rational behavior in fact reflects the convention or intuition that those principles provide the best way to solve problems given limited information or provide the best way to organize preferences, thus giving rise to debate about whose intuitions and conventions should control. For discussion of this issue, see the target article by L.J. Cohen, the open peer commentaries on Cohen's article, and Cohen's response to the commentaries in volume 4 of *Behavioral and Brain Sciences*. L. Jonathan Cohen, *Can Human Irrationality be Experimentally Demonstrated?*, 4 *BEHAV. & BRAIN SCI.* 317-31 (1981); *Open Peer Commentary*, *id.* at 331-59; *Author's Response*, *id.* at 359-70.

57. See, e.g., Philip E. Tetlock, *The Impact of Accountability on Judgment and Choice: Toward a Social Contingency Model*, 25 *ADVANCES IN EXPERIMENTAL SOC. PSYCHOL.* 331, 335 (1992). Professor Tetlock notes:

Subjects in laboratory studies of cognitive processes rarely feel accountable to others for the positions they take. They function in a social vacuum (or as close

Application of these normative principles provides a logical or coherent answer within the closed universe of the information available to the decision maker and given the expressed preferences of the decision maker, but application of these principles does not necessarily lead to an empirically accurate or socially adept answer.⁵⁸ For instance, one may reach a logically correct conclusion

an approximation to a social vacuum as can be achieved) in which they do not need to worry about the interpersonal consequences of their conduct ("How will others react if I do this? How effectively can I justify my views if challenged?"). Such issues are simply seen as irrelevant to the explanatory goals of the [behavioral decision theory] research program.

Id.

58. See Kenneth R. Hammond, *Coherence and Correspondence Theories in Judgment and Decision Making*, in *JUDGMENT AND DECISION MAKING: AN INTERDISCIPLINARY READER* 53, 53 (Terry Connolly et al. eds., 2d ed. 2000).

The goal of a correspondence metatheory is to describe and explain the process by which a person's judgments achieve *empirical accuracy*. The goal of a coherence metatheory of judgment, in contrast, is to describe and explain the process by which a person's judgments achieve logical, or mathematical, or statistical *rationality*. . . .

It may come as a surprise to the reader that rationality does not directly imply accuracy and vice versa, but brief reflection shows that this is the case. Rationality always operates in a closed system; given the premises, certain conclusions always follow if a rational reasoning process is followed. When the reasoning process satisfies a logical test, the system is termed coherent, and that is all it is and all it claims to be.

Id.; see also J.St. B.T. Evans, *Bias and Rationality*, in *RATIONALITY: PSYCHOLOGICAL AND PHILOSOPHICAL PERSPECTIVES* 7-15 (K.I. Manktelow & D.E. Over eds., 1993) (comparing "rationality of purpose" and "rationality of process" in cognitive research); Herbert A. Simon, *Rationality in Psychology and Economics*, in *RATIONAL CHOICE: THE CONTRAST BETWEEN ECONOMICS AND PSYCHOLOGY* 27 (Robin M. Hogarth & Melvin W. Reder eds., 1986) [hereinafter *RATIONAL CHOICE*] ("The rational person of cognitive psychology goes about making his or her decisions in a way that is procedurally reasonable in the light of the available knowledge and means of computation.").

A coherence account of rationality

does not specify any particular goal as humanly universal. . . . In the rational choice model, . . . there is no particular goal. Rather, there is a set of possible goals that must be ordered. . . . It is . . . important to point out that whether the goals in the set are good or evil or neither is irrelevant, as long as the participants order them.

William H. Riker, *The Political Psychology of Rational Choice Theory*, 16 *POL. PSYCHOL.* 23, 24-25 (1995). Many economic models of rationality assume that people are motivated by selfish interests (or by utility functions), and that people make choices and take actions designed to realize their selfish interests consistently and to the greatest extent possible, but the content of these selfish interests is generally unspecified. Arlen, *supra* note 13, at 1766 ("Conventional law and economics assumes that people exhibit rational choice: that people are self-interested utility maximizers with stable preferences and the capacity to optimally

based on reasoning from false premises, but this logical conclusion may be nonsensical or demonstrably false in relation to conditions in the real world.⁵⁹ Thus, in behavioral decision theory, the touchstone of rationality is either (1) consistency in the face of normatively irrelevant features of the decision situation or (2) appropriate change in judgments and decisions as normatively relevant features of the situation change.⁶⁰

accumulate and assess information.”) (footnote omitted); Jane J. Mansbridge, *The Rise and Fall of Self-Interest in the Explanation of Political Life*, in *BEYOND SELF-INTEREST*, *supra* note 33, at 12 (“Shortly after its inception the rational choice school evolved a set of practitioners who did not require the self-interest assumption but only the assumption that rational actors would act consistently and maximize whatever ends they preferred. Yet the version of rational choice that restricted its modeling to self-interest predominated at the start.”). Much debate about economic models of rationality concerns not whether people are good at maximizing their selfish interests but rather the proper meaning of utility or selfishness. *E.g.*, Linnda R. Caporael et al., *Selfishness examined: Cooperation in the absence of egoistic incentives*, 12 *BEHAV. & BRAIN SCI.* 683, 683 (1989) (examining the empirical basis for the assumption that “human nature is basically selfish and individualistic”).

Of greatest interest to the legal decision theorists has been the procedural aspect of rationality. See Lynn A. Stout, *Other-Regarding Preferences and Social Norms* 3 (Mar. 2001) (Georgetown University Law Center, Working Paper No. 265902) (“[C]ontemporary challenges to the rational selfishness model of human behavior tend to focus far more on the first adjective—the assumption of rationality—than on the second—the assumption of selfishness”), available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=265902. Some legal decision theorists, however, also consider the motivational component. For instance, Professors Jolls, Sunstein, and Thaler touch on the motivational aspect of rational choice models in their discussions of “bounded willpower” and “bounded self-interest.” See Jolls et al., *supra* note 9, at 1479; see also Korobkin & Ulen, *supra* note 4, at 1126-43 (considering “deviations from self-interest”).

59. See *BARON*, *supra* note 20, at 55 (“Rationality is . . . not the same as accuracy, and irrationality is not the same as error. We can use good methods and reach erroneous conclusions, or we can use poor methods and be lucky, getting a correct answer.”).

60. See *Evans*, *supra* note 58, at 16 (“Interest focuses on tasks which are normally sufficiently simple in structure that there is a tacit agreement that subjects ‘ought’ to be able to solve them. When subjects’ responses deviate from the answer which is prescribed by standard logic or by some other normative system . . . then subjects are typically reported as making errors. A ‘bias’ is usually defined as systematic attention to some logically irrelevant features of the task, or systematic neglect of a relevant feature.”); Daniel Kahneman, *Judgment and Decision Making: A Personal View*, 2 *PSYCHOL. SCI.* 142, 143-44 (1991) (“A salient characteristic of current research in judgment and decision making is its emphasis on anomalies that violate normative standards of belief or choice. A common strategy is to identify a factor that should *not* affect judgment or decision, then design an experiment in which it does. Thus, an uninformative message can produce anchoring; subjects exposed to different anchors will make different estimates of, say, the population of Turkey or the likely selling price of a house. Similarly, the framing of outcome statistics in terms of mortality or survival can affect the attractiveness of medical treatments: 10% mortality is more frightening than 90% survival. In these cases and in many others, subjects are influenced by

"[T]he existence of a single correct answer is a *necessary* assumption" in behavioral decision theory research,⁶¹ because "[e]rrors are always defined relative to a normative framework that makes assumptions about human goals."⁶² If the "rational" solution to a decision problem is indeterminate, then conclusions about the rationality of behavior cannot be uncontroversially drawn.

Using this internal coherence/logical consistency approach to rationality, behavioral decision researchers have identified a great number of apparent anomalies in judgment and choice, instances where observed conduct deviates from that predicted by one of the normative principles of rationality.⁶³ Evidence of these cognitive biases, or "cognitive quirks" as Richard Posner somewhat dismissively calls them,⁶⁴ brings into question the descriptive accuracy of a view of human behavior as perfectly rational.⁶⁵ Behavioral decision theory thus offers a descriptive account of judgment and decision making that stands in stark contrast to an ideal conception of judgment and decision making derived from norms of procedural rationality.

Behavioral decision theory's portrait of the fundamentally flawed decision maker is not sharply drawn, however, and presents a compelling picture of pervasive nonrational behavior only when viewed from a distance (or, more precisely, behavioral decision theory's claims are most compelling when stated only in the most general terms without appropriate qualifiers). When one examines more closely the details of the portrait painted by behavioral

a factor that they would wish to ignore, or in some way fail to meet a standard that, upon reflection, they would accept.").

61. GERD GIGERENZER ET AL., *THE EMPIRE OF CHANCE: HOW PROBABILITY CHANGED SCIENCE AND EVERYDAY LIFE* 227 (1989); see also *id.* ("A bias is defined as the *deviation* from a normatively correct answer, and this makes the assumption of a single correct answer essential.").

62. Mellers et al., *supra* note 52, at 449.

63. See *supra* note 14.

64. Posner, *supra* note 47, at 1553 ("The cognitive quirks that set bounds on rational maximizing include the availability heuristic, overoptimism, the sunk-cost fallacy, loss aversion, and framing effects . . .").

65. See, e.g., Amos Tversky & Daniel Kahneman, *Rational Choice and the Framing of Decisions*, in *RATIONAL CHOICE*, *supra* note 58, at 68 ("We argue that the deviations of actual behavior from the normative model are too widespread to be ignored, too systematic to be dismissed as random error, and too fundamental to be accommodated by relaxing the normative system. . . . We conclude from these findings that the normative and the descriptive analyses cannot be reconciled.").

decision theorists and the tools used to paint the portrait, one sees that the portrait is unfinished in many areas and that the tools used are blunt and at times inappropriate to the task.

Nevertheless, many legal decision theorists present behavioral decision theory as compelling grounds to adjudge all legal decision makers fundamentally flawed, for whom norms of rational behavior are inappropriate models of behavior. Consider, as an example of unwarranted pessimism drawn from selected behavioral decision studies, the claim made by some legal decision theorists that people consistently fail to take proper account of base rate information when judging probabilities.⁶⁶ As Jonathan Koehler's review of base-rate studies demonstrates, this pessimistic view of base-rate reasoning is unjustified and misleading: "Not only is there little evidence that base rates are routinely ignored, but a critical review of the recent literature shows that base rates usually influence judgments and often do so in reasonable ways."⁶⁷ Koehler does not deny that base-rate information is sometimes neglected, but, in a statement that applies with equal force to claims made by psychologists and legal scholars, asserts that "the ubiquitous summary statements of base rate neglect distort the empirical literature."⁶⁸

66. *E.g.*, Korobkin & Ulen, *supra* note 4, at 1085 ("To accurately predict the probability of future events, actors must consider the statistical probability that an event will occur and 'update' (adjust) this 'base rate' with any available particularized information about a specific situation. There is significant evidence, however, that actors systematically underuse base rates when making probability predictions, or ignore them altogether.") (footnote omitted); Cass R. Sunstein, *Which Risks First?*, 1997 U. CHI. LEGAL F. 101, 119 ("People tend to be insensitive to the sample size, to misunderstand the phenomenon of regression to the mean, to have excessive confidence in their own judgments, and to misunderstand the effect on probability of base-rate frequency.") (footnote omitted).

67. Jonathan J. Koehler, *The Base Rate Fallacy Reconsidered: Descriptive, Normative, and Methodological Challenges*, 19 BEHAV. & BRAIN SCI. 1, 2 (1996); *see id.* at 5 ("In sum, there is little evidence either from the lawyer-engineer problem or from the stereotype literature to support the strong claim that base rates are routinely ignored when individuating information is made available. Indeed, when care is taken to develop an appropriately individualized criterion, even weaker forms of the base rate fallacy do not receive clear and convincing support.").

68. *Id.*; *see also* Edwards & von Winterfeldt, *supra* note 9, at 233 ("While these studies generally supported the existence of a base rate fallacy, they also showed that base rates are sometimes taken into account: when the link between base rate and target event is causal, when base rates appear relevant, when the base rates related to individuating information, and when both diagnostic and base rate information are essentially statistical.").

As this first example suggests, and as we will see is the case in many other instances, legal decision theorists fail to appreciate or fail to report the qualified nature of many of behavioral decision theory's claims. In the following sections, I discuss three categories of limitations on behavioral decision theory that counsel against the rush to pronounce legal decision makers as predictably hobbled by pervasive and unavoidable cognitive biases and errors. When these limitations on behavioral decision theory are taken into account, a foggy but in many ways less dire picture of human rationality emerges.

A. Features of Behavioral Decision Theory That Mask Individual and Situational Differences in Rational Behavior and Distort Perceptions of the Prevalence of Irrational Behavior

Rarely, if ever, does behavioral decision research find uniform results across all experimental subjects.⁶⁹ In fact, when one examines the actual data gathered by decision researchers rather than just summary presentations of the data, one finds that at least a significant minority and often a significant majority of the subjects provided the "right," or rational, answer to the judgment or decision problem under consideration.⁷⁰ Yet because of the way that behavioral decision research is conducted and reported, this

69. It is standard usage nowadays in experimental psychology to refer to experimental "participants" rather than subjects. See AMERICAN PSYCHOLOGICAL ASS'N, PUBLICATION MANUAL OF THE AMERICAN PSYCHOLOGICAL ASSOCIATION 49 (4th ed. 1994) [hereinafter APA MANUAL] ("Replace the impersonal term *subjects* with a more descriptive term when possible—*participants, individuals, college students, children, or respondents*, for example (*subjects* and *sample* are appropriate when discussing statistics)."). Nevertheless, here I will usually refer to experimental subjects rather than participants because a legal audience is likely to better understand the meaning of the former.

70. See Riker, *supra* note 58, at 36 ("None of the experiments displaying inconsistencies in choice portray all subjects as inconsistent. For experimenters to recommend the abandonment of expected utility theory when the experiments themselves show that many people—often well over half, as in the preference reversal experiments—are indeed expected utility maximizers is to ignore the evidence that the experimenters have themselves created."); Richard F. West & Keith E. Stanovich, *The Domain Specificity and Generality of Overconfidence: Individual Differences in Performance Estimation Bias*, 4 PSYCHONOMIC BULL. & REV. 387, 387 (1997) ("Despite this overall finding of overconfidence on many tasks, performance across a sample of participants is almost always characterized by enormous variability. It is almost always the case that some participants show no global bias toward overconfidence.").

variability in responses and the presence of normative behavior on the part of some or many of the subjects is often neglected or downplayed. Indeed, a mythology of decision making as rampantly and fundamentally flawed has developed through the repeated use of standard research paradigms that are designed to show biased behavior and through the use of statistical methodology that stacks the decks in favor of finding biased behavior without concern for the practical importance of the behavior outside of the laboratory and, perhaps most surprisingly, *largely without documenting that any particular individuals actually acted irrationally in the experiments.* The end result of this lack of attention to how behavioral decision theory tests for irrationality is that findings from aggregated-data experiments—in which no particular individuals may be shown to act irrationally and in which the findings are of statistical significance but no demonstrated *practical* significance—are being improperly used as the basis for generalizations about the conduct of all legal actors across legal settings.

1. Between-Subjects Designs and Emphasis on the “Average Judge”

Behavioral decision theory research focuses on percentages and averages.⁷¹ Did a statistically significant percentage of experimental subjects fail to make the “rational” choice? Did an experimental manipulation that should not affect behavior according to rational actor theory (such as stating the choice options in terms of willingness to pay versus willingness to accept) lead to a statistically significant difference between the mean responses of the experimental groups on the dependent variable? Whether individuals within experimental groups predictably differed in their reasoning and choices and, if so, why, has been of little traditional concern to most behavioral decision researchers.⁷² Instead,

71. See Keith E. Stanovich, *Individual Differences in Cognitive Biases: Commentary on Krueger on Social-Bias*, 9 PSYCOLOGY 11, ¶ 7 (1998), at <http://www.cogsci.soton.ac.uk/psych-bin/newpsy?9.75> (noting that “proponents of the heuristics and biases approach (and equally their critics) have focused entirely on the central tendency of responses (usually the mean or modal performance tendency)”).

72. Early behavioral decision theory acknowledged the importance of personality variables in decision making, see Edwards, *supra* note 53, at 487-88 (reviewing studies of personality variables in decision making), and researchers have shown some continuing interest in the

researchers usually concern themselves with differences in average responses across experimental conditions:

A research strategy that emphasizes individual differences in cognitive and perceptual judgments is at variance with more classical research approaches. In the tradition of experimental psychology, the judgments obtained are summarized by averaging across all experimental subjects. Should individual differences among judges emerge, these differences would be treated as error and an "average judge" would be considered the most meaningful summary of all the judges. That approach has the advantage of ensuring generalizability to other judges, similar to the ones being studied, at the expense of ignoring individual differences. It is assumed that the mean judgment is representative of all the judges in the sample, or alternatively,

role of risk attitudes and risk perceptions in risky choice behavior, see Mellers et al., *supra* note 52, at 451-53 (reviewing behavioral decision research into risk attitudes and perceptions). But with the ascendancy of cognitive psychology in the middle to latter half of the twentieth century, along with its attendant search for fundamental cognitive processes to explain behavior, an individual differences/correlational approach to judgment and decision making received little attention for many years. See Alan Page Fiske et al., *The Cultural Matrix of Social Psychology*, in 2 THE HANDBOOK OF SOCIAL PSYCHOLOGY, *supra* note 9, at 918 ("Until recently, the cognitive view that has dominated the discipline of psychology in the last four decades typically adopted a mechanistic metaphor, describing the mind as a machine or computer that is the same in all times and places, while only the raw materials processed by the machinery or the data in the computer vary."); Keith E. Stanovich & Richard F. West, *Reasoning Independently of Prior Belief and Individual Differences in Actively Open-Minded Thinking*, 89 J. EDUC. PSYCHOL. 342, 342 (1997) ("Although investigators have observed group trends indicating belief-bias effects in the literatures of both social psychology and cognitive psychology, only recently have they turned their attention to individual differences in these effects.") (citations omitted); Elke U. Weber & Christopher K. Hsee, *Models and mosaics: Investigating cross-cultural differences in risk perception and risk preference*, 6 PSYCHONOMIC BULL. & REV. 611, 614 (1999) ("Within experimental and cognitive psychology, the study of individual differences has been relatively ignored, in favor of attempts to establish universal processes."). For a discussion of the development of cognitive psychology and the "cognitive revolution" see ERNEST R. HILGARD, *PSYCHOLOGY IN AMERICA: A HISTORICAL SURVEY* 221-67 (1987). For a discussion of the development of the subfield of social cognition within social and cognitive psychology see Hazel Markus & R. B. Zajonc, *The Cognitive Perspective in Social Psychology*, in 1 THE HANDBOOK OF SOCIAL PSYCHOLOGY, *supra* note 9, at 137-41. This trend regarding attention to individual differences seems to be changing. Indeed, ten years ago Herbert Simon, one of the pioneers of behavioral economics, called for greater attention to individual differences in information-processing and decision making. Herbert A. Simon, *Invariants of Human Behavior*, 41 ANN. REV. PSYCHOL. 1, 14-16 (1990) (discussing the potential insights from greater attention to individual differences across task domains, particularly with respect to differences between experts and novices).

that the judges are replications of one another within error of measurement.⁷³

Therefore, rather than examine individual variation in judgment and choice, behavioral decision theorists typically assume that, "to a first approximation, the thought processes of most un-institutionalized adults are quite similar,"⁷⁴ and any variation in subjects' responses is attributed to measurement or random error.⁷⁵

Consistent with this focus on average responses and with the assumption of substantial similarity in mental processing, many behavioral decision-making studies use "between-subjects" comparisons rather than "within-subjects" comparisons.⁷⁶ Thus, rather than examine how individual decision strategies or behaviors shift depending on the experimental manipulations, *the average responses of different groups of individuals under different experimental manipulations are compared*. As Professor Dawes notes, however, "[w]hen discussing anomalies, we are referring to inconsistencies that occur *within* individuals."⁷⁷ The between-subjects comparisons

73. Nancy Wiggins, *Individual Differences in Human Judgments: A Multivariate Approach*, in HUMAN JUDGMENT AND SOCIAL INTERACTION 110 (Leon Rappoport & David A. Summers eds., 1973); see also Simon, *supra* note 72, at 15 (noting that experimental studies of cognition "that take averages over sets of subjects" are not receptive to studies of individual differences).

74. Dawes, *supra* note 13, at 498 (quoting Baruch Fischhoff, *Strategic Policy Preferences: A Behavioral Decision Theory Perspective*, 39 J. SOC. ISSUES 133, 135 (1983)). Dawes adds that this assumption could be extended to all individuals, institutionalized or not. *Id.*

75. See Earl Hunt & Marcy Lansman, *Cognitive Theory Applied to Individual Differences*, in 1 HANDBOOK OF LEARNING AND COGNITIVE PROCESSES 81, 107 (W. K. Estes ed., 1975) ("In the past, many psychologists have seen individual differences as sources of error variance to be eliminated through the use of efficient experimental designs. The complexity of the causal chain between stimulus and response has forced them to turn their attention away from individual data to group means").

76. In a between-subjects design, the experimenter examines whether the mean, or average, response of one group of subjects exposed to one experimental condition differs significantly (in a statistical sense, where the likelihood of a difference being due to chance is estimated) from the mean response of a different group of subjects exposed to a different experimental condition. Alternatively, hypotheses may be tested using a within-subjects (or repeated-measures) design, in which individual subjects participate in each of the experimental conditions (or, less commonly, repeat trials within conditions) and then variations in behavior across conditions are examined, or using a mixed design, in which some independent variables of the experiment are examined between subjects and some are examined within subjects.

77. Dawes, *supra* note 13, at 503; see also Einhorn & Hogarth, *supra* note 20, at 82 (stating that "while group data may indicate large effects unless sufficient stimuli are

only allow us to infer that contradictions occur within individuals based on mean group scores.⁷⁸ That is, in a between-subjects design, no individual is actually shown to be inconsistent because no individual participates in all experimental conditions, but instead, individual inconsistency in judgments or choices is inferred if there are differences between the average responses of the different

sampled, no single individual can be shown to exhibit the bias"). But see Daniel Kahneman, *A psychological point of view: Violations of rational rules as a diagnostic of mental processes*, 23 BEHAV. & BRAIN SCI. 681, 682 (2000).

The between-subjects test of coherence is much stricter. It requires respondents to be disposed to produce the *same* judgments of probability, regardless of whether the questions about the week or the year are asked together or separately. Furthermore, coherence requires choices and beliefs to be immune to variations of framing and context. This is a lot to ask for, but an inability to pass between-subjects tests of coherence is indeed a significant flaw. Knowing rules and being able to apply them is good, but not sufficient, because much of life resembles a between-subjects experiment. Questions about preferences and beliefs arise one at a time, in variable frames and contexts, and without the information needed to apply relevant rules. A perfect reasoner whose judgments and choices are susceptible to framing and context will make many errors in the game of life.

Id. Some decision-making roles within the legal system would seem to resemble the role of subjects in repeated-measures/within-subjects designs (e.g., prosecutors, judges), whereas other roles seem closer to subjects in single-measure/between-subjects designs (e.g., juries, many civil litigants).

78. Dawes, *supra* note 13, at 503 ("Often ... we can only test different people with different questions and *infer* contradictions within people."). To be sure, between-subjects designs may provide a good examination of "average judge" performance on one-shot decision tasks and arguably provide "a clean test of the hypothesis that subjects rely on a given heuristic" by avoiding contextual effects arising from exposure to multiple experimental conditions. Daniel Kahneman & Amos Tversky, *On the Reality of Cognitive Illusions*, 103 PSYCHOL. REV. 582, 587 (1996); see also Gideon B. Keren & Jeroen G. W. Raaijmakers, *On Between-Subjects versus Within-Subjects Comparisons in Testing Utility Theory*, 41 ORGANIZATIONAL BEHAV. & HUM. DECISION PROCESSES 233 (1988) (arguing that within-subjects designs are often inappropriate in tests of expected utility theory). Where a deviation from rationality implies a violation of some rule of internal consistency, however, such as with loss/gain framing effects (equivalent expected utilities are judged differently depending on framing as a loss or gain) or the hindsight bias (where the probability of an event is judged greater after the outcome is known than before the outcome is known), between-subjects designs can only provide indirect evidence of deviations from rationality. See Kahneman & Tversky, *supra*, at 587 (discussing direct versus indirect tests of rational choice through within- and between-subjects designs); see also John C. Hershey & Paul J. H. Schoemaker, *Prospect Theory's Reflection Hypothesis: A Critical Examination*, 25 ORGANIZATIONAL BEHAV. & HUM. PERFORMANCE 395 (1980) (comparing between- and within-subject approaches to the study of prospect theory's postulates). Furthermore, between-subjects designs do not allow researchers to examine how "conflicts between rules and heuristics are resolved" within individuals. Kahneman & Tversky, *supra*, at 587 (footnote omitted).

treatment groups. For instance, in a test of context effects on preferences, if a majority of subjects in experimental group A chooses option x from option set $\{x,y\}$ but a majority of subjects in experimental group B chooses option y from option set $\{x,y,z\}$, then the behavioral decision theorist will claim that the subjects in this experiment exhibited irrational preference reversals (because the addition of option z should not alter the dominance of option x over option y), even though no individual subjects actually reversed their preferences between the two option sets (because no individuals made choices from both sets).

Given this practice of inferring individual irrationality from group differences, "[t]here is a real danger that the laws revealed by these mean scores are not true of any individual subject, but are the result of averaging two or more distinct, nonmodal patterns of behavior."⁷⁹ Moreover, "aggregating over attributes across individuals makes it hard or impossible to find some psychological regularities. This is so although the psychological process may be regular across subjects," for "different choices can result from the same regular decision process."⁸⁰

When individuals are tested in within-subjects experiments, in which individual inconsistency can actually be assessed, one finds that true intra-individual contradictions are not as prevalent nor as automatic as the legal decision theorists assume.⁸¹ For example,

79. Hunt & Lansman, *supra* note 75, at 107; see also Joachim Krueger, *Individual Differences and Pearson's r: Rationality Revealed?*, 23 BEHAV. & BRAIN SCI. 684, 684 (2000) ("Group-level analyses not only ignore systematic variations among people, they also work against the vindication of human judgment.") (citation omitted).

80. Ola Svenson, *Decision Making and the Search for Fundamental Psychological Regularities: What Can Be Learned from a Process Perspective?*, 65 ORGANIZATIONAL BEHAV. & HUM. DECISION PROCESSES 252, 260-61 (1996). Svenson recommends that

future [process] research should attempt to avoid confounding interindividual differences in goal elicitation and goal mapping with decision making processes as performed by single individuals and attempt to use within subjects designs whenever feasible. Otherwise we run the risk of only understanding processes about decision problems in which everyone agrees about values and the mapping of values (and of getting confused when there is a variety of goal representations).

Id.

81. Evidence of errors and biases found in between-subjects designs is often attenuated in experiments using within-subjects designs. See Kahneman & Tversky, *supra* note 78, at 586-87. For example, Stapel, Reicher and Spears found that using a within-subjects design for Kahneman and Tversky's "famous names" study, instead of the between-subjects design

Robyn Dawes discusses a within-subjects experiment in which alteration in the framing of a decision problem led to a statistically

used by Kahneman and Tversky, led to order effects in the operation of the availability heuristic, such that performance showed less evidence of bias over trials. Diederik A. Stapel et al., *Contextual determinants of strategic choice: Some moderators of the availability bias*, 25 EUR. J. SOC. PSYCHOL. 141, 144-47 (1995) (from Study 1). These authors emphasize the ability of subjects to adapt to the judgment task:

Our explanation for this progressive change is that, on the initial trial, there is a discrepancy between the task implied in the instruction to 'listen attentively' to the tapes and the actual task of estimating frequencies of men and women. As subjects became aware of the actual nature of the task, uncertainty is reduced, they adopt more appropriate judgmental strategies and the availability bias disappears.

Id. at 147.

When attention is drawn to independent variables or when subjects are given a chance to detect and correct possible errors, performance often moves toward the normative response. See Kahneman & Tversky, *supra* note 78, at 587; see also Michael H. Birnbaum & Barbara A. Mellers, *Bayesian Inference: Combining Base Rates With Opinions of Sources Who Vary in Credibility*, 45 J. PERSONALITY & SOC. PSYCHOL. 792, 796 (1983) (finding less than ten percent of subjects in single-judgment task using base rate information, but finding great majority of subjects using base rate information in multiple-judgment task); Baruch Fischhoff et al., *Subjective Sensitivity Analysis*, 23 ORGANIZATIONAL BEHAV. & HUM. PERFORMANCE 339, 356-57 (1979) ("On a theoretical level, the contrast between within-subject and between-subject designs clarifies the conditions under which various kinds of information are used. Although effectively ignored when embedded in the context of other information, both base-rate and validity information elicit somewhat appropriate responses when varied systematically (for most subjects). Of course, there is an implicit demand not to respond the same way each time. But the differences in responses would not be properly ordered if subjects did not know (or were not able to figure out) the meaning of that information for their inferences."); Carol A. Varey et al., *Judgments of Proportions*, 16 J. EXPERIMENTAL PSYCHOL.: HUM. PERCEPTION & PERFORMANCE 613, 623 (1990) ("The interpretation of base rate 'neglect' is based on the finding that in between-subjects comparisons, the effect of base rate is too small. In within-subjects comparisons, however, subjects use the base rate, and the evidence for a 'fallacy' disappears.") (citations omitted). In short, "by ignoring the nature of situational transitions, between-subjects designs can be misleading unless it is specifically desired to generalize to situations where judgments and choices are made in unique circumstances." Robin M. Hogarth, *Beyond Discrete Biases: Functional and Dysfunctional Aspects of Judgmental Heuristics*, 90 PSYCHOL. BULL. 197, 212 (1981).

On the other hand, the use of within-subjects designs may amplify biased or erroneous responses where the test is focused on determining the influence of irrelevant information on judgment and choice. In a within-subjects design, the implicit demand to use irrelevant but varying information will be greater because the salience of the irrelevant information will increase as this information is varied across within-subject conditions (e.g., individuating information that changes across conditions may become more salient and presumably more important than base rate information held constant across the conditions). See NORBERT SCHWARZ, COGNITION AND COMMUNICATION: JUDGMENTAL BIASES, RESEARCH METHODS, AND THE LOGIC OF CONVERSATION 23 (1996) ("In many base-rate studies, the apparent relevance of individuating information is further enhanced by the use of a within-subjects design.").

significant difference in choices across conditions, yet the statistically significant difference was attributable to only fifteen percent of the subjects changing their choices (i.e., only fifteen percent violated the normative principle that choice options with the same expected utility should be treated the same regardless of how the options are framed).⁸² Although this fifteen percent of the population may be a statistically significant number of respondents displaying inconsistency or "error" (i.e., we may have a high degree of confidence that this percentage was not simply the result of random error but rather the result of some psychological process at work in this subset of subjects), should we in this instance label the normative principle descriptively incorrect on the basis of this rather small, internally inconsistent minority?⁸³

Additionally, between-subjects designs present difficult interpretation problems because they allow subjects to impose their own, uncontrolled contexts onto the different tasks. As psychologists Michael Birnbaum and Barbara Mellers explain:

It is difficult to interpret the effect or lack of effect of variables that have been manipulated between subjects. The problem is

82. Dawes, *supra* note 13, at 503-04. An example of a gain/loss framing that should not affect choice under rational choice theory would be the presentation of a medical treatment program as saving 200 of 600 lives or as losing 400 of 600 lives, as in Kahneman and Tversky's well-known "Asian disease problem." See Michael J. Zickar & Scott Highhouse, *Looking Closer at the Effects of Framing on Risky Choice: An Item Response Theory Analysis*, 75 ORGANIZATIONAL BEHAV. & HUM. DECISION PROCESSES 75, 75-76 (1998) ("Most framing items are a variation on the classic 'Asian disease problem' of Tversky and Kahneman . . .") (citation omitted).

83. In between-subjects designs, one always finds some subjects who provide the normative response, even on the "hard" versions of the tasks where there have been no situational modifications designed to naturalize, clarify, or simplify the decision environment. See Keith E. Stanovich & Robert F. West, *Individual Differences in Rational Thought*, 127 J. EXPERIMENTAL PSYCHOL. 161, 161 (1998) ("What has largely been ignored is that—although the average person in these experiments might well display an overconfidence effect, underuse base rates, choose P and Q in the selection task, commit the conjunction fallacy, etc.—on each of these tasks, some people give the standard normative response."). Often the actual percentage of normative or correct responses in any given between-subjects decision-making experiment constitutes a substantial minority and, sometimes, even a majority of the responses given. See *id.* at 164-65 ("Unlike the case of the contrary-to-fact syllogisms task and statistical reasoning task, whereby a substantial number of people give the normative response, the gap between the descriptive and normative for the abstract selection task is unusually large. Less than 10% of the participants give the normative response on the abstract version."); see also *supra* note 70.

that when different subjects experience different stimulus contexts, responses cannot be compared without taking the different contexts into account. For example, by comparing judgments between subjects, it has been found that the number 450 can be judged greater than the number 550; however, one need not conclude that 450 actually seems greater than 550, because when judgments are compared within subjects, one finds that 550 is indeed judged greater than 450.⁸⁴

In other words, "[b]etween-subjects comparisons confound the stimulus and the context by allowing the stimulus to evoke its own context."⁸⁵ The lack of a uniform context paired with the averaging of results in between-subjects designs can lead to obvious contradictions that single individuals would not make if tested in all experimental conditions, such as Birnbaum and Mellers' example where the number 450 is judged greater than the number 550 in a between-subjects setting because the relative magnitude of the two numbers is never directly compared by the same individuals.

Therefore, "one should be extremely cautious when drawing inferences from between-subjects comparisons of judgments."⁸⁶ The averaging of results from between-subjects designs may mask important differences in the conduct of subjects and may cloud the

84. Birnbaum & Mellers, *supra* note 81, at 800. Berkeley and Humphreys make a similar point in their discussion of the importance of sources of uncertainty in decision problems:

[I]n general it is almost impossible to write descriptions of "real life" decision making situations which guarantee that all subjects will locate the problem within the same small world.

Comparisons of outputs from the formal model and the intuitive model of the decision problem may then reflect differences in problem structuring and decomposition, or different scaling of inputs, or use of different composition rules, the existence of incoherence, or any mixture of these.

Berkeley & Humphreys, *supra* note 25, at 225-27. Within-subjects designs present different context effect concerns. See Anthony G. Greenwald, *Within-Subjects Designs: To Use or Not To Use?*, 83 PSYCHOL. BULL. 314, 316-18 (1976).

85. Varey et al., *supra* note 81, at 623.

86. *Id.* The cognitive psychologist Ralph Hertwig likewise expresses strong concern about the inferences drawn from between-subjects designs: "Between-subjects designs are strange animals. They can lead to paradoxical, even absurd inferences." Barbara Mellers et al., *Do Frequency Representations Eliminate Conjunction Effects? An Exercise in Adversarial Collaboration*, 12 PSYCHOL. SCI. 269, 273 (2001) (citation omitted). Again, this is not to say that between-subjects designs are necessarily inappropriate—indeed, they may be the preferred design in some instances—but one must be cautious about the inferences drawn from them.

fact that often a good number of subjects act rationally in the very experiments that are cited by the legal decision theorists as proof of pervasive irrationality.

*2. Emphasis on Null Hypothesis Significance Testing:
Statistical Significance ≠ Practical Significance*

The mission of most behavioral decision theorists is to test hypotheses rather than estimate the prevalence of cognitive biases and errors within any given population:

[Behavioral decision research] is concerned with people in general. The studies are designed to demonstrate some effect Then further studies are done to analyze the effect. All that is required for these studies is some group of subjects who show the effect in question. Most researchers are not very interested in the prevalence of the effect.⁸⁷

To test their hypotheses, behavioral decision theorists employ null hypothesis significance testing (NHST), in which acceptance of the null hypothesis (which occurs when there is no statistically significant difference between the behavior of the control and experimental groups) is equated with rational behavior and a rejection of the null hypothesis (which occurs when there is a statistically significant difference between the behavior of the control and experimental groups) is equated with irrational behavior. Equating rationality with the null hypothesis in this way stacks the deck in favor of finding deviations from rationality.

Performing the rituals of [NHST], investigators stake their substantive claims about bias on appeals to statistical significance. The theoretical notion of rational or unbiased reasoning assumes the feeble status of a point specific null hypothesis, whereas bias lies in any significant departure from this point. [Subjects] have ample room to err, but only one place to be correct. Not surprisingly, NHST reveals that [subjects] "significantly" miss the point of no bias. With this asymmetric testing, there is a growing conviction that people are cognitively

87. BARON, *supra* note 20, at 45-46.

limited or miserly. Investigators demonstrate bias by detecting it. They rarely attempt to detect rational judgment. In the typical study, the detectability of bias increases with improved apparatus, the application of robust statistics, and sheer statistical power. These methodological improvements can dignify even tiny effects with the predicate of statistical significance.⁸⁸

As with the use of between-subjects designs, the great emphasis on NHST in behavioral decision research makes it difficult to determine the frequency and breadth with which irrational behavior occurs, because rejection of the null hypothesis (by a finding of a statistically significant difference between experimental groups) provides little information about the prevalence or strength of any given cognitive bias for three reasons: (1) when comparing group means in a between-subjects experiment, the ostensibly biased behavior of a fairly small number of participants may be the difference between the rejection of, and the failure to reject, the null hypothesis (as in Dawes' example discussed above, where inconsistent conduct by only fifteen percent of the subjects led to rejection of the null hypothesis); (2) even rather minor deviations from the rational choice prediction can lead to rejections of the null hypothesis depending on the sensitivity of the scales and other methodological features of the experiment; and (3) the level of statistical significance obtained (the "*p* value") depends on both treatment magnitude and sample size and therefore does not provide a true estimate of effect size (although both psychologists and legal decision theorists at times seem to forget this).⁸⁹ In short,

88. Joachim Krueger, *The Bet on Bias: A Foregone Conclusion?*, 9 PSYCOLOGY 1, ¶ 4 (1998), at <http://www.cogsci.soton.ac.uk/cgi/psyc/newpsy?9.46>; see also Reid Hastie, *A Review from a High Place: The Field of Judgment and Decision Making as Revealed in its Current Textbooks*, 2 PSYCHOL. SCI. 135, 138 (1991). Professor Hastie notes:

Many researchers also exhibit a detrimental tendency to plan empirical research to test the null hypothesis that human behavior is optimally rational, which frequently diverts research from the most important psychological issues. After all, precise null hypotheses are almost always refutable, with large enough samples of subjects or detailed enough measures of single subjects' behavior. The obsession with the rational null hypothesis has yielded a large harvest of "significant," but unimportant "proofs" that humans are irrational.

Id.

89. See generally Raymond S. Nickerson, *Null Hypothesis Significance Testing: A Review of an Old and Continuing Controversy*, 5 PSYCHOL. METHODS 241 (2000) (explaining concepts

relating to null hypothesis significance testing and critically reviewing controversies surrounding such testing). As Nickerson explains,

the value of p obtained from a null hypothesis statistical test is not the probability that [the null hypothesis] is true; to reject the null hypothesis at a confidence level of, say, .05 is not to say that given the data the probability that the null hypothesis is true is .05 or less. Furthermore, inasmuch as p does not represent the probability that the null hypothesis is true, its complement is not the probability that the alternative hypothesis . . . is true.

Id. at 246.

All too frequently, we find researchers comparing an F test that is significant at $p < .00001$ with one that is significant at $p < .05$ and concluding that the first experiment represents an impressive degree of prediction while the second experiment commands only passing interest. The problem with such a comparison of F statistics is that the size of the F ratio is affected by other factors in addition to the size of the treatment effects, the most obvious of which is *sample size*. Thus, a large F may imply that treatment effects are large, or that sample size was large, or that both factors are contributing to the observed value of F .

GEOFFREY KEPPEL, *DESIGN & ANALYSIS: A RESEARCHER'S HANDBOOK* 89 (2d ed. 1982); *see also* APA MANUAL, *supra* note 69, at 18 ("Neither of the two types of probability values reflects the importance or magnitude of an effect because both depend on sample size."); David Bakan, *The Test of Significance in Psychological Research*, 66 PSYCHOL. BULL. 423, 428 (1966) ("A common misinterpretation of the test of significance is to regard it as a 'measure' of significance."); E. Rae Harcum, *The Highly Inappropriate Calibrations of Statistical Significance*, 44 AM. PSYCHOLOGIST 964, 964 (1989) (noting that it is inappropriate to describe empirical results as "highly significant" because "(a) [t]he convention in psychology is to dichotomize results as significant or nonsignificant, and (b) [s]peaking in such terms] substitutes the probability of an effect for a measure of the size or importance of the effect"); Nickerson, *supra*, at 257 ("The value of p is not a reliable indication of the magnitude of an effect . . ."); Amos Tversky & Daniel Kahneman, *Belief in the Law of Small Numbers*, in A HANDBOOK FOR DATA ANALYSIS IN THE BEHAVIORAL SCIENCES: METHODOLOGICAL ISSUES 341, 348 (Gideon Keren & Charles Lewis eds., 1993) ("The emphasis on significance levels tends to obscure a fundamental distinction between the size of an effect and its statistical significance. Regardless of sample size, the size of an effect in one study is a reasonable estimate of the size of the effect in replication. In contrast, the estimated significance level in a replication depends critically on sample size.").

Like psychologists, legal scholars (including legal decision theorists and adherents to law and economics) often inappropriately describe a statistical test as yielding not just a statistically significant result but a "highly significant" or a "highly statistically significant" result, ostensibly to imply some special reliability or magnitude that may well not exist and that cannot properly be assessed with only a test of statistical significance. For some examples of the use of such inappropriate terminology in recent legal articles see Jean-Claude Bosch et al., *The Competitive Impact of Air Crashes: Stock Market Evidence*, 41 J.L. & ECON. 503, 514 (1998); Jim Chen & Edward S. Adams, *Feudalism Unmodified: Discourses on Farms and Firms*, 45 DRAKE L. REV. 361, 428 (1997); Kevin M. Clermont & Theodore Eisenberg, *Xenophilia in American Courts*, 109 HARV. L. REV. 1120, 1131 (1996); Guthrie, *supra* note 15, at 75; Korobkin, *supra* note 9, at 639; S. Raja Krishnamoorthi, *Making Local School Councils Work: The Implementation of Local School Councils in Chicago Public Elementary Schools*, 29 J.L. & EDUC. 285, 298 (2000); George L. Priest, *Private Litigants and the Court Congestion*

a "statistically significant" experimental result is not synonymous with a "practically significant" finding.⁹⁰

Most behavioral decision research provides little guidance to practical significance because of the predominance of NHST:

Many studies in the judgment literature merely indicate whether a bias exists according to a particular statistical level of probability. This knowledge, however, is not adequate information for a practitioner deciding whether to be concerned about a bias. Any attempts by a practitioner to correct or compensate for a bias invariably requires the investment of time, energy, and/or financial expenses—costs which the practitioner would prefer to expend only when it was likely that they would be exceeded by the benefits generated from the correction. This cost/benefit analysis faced by the practitioner is, by definition, situation-specific and subjective in nature. It depends upon the decision maker's utilities. There is no uniform index of practical significance.⁹¹

Problem, 69 B.U. L. REV. 527, 556 (1989); Jeffrey J. Rachlinski, *Gains, Losses, and the Psychology of Litigation*, 70 S. CAL. L. REV. 113, 154 (1996).

90. Rather, a finding of statistical significance indicates only that there is a less than one percent or five percent probability that the observed difference between groups as compared to the null hypothesis of no difference would likely occur by chance in samples randomly drawn from the same population. See KEPPEL, *supra* note 89, at 54 (noting that social scientists most commonly employ a .05 level of significance as the decision criterion for rejection of the null hypothesis); ROBERT ROSENTHAL ET AL., *CONTRASTS AND EFFECT SIZES IN BEHAVIORAL RESEARCH: A CORRELATIONAL APPROACH* 4 (2000) ("A result that is statistically significant at conventional levels is not necessarily 'practically significant' as judged by the magnitude of the effect."); Fred S. McChesney, *Statistics: The Language of Science (Part II)*, 9 KAN. J.L. & PUB. POL'Y 75, 84 (1999) ("Statistically 'significant' ... does not necessarily translate into 'important.' It is a term of art and it simply means, at a relevant level of confidence, we can exclude the possibility of a value of zero being attached to whatever phenomenon is of interest ...").

Paul Meehl bluntly evaluated psychology's fetish for NHST: "[T]he almost universal reliance on merely refuting the null hypothesis ... is a terrible mistake, is basically unsound, poor scientific strategy, and one of the worst things that ever happened in the history of psychology." Paul E. Meehl, *Theoretical Risks and Tabular Asterisks: Sir Karl, Sir Ronald, and the Slow Progress of Soft Psychology*, 46 J. CONSULTING & CLINICAL PSYCHOL. 806, 817 (1978).

91. Jay J. Christensen-Szalanski & Cynthia Fobian Willham, *The Hindsight Bias: A Meta-analysis*, 48 ORGANIZATIONAL BEHAV. & HUM. DECISION PROCESSES 147, 149 (1991) (citations omitted); see also John Ruscio, *Applying What We Have Learned: Understanding and Correcting Biased Judgment: Commentary on Krueger on Social-Bias*, 9 PSYCOLOGY 7, ¶ 2 (1998), at <http://www.cogsi.soton.ac.uk/cgi/psyc/newpsy?9.69> ("Researchers in the area of judgment and decision making have generated impressive lists of specific biases or cognitive

If the deleterious effects of cognitive biases found in the laboratory are not substantial or widespread, then little justification exists for system-wide reforms to address these biases.⁹² To assess practical significance, we must pay greater attention to the size of the effects

errors observed in human judges. The limitation of null hypothesis significance testing (NHST) in this field, as in many others, is that it makes a fairly unimportant, categorical decision regarding the presence or absence of a phenomenon while failing to provide an important, quantitative estimate of its magnitude.”); cf. Rachlinski, *supra* note 23, at 102 (“Absent an accurate quantification of the size of the bias that representativeness heuristic creates relative to the normative influence of statistical forensic evidence, a court could not be certain that withholding such evidence might do more to undermine the accuracy of the process than admitting such evidence. Likewise, without quantification of the size of the hindsight bias, it is not clear that withholding the fact of subsequent remedial measures is appropriate.”).

92. Cf. James P. Byrnes, *Analyzing Perspectives on Rationality and Critical Thinking: A Commentary on the Merrill-Palmer Quarterly Invitational Issue*, 39 MERRILL-PALMER Q. 159, 167-68 (1993).

[R]esearchers need to consider what would happen if someone fell prey to decision-making biases, failed to make fully correct conditional inferences, failed to reason like a working scientist, or failed to exhibit the dispositional tendencies of a good thinker. The point is that there is nothing to worry about if an unimportant skill is typically carried out in a suboptimal way, and there is no reason to expend a great deal of effort to try to improve this skill.

It would seem that in order to answer that “what would happen if” question, one would need to introduce two notions: (a) the frequency with which some fell prey to biases, failed to make conditional inferences, and so on; and (b) the frequencies of these behaviors that are minimally required for a given environment.

Id. Using Monte Carlo simulations, McKenzie examined the frequency and degree of error associated with use of intuitive judgment strategies in covariation assessment (assessing relations between variables) and Bayesian inference (updating of beliefs in light of new evidence). See Craig R. M. McKenzie, *The Accuracy of Intuitive Judgment Strategies: Covariation Assessment and Bayesian Inference*, 26 COGNITIVE PSYCHOL. 209 (1994). He concluded that, among other things: (1) “When examined under the general conditions of the simulations, all the apparent intuitive strategies performed much better than chance,” *id.* at 228, (2) “The simulation results also show that most of the intuitive strategies’ accuracy is strongly influenced by some simple environmental variables. . . . Conclusions regarding the accuracy of intuitive judgment strategies appear incomplete without taking environmental conditions into account,” *id.* at 229 (citations omitted), and (3) “[S]ubjects usually cope with covariation assessment and Bayesian inference by evaluating a single hypothesis, but that they compare two hypotheses when that is the perceived task. Although considering the alternative hypothesis does not entail using a normative strategy, it can lead to using intuitive strategies that appear relatively simple and are accurate across environments. Using these strategies may be the most efficient means of ensuring highly accurate judgments in these important tasks.” *Id.* at 232 (citations omitted).

associated with particular cognitive biases and the parameters under which the biases most reliably appear.⁹³

Moving behavioral decision theory research into the real world therefore necessitates a greater emphasis on assessing the actual impact of cognitive biases and errors on behavior and less emphasis on statistical significance. Factors having small but statistically significant effects in the laboratory may pale in comparison to the force of other factors in real world settings. Conversely, it may be possible to isolate those particular situations in which cognitive biases exert considerable negative causal force on particular people, even if the effect size appears small.⁹⁴ Until we have a better grasp on the level of unwanted influence of cognitive biases and errors relative to other factors in decision-making settings, it is difficult to determine whether the benefits of interventions aimed at averting or avoiding the effects of cognitive biases outweigh the costs of these interventions and reforms.

Meta-analytic reviews of experimental data are particularly useful in determining the known parameters and effect sizes associated with behavioral phenomena,⁹⁵ and meta-analyses provide

93. "An effect size estimate denotes the size of an effect in standard units. It provides a first step toward evaluating the practical importance of a finding." Kathleen McCartney & Robert Rosenthal, *Effect Size, Practical Importance, and Social Policy for Children*, 71 CHILD DEV. 173, 174 (2000). There are a variety of effect size indicators that may be used depending on the data in question. For discussion of various effect size indicators see *id.* at 174-75; Kenneth O. McGraw & S. P. Wong, *A Common Language Effect Size Statistic*, 111 PSYCHOL. BULL. 361, 361-65 (1992); R. Rosenthal & M. R. DiMatteo, *Meta-Analysis: Recent Developments in Quantitative Methods for Literature Reviews*, 52 ANN. REV. PSYCHOL. 59, 70-77 (2001). See generally ROSENTHAL ET AL., *supra* note 90, *passim*.

94. Cf. ROSENTHAL ET AL., *supra* note 90, at 4 (claiming that "even if the effect size were considered quantitatively unimpressive, it might still have important practical implications"). McCartney and Rosenthal assert that

neither experienced researchers nor experienced statisticians have a good intuitive feel for the practical meaning of common effect size estimates. As a result, most effects are labeled small or trivial, which is a discouraging conclusion; it is also a mistaken one. This state of affairs needs to be rectified not only with respect to policy analysis but also more generally.

McCartney & Rosenthal, *supra* note 93, at 174.

95.

Meta-analysis is more than a statistical technique; it is a methodology for systematically examining a body of research, carefully formulating hypotheses, conducting an exhaustive search and establishing inclusion/exclusion criteria for articles, recording and statistically synthesizing and combining data and effect sizes from these studies, searching for moderator and mediator variables to explain effects of interest, and reporting results.

a better means for assessing the real-world importance of such phenomena than do ordinary literature reviews.⁹⁶ In fact, meta-analytic review of research may reveal that the performance of experimental subjects over many studies is largely consistent with the proper utilization and processing of information from a legal standpoint, contrary to conjectures that might be drawn from data in any given individual study.⁹⁷ Unfortunately, only a small number of meta-analyses of behavioral decision research are available. As discussed below, the available meta-analytic data indicate, however, that the effect sizes associated with certain of the cognitive biases are small to moderate, as so often is the case with social science research.⁹⁸

Rosenthal & DiMatteo, *supra* note 93, at 62.

96. No experimental study can assess practical significance, but meta-analyses tend to provide better information for use in this assessment. *See* Christensen-Szalanski & Willham, *supra* note 91, at 149 ("Given the subjective nature of practical significance, neither meta-analysis nor testing for statistical significance can determine whether an observed effect is practically significant. However, meta-analysis can provide a measure of the size of the effect: information that is more useful than the indication that a bias exists according to a particular statistical level of probability."); *see also* Rosenthal & DiMatteo, *supra* note 93, at 63-68 (discussing advantages and criticisms of meta-analysis). Reliance on non-meta-analytic reviews of behavioral decision research for information about alleged errors and biases can be particularly misleading because such reviews are

usually based on the statistical tests reported in the original articles—significance or nonsignificance—with little or no attention paid to the patterns of the results or to the degree of support provided by the different experiments. As a consequence, potentially important information contained in the primary research literature is generally disregarded in the review process.

KEPPEL, *supra* note 89, at 77.

97. *See, e.g.,* Jennifer K. Robbennolt, *Outcome Severity and Judgments of "Responsibility": A Meta-Analytic Review*, 30 J. APPLIED SOC. PSYCHOL. 2575, 2602-03 (2000). Robbennolt notes:

The results of the present study are of interest in relation to assumptions on which the legal system is based. While the severity of the injury resulting from an action is legally relevant to some judgments that must be made by jurors, including damage-award judgments and the assessment of punishment, the legal system is structured such that injury severity should be largely irrelevant to other legal determinations, such as civil liability or criminal guilt. The results presented here indicate that decision makers may use information about the severity of an injury to inform both those decisions for which this information is relevant and those for which this information is not relevant. However, it appears that for those decisions in which outcome information should not play a role, the strength of the influence of outcome severity is weaker and is relatively small.

Id.

98. *See infra* sections II.A.2.a.-b (discussing meta-analyses of framing effect and hindsight bias research); *see also* Rostain, *supra* note 12, at 989 ("[T]he best empirical social science

a. Framing Effects

Rational choice theory holds that presenting risky choice options in different but formally equivalent terms should not affect choice, because people should evaluate options in terms of overall wealth or utility. One of the chief claims of legal decision theorists drawn from behavioral decision theory is that people do not evaluate choices in terms of overall wealth or utility but rather in terms of losses and gains around reference points, thus asserting that the rational choice account is descriptively incorrect. Legal decision theorists contend that the manner in which a risky choice is presented, or "framed," in relation to the status quo or some other reference point, has a large impact on what choices are made. The presumption is that there is a widespread tendency toward risk aversion for positively framed decisions and toward risk seeking for negatively framed decisions.⁹⁹ Thus, if presented with the choice between a

models cannot explain most of the variation seen in the variable under investigation."). The convention associated with one of the most common measures of effect size, the Pearson product moment correlation (r), is that "an r of .10 is small, .30 moderate, and .50 large." McCartney & Rosenthal, *supra* note 93, at 175. Within the social sciences, researchers rarely obtain r s reaching the .50 level associated with a large effect size, but this may be due in part to weak power in the experimental test rather than weakness of the effects. *Id.* at 175 (noting that "researchers examining the association between two variables seldom obtain r s as large as .50"); *id.* at 176 ("Not surprisingly, better measurement, design, and method all tend to produce larger effects."); *see also* Mark A. Mone et al., *The Perceptions and Usage of Statistical Power in Applied Psychology and Management Research*, 49 PERSONNEL PSYCHOL. 103, 113-18 (1996) (finding low statistical power in sample of studies published in seven leading applied psychology and management journals, which leads to an increased chance of failing to detect a significant effect size). *But see* Bakan, *supra* note 89, at 425-28 (discussing "a priori reasons for believing that the null hypothesis is generally false anyway," discussing problems with too great a reliance on significance testing, and stating that "[t]he fact of the matter is that there is really no good reason to expect the null hypothesis to be true in any population") (emphasis omitted).

99. *E.g.*, Hanson & Kysar, *Behavioralism I*, *supra* note 1, at 685 ("Throughout these [framing effect] studies, subjects have shown a remarkable proclivity to exhibit different preferences based solely on the manner in which the options are presented."); Hanson & Kysar, *Behavioralism II*, *supra* note 2, at 1572 ("We do know . . . that consumers *perceive* the product differently depending on the manufacturer's use of a non-substantive information frame."); Korobkin & Ulen, *supra* note 4, at 1104-05 ("When decisions options are perceived as 'gains' relative to the reference point, individuals are risk averse; that is, they prefer more certain options to gambles with the same expected value. But when decision options are perceived as 'losses' relative to the reference point, the same individuals will be risk-seeking; that is, they will prefer a gamble to the certain option when both have the same expected

sure gain of \$240 versus a 25% chance to gain \$1000 and a 75% chance to gain nothing, most people will supposedly take the sure gain. Yet, if presented with a sure loss of \$750 versus a 75% chance to lose \$1000 and a 25% chance to lose nothing, most people will supposedly take the gamble.¹⁰⁰

Despite the portrayal of framing effects as pervasive and powerful, a recent meta-analysis of 136 framing-effects studies found evidence of a framing effect across studies, "but this effect [was] of only small to moderate size."¹⁰¹ Even this small to moderate effect size associated with framing effects may be overestimated

value.") (footnote omitted); Kimberly D. Krawiec, *Accounting for Greed: Unraveling the Rogue Trader Mystery*, 79 OR. L. REV. 301, 319 (2000) ("A growing body of evidence indicates that individuals are loss averse, meaning that they tend to be risk averse to positively framed problems and risk seeking to negatively framed problems."); Prentice, *supra* note 2, at 156 ("One of the most pervasive influences on human judgment is framing. ... [P]references reverse completely depending on whether the chance is framed as a potential gain or a potential loss."); Rachlinski, *New Law & Psychology*, *supra* note 9, at 744 (stating that "plaintiffs, who regularly choose among gains, are risk-averse, while defendants, who regularly choose among losses, are risk-seeking").

100. This example is taken from Kahneman & Tversky, *Choices, Values, and Frames*, *supra* note 29, at 6.

101. Anton Kühberger, *The Influence of Framing on Risky Decisions: A Meta-analysis*, 75 ORGANIZATIONAL BEHAV. & HUM. DECISION PROCESSES 23, 42, 47 (1998) (examining 230 single effect sizes and reporting a between-conditions mean effect size of $d = .31$).

Another meta-analysis by Kühberger of the subset of framing studies that involve some version of Kahneman and Tversky's "Asian-disease" decision problem found significant bidirectional framing effects. These framing effects were found, however, in only approximately sixty percent of the subjects across all experiments and not in all subjects or even the vast majority of subjects as accounts of the framing research might lead one to believe. Anton Kühberger et al., *The Effects of Framing, Reflection, Probability, and Payoff on Risk Preference in Choice Tasks*, 78 ORGANIZATIONAL BEHAV. & HUM. DECISION PROCESSES 204, 219 (1999). The authors note:

The present meta-analysis found significant bidirectional framing effects. Presenting problems as gains leads participants to choose predominantly in a risk-averse manner (about 60% of all participants chose the sure gain and only 40% chose the risky gain). With losses, risk seeking predominates (about 40% of the participants chose the sure loss, while 60% chose the risky loss).

Id. Furthermore, this finding of an effect for sixty percent of the participants does not mean framing manipulations exert a large or determinative effect in any given situation (as noted above, Kühberger's larger meta-analysis found a small to moderate effect size). In fact, risk preferences on these Asian-disease types of problems "depend on the size and quality of the payoffs used." *Id.* at 227. Moreover, this finding does not necessarily mean that sixty percent of participants would individually change preferences across different frames of the disease problem.

because the reported studies may not be a representative sample, given the bias toward reports of successful replications.¹⁰²

What this means is that the influence of gain/loss framing on choice as a determinative factor may be fairly infrequent and, furthermore, that framing effects may be easily overwhelmed by other factors in the decision situation. Indeed, there is a "growing body of literature indicating that framing effects often are not as pervasive or robust as had once been considered."¹⁰³ The tenuous nature of framing effects is illustrated, for example, by the finding that simply adding pertinent social context information to standard framing-effect questions can greatly reduce or even eliminate framing effects.¹⁰⁴ To sum up, we find that the legal decision theorists overstate the certainty with which we can expect decision frames to influence choices.

b. The Hindsight Bias

Another prominent claim of the legal decision theorists is that legal actors often fall prey to the hindsight bias, which "causes decision makers to overestimate *ex post* their *ex ante* prediction

102. Kühberger et al., *supra* note 101, at 47; see also *infra* section II.A.3.

103. David R. Mandel, *Gain-Loss Framing and Choice: Separating Outcome Formulations from Descriptor Formulations*, 85 ORGANIZATIONAL BEHAV. & HUM. DECISION PROCESSES 56, 70 (2001) (citations omitted); see also Jerwen Jou et al., *An Information Processing View of Framing Effects: The Role of Causal Schemas in Decision Making*, 24 MEMORY & COGNITION 1, 10 (1996) ("The results . . . suggest that framing effects in real life may not be as prevalent as some decision studies have led us to believe.").

104. See James N. Druckman, *Using Credible Advice to Overcome Framing Effects*, 17 J.L. ECON. & ORG. 62 (2001). Druckman found that adding information about political party endorsements to the program options that subjects must choose between in the Asian disease problem and a variant on this problem greatly affected the likelihood of finding framing effects. *Id.* at 66-73 (reporting methods and results of Experiment 1); *id.* at 74-77 (reporting methods and results of Experiment 2). Druckman concludes:

The experimental results demonstrate that the availability of credible advice dramatically decreases, and sometimes eliminates, framing effects. Instead of basing their preferences on arbitrary question wording, people tend to rely on what they believe is credible information. Moreover, it seems plausible that outside of the laboratory, people do in fact access and use advice from others, especially in situations where they have ill-informed preferences. Many previous framing effect experiments ignore this possibility, and as a result, overstate the pervasiveness of framing effects.

Id. at 77.

about the likelihood of an event taking place."¹⁰⁵ Jeffrey Rachlinski, one of the leading legal decision theorists who has written extensively on the hindsight bias and its relevance for the legal system,¹⁰⁶ asserts that "[e]very published empirical test of the hindsight bias replicates the phenomenon."¹⁰⁷ The artful wording of this claim is itself interesting, in light of the fact that the hindsight bias has not been found under all experimental conditions in which it could have appeared.¹⁰⁸ More to the point, Professor Rachlinski fails to provide information about the overall effect size associated with the bias, moderating variables, or the likely practical significance of this bias—despite the fact that, to support his claim that "[p]sychologists have conducted nearly 130 experiments demonstrating the existence of the hindsight bias using a variety of different methods, materials, and subjects,"¹⁰⁹ Professor Rachlinski cites a meta-analytic study that provides just such information.¹¹⁰

When we look at the conclusions of this meta-analysis of over 120 hindsight experiments, we see why Professor Rachlinski may have omitted this information. This meta-analysis reveals that the hindsight bias is not necessarily the juggernaut that Professor Rachlinski and other legal decision theorists portray it to be:

The results of this meta-analysis revealed that the overall effect size ($r = .17$; corrected $r = .25$) of the hindsight bias is not large. This does not mean that the bias should be ignored since, depending upon the costs and benefits of making a correct and incorrect decision, effect sizes much smaller than this can still be of practical significance. At the same time, given the small observed effect size of the hindsight bias, its effect will more likely be washed out by the random error inherent in the real

105. Korobkin, *supra* note 6, at 48; *see also supra* note 45.

106. *E.g.*, Rachlinski, *Hindsight*, *supra* note 9; Rachlinski, *supra* note 23, at 66-82, 94-96; Jeffrey J. Rachlinski, *Regulating in Foresight Versus Judging Liability in Hindsight: The Case of Tobacco*, 33 GA. L. REV. 813 (1999).

107. Rachlinski, *supra* note 23, at 67.

108. *E.g.*, David Wasserman et al., *Hindsight and Causality*, 17 PERSONALITY & SOC. PSYCHOL. BULL. 30, 33 (1991) (subjects in the "chance explanation" condition did not show the hindsight bias). The assertion is particularly odd in light of Rachlinski's earlier acknowledgement of studies reporting a "reverse hindsight bias," though he did criticize these studies and downplay their results. *See* Rachlinski, *Hindsight*, *supra* note 9, at 580 n.31.

109. Rachlinski, *supra* note 23, at 67.

110. *Id.* at 68 & n.31 (citing Christensen-Szalanski & Willham, *supra* note 91, at 150).

world than would have occurred had the effect size been larger. Consequently, before issuing warnings to correct for the hindsight bias one needs to examine closely its potential impact for the specific situation of concern.

This study also showed that people who retrospectively claim that they "knew-all-along" what would occur are not necessarily exhibiting the hindsight bias. ... [M]any people faced with a dichotomous choice (e.g., act, not act), who retrospectively claim to have known all along that the observed outcome would occur, might have acted the same way even in *foresight*.

The magnitude of the hindsight effect can be moderated by the type of outcome information provided and the person's familiarity with a task. ...

In addition to these moderating variables, the location of the foresight probability estimate in respect to the threshold probability can affect the impact of the hindsight bias. ... Depending upon the familiarity of the task and type of outcome information presented, anywhere from a minimum of 0% to a maximum of 7-27% of the population may make different decisions because of the hindsight bias.¹¹¹

Although Professor Rachlinski cites this very study, he does not provide the above information, does not suggest that the reader examine the above information, nor does he qualify or explain for a lay audience his broad statement that the hindsight bias is a "robust phenomenon."¹¹² It seems apparent that inclusion of the above-quoted information inside Professor Rachlinski's article would detract from the legal importance assigned to the hindsight bias by Professor Rachlinski and would, at a minimum, counsel the reader to examine more closely claims assigning determinative causal force to the hindsight bias in the decisions of judges and juries.¹¹³

111. Christensen-Szalanski & Willham, *supra* note 91, at 162 (citations omitted).

112. See Rachlinski, *supra* note 23, at 68.

113. In another article, Professor Rachlinski contends that the hindsight bias "gives an average of a 15 percent 'boost' to the assessed probability in foresight—easily enough to enable the plaintiff to surmount the 50 percent threshold needed to establish liability in a close case." Rachlinski, *Hindsight*, *supra* note 9, at 606 (footnotes omitted). Professor Rachlinski does not specify his methodology for this estimation, nor does he cite any particular source to support the estimate. Perhaps Professor Rachlinski's estimate on probability alone is correct—though it seems bold in light of the Christensen-Szalanski and Willham meta-analytic results and given the great complexity of behavior—but at a minimum the reader should be provided with more information about his estimation methodology or

3. *The File Drawer Problem*

The perception that experimental subjects are hopelessly prone to error may be further heightened by the phenomenon known as the "file drawer problem":

Both behavioral researchers and statisticians have long suspected that the studies published in the behavioral sciences are a biased sample of the studies that are actually carried out. The extreme view of this problem, the "file drawer problem," is that the journals are filled with the 5% of the studies that show Type I errors, while the file drawers back at the lab are filled with the 95% of the studies that show nonsignificant (e.g., $p > .05$) results.¹¹⁴

There is evidence that psychologists are reluctant to submit and referees are reluctant to accept for publication articles failing to reject the null hypothesis.¹¹⁵

other bases for the claim (for example, what is the justification for extrapolating from hindsight bias studies of individual judgment to the context of group judgments by juries? See *supra* note 45). A full disclosure of the findings of the meta-analysis would have provided the reader with a more balanced view of research on the hindsight bias.

114. Robert Rosenthal, *The "File Drawer Problem" and Tolerance for Null Results*, 86 PSYCHOL. BULL. 638, 638 (1979) (citations omitted); see also ROBERT ROSENTHAL, JUDGMENT STUDIES: DESIGN, ANALYSIS, AND META-ANALYSIS 223-25 (1987) (providing an analytical tool for estimating "the degree of damage to any research conclusion that could be done by the file drawer problem").

115. Donald R. Atkinson et al., *Statistical Significance, Reviewer Evaluations, and the Scientific Process: Is There a (Statistically) Significant Relationship?*, 29 J. COUNSELING PSYCHOL. 189, 192 (1982) ("The results of this study tend to substantiate that, other things being equal, a research manuscript reporting statistically significant findings is more likely to be recommended for publication than is a manuscript reporting statistically nonsignificant findings."); Steven Kerr et al., *Manuscript Characteristics Which Influence Acceptance for Management and Social Science Journals*, 20 ACAD. MGMT. J. 132, 141 (1977) (finding that "a number of characteristics were considered to seriously impair publication chances," including "results which are statistically insignificant"); Michael J. Mahoney, *Publication Prejudices: An Experimental Study of Confirmatory Bias in the Peer Review System*, 1 COGNITIVE THERAPY & RES. 161, 173 (1977) (finding that "referee evaluations may be dramatically influenced by such factors as experimental outcome"); James Rotton et al., *Publication Practices and the File Drawer Problem: A Survey of Published Authors*, 10 J. SOC. BEHAV. & PERSONALITY 1, 9 (1993) (finding in a survey of behavioral scientists that the most frequently cited reason for deciding against publication was "the results were nonsignificant or unimpressive").

If studies failing to reject the null hypothesis are published less frequently than studies rejecting the null hypothesis (i.e., if these former studies remain in the “file drawer”), then we see a disproportionately low number of studies reporting rational behavior, because a finding of rational behavior is typically equated with a failure to reject the null hypothesis.¹¹⁶ So long as social and cognitive psychology maintains a “negativistic paradigm” that focuses on finding bias and error while assigning rationality to the easily rejected null hypothesis,¹¹⁷ we can expect the file drawer problem to bias publication against reports of rational behavior.

4. *Summary: The Aggregation-Generalization Fallacy*

When decision researchers and legal scholars treat a statistically significant result from a between-subjects experiment (i.e., a finding that group means differ significantly) as if it were a finding of a general effect of practical significance (e.g., that virtually all juries will fall prey to the hindsight bias), they commit what may be called the “aggregation-generalization fallacy”: confusing findings drawn from aggregate data with generally applicable findings or findings showing constant effects across individuals.¹¹⁸ “A *general-type proposition* asserts something which is presumably true of each and

116. See *supra* note 88 and accompanying text.

117. Krueger, *supra* note 88, ¶ 1 (“Much of the social psychological research effort has concentrated on the shortcomings of everyday behavior and thinking. Although topics such as attraction, altruism, and accuracy in person perception have also been of interest, the bulk of these findings and the interpretation of these findings reflect a negativistic paradigm.”); see also Ralph Hertwig & Peter M. Todd, *Biases to the Left, Fallacies to the Right: Stuck in the Middle with Null Hypothesis Significance Testing: Commentary on Krueger on Social Bias*, 11 PSYCOLOGY 20, ¶ 3 (2000), at <http://www.cogsci.soton.ac.uk/psyc-bin/newpsy?11.028> (“[P]sychology’s way of testing hypotheses stacks the deck in favor of rejecting predictions derived from normative principles (i.e., the null hypothesis). In other words: Asymmetric hypothesis testing has a built-in bias to find biases.”).

118. “Effects that are identical for all may be said to be *constant*, whereas those that vary with individuals may be said to be *variable*.” David Sohn, *Experimental Effects: Are They Constant or Variable Across Individuals?*, 9 THEORY & PSYCHOL. 625, 626 (1999). Legal decision theorists are not unique: “What frequently happens is that the scientist uses research findings about means, which . . . are aggregate propositions, to support scientific propositions of the *general* variety: ‘the data are treated as aggregates [i.e. means] while the experimenter is trying to infer general propositions.’” *Id.* at 627 (citation omitted); see also RAYMOND S. NICKERSON, REFLECTIONS ON REASONING 126 (1986) (“A common reasoning fallacy that is encountered in everyday life is that of over generalization, or what might be called the problem of imputing to a class the characteristics of one or a few of its members.”).

every member of a designable class. An *aggregate-type proposition* asserts something which is presumably true of the class considered as an aggregate."¹¹⁹ General-type propositions do not necessarily follow from aggregate-type propositions. Similarity in aggregate responses by two or more population subgroups does not mean that the effects of a manipulation are constant within, or general to, all members of the subgroups, much less that the aggregate and individual responses of other identifiable subgroups will be the same.

The legal decision theorists' strategy of inferring generalized patterns of behavior from aggregate data is ironic given this strategy's similarity to the economists' strategy of arguing that the effects of cognitive biases and errors will "wash out" when aggregate market behavior is examined (i.e., competition will weed out the irrational in the aggregate).¹²⁰ Legal decision theorists rebut economists' "errors wash out in the aggregate" argument by asserting that cognitive biases and errors are systematic and pervasive, rather than randomly occurring, and therefore, the influence of these nonrational behaviors persists even in the aggregate.¹²¹ Yet to support this claim that cognitive biases and errors are systematic and pervasive, legal decision theorists rely on aggregate data from between-subjects experiments that says very little about the strength, pervasiveness, and systematicness of cognitive biases and errors. In the end, this leaves legal decision theorists making an argument that mirrors that of the economists:

119. David Bakan, *The General and the Aggregate: A Methodological Distinction*, 5 PERCEPTUAL & MOTOR SKILLS 211, 211 (1955); see also Bakan, *supra* note 89, at 433. Bakan asserts:

One of the most unfortunate characteristics of many studies in psychology, especially in experimental psychology, is that the data are treated as aggregates while the experimenter is trying to infer general propositions. There is hardly an issue of most of the major psychological journals reporting experimentation in which this confusion does not appear several times; and in which the test of significance, which has some value in connection with the study of aggregates, is not interpreted as a measure of the credibility of the general proposition in which the investigator is interested.

Id.; see also Sohn, *supra* note 118, at 628 ("Psychologists seem to be content with statements of findings or propositions of the form that A affects B or an independent variable has an effect on a dependent variable. Such statements leave the reader in the dark about the characteristics of effects, whether they are constant or variable.").

120. See *supra* note 22 and accompanying text.

121. *Id.*

We should not be concerned with the fact that some individuals make normatively *correct* judgments or choices because, in the aggregate, these rational responses “wash out.” Both the economists’ and the legal decision theorists’ arguments raise unanswered empirical questions about the real-world significance of the bias and error research.

Inferring generalizations about behavior from aggregated data may be a defensible, if imprecise, practice, but only when this inferential leap and its basis are made clear.¹²² Virtually all of the claims of the legal decision theorists are general-type propositions inferred from aggregate-type propositions without explication as such, and without any discussion of how findings of statistical significance alone can support the generalizations.¹²³ One finds, at most, linguistic hedges, such as the data “suggest” some effect or some effect “generally” occurs, but not outright admissions that legal decision theory is founded on generalizations that are shakily inferred from aggregated data in between-subjects experiments.

The point here is not to assert that important anomalies in judgment and decision making never occur—they clearly do for some subjects, even in within-subjects designs¹²⁴—but rather to emphasize that subjects vary widely in their approaches to, and success on, behavioral decision-making tasks and to argue that generalizing from the conduct of the “average judge” composed from between-subjects experiments can be very misleading. The immediate danger of behavioral decision (and hence legal decision) theory’s focus on the “average judge” is twofold: (1) this approach

122. Cf. Bakan, *supra* note 119, at 212 (“The distinction between [general-type and aggregate-type] propositions does not preclude the possibility of using one type of proposition as a basis for *inference* with respect to the other type.”).

123. See section I.A (documenting the broad and pessimistic claims regarding human cognition made by legal decision theorists).

124. The conjunction fallacy (e.g., judging the likelihood of some person being a feminist bank teller to be greater than the likelihood that the person is simply a bank teller, which is logically more likely given that the latter set includes the former subset) is an example of an anomaly that may persist even in within-subjects designs. See Amos Tversky & Daniel Kahneman, *Extensional Versus Intuitive Reasoning: The Conjunction Fallacy in Probability Judgment*, 90 PSYCHOL. REV. 293, 299-300 (1983) (reporting significant percentages of subjects continuing to commit the conjunction fallacy even when presented with both response options in the “Linda problem” in an attempt to make the problem more transparent). However, commission of the conjunction fallacy may be quite dependent on how the problem is phrased and presented. See *infra* notes 155-57 and accompanying text.

implies incorrectly that no important or consistent individual differences are associated with different reasoning skills or decision tendencies¹²⁵ and that few subjects provide the correct responses in behavioral decision research; and (2) tests for statistical significance between treatment means or tests of the significance of different frequencies of responses between experimental groups—without the reporting of treatment magnitudes, and without detailed discussion of the absolute numbers of subjects getting the “right” versus the “wrong” answers and moderating or mediating variables—may lead one to overestimate the scope and importance of an experimental manipulation, and thus, the prevalence and systematicness of the anomaly being studied.

Because the bulk of behavioral decision research involves conclusions about average judges in between-subjects designs without estimates of the magnitude or prevalence of effects,¹²⁶ legal decision theorists should exercise much greater care in the generalizations drawn from behavioral decision theory research. The ultimate danger of this lack of care and precision in the claims being made about the irrationality of legal behavior is that “simple and general statements about a literature can become more authoritative than either the existing data or the claims made about the data by the original authors.”¹²⁷

125. See generally Mitchell, *supra* note 8.

126. See, e.g., Albert Erlebacher, *Design and Analysis of Experiments Contrasting the Within- and Between-Subjects Manipulation of the Independent Variable*, 84 PSYCHOL. BULL. 212, 212 (1977) (“Most experiments in personality or social psychology tend to be performed with a between-subjects design.”); Paul A. Klaczynski et al., *Goal-Oriented Critical Reasoning and Individual Differences in Critical Reasoning Biases*, 89 J. EDUC. PSYCHOL. 470, 472 (1997) (“Research on the biased application of reasoning strategies has typically been conducted with between-subjects designs. Reasoning biases are demonstrated when one group readily assimilates the enhancing information to their existing beliefs and a second group uses complex reasoning to reject threatening information.”) (citations omitted).

127. Jonathon J. Koehler, *The Base Rate Fallacy Myth*, 4 PSYCOLOQUY 1, ¶ 4 (1993), at <http://www.cogsci.soton.ac.uk/cgi/psyc/newpsy?4.49>.

Once in print, generalizations and bold claims take on a life of their own. Koehler analyzes, for instance, how “the ‘base rates are ignored’ misperception” survives in the heuristics and biases literature, and he notes that the base rate literature “has been simplified, misinterpreted, and selectively cited by observers, researchers, and reviewers alike,” and that “the ubiquitous summary statements of base rate neglect distort the empirical literature.” Koehler, *supra* note 67, at 5. Likewise, Vicente argues that

proponents of the heuristics and biases view have a thesis that they wish to promote vigorously—that human decision making is flawed. As a result, they may have inadvertently reconstructed the findings of seminal studies in the base

B. Features of Behavioral Decision Theory That Increase the Likelihood of Irrational Behavior in Experimental Settings

Certain features of behavioral decision theory research increase the likelihood that experimental subjects will provide the non-normative, or "irrational," response. Like the features just discussed, which mask the variation in behavior among experimental subjects and thus distort the perception of the prevalence of irrational behavior, this next set of features, by causing irrational behavior to be oversampled, also distorts the perceived frequency of irrational behavior.

1. Experiments Designed to Elicit Non-normative Responses

The large number of behavioral decision theory experiments that demonstrate irrational behavior should not be particularly surprising when one realizes that many of these experiments are designed specifically to elicit error. That is, many experiments have been designed and refined with an eye towards presenting subjects with situations and problems likely to cause at least some subjects to provide the normatively incorrect response.

rate literature in such a way as to be consistent with their thesis. In doing so, they have created yet another "virus" of miscitations of a well-known body of psychological research.

Kim J. Vicente, *The Perils of Reconstructive Remembering and the Value of Representative Design*, 19 BEHAV. & BRAIN SCI. 40, 40 (1996). For additional discussion and examples of distortion of research findings in secondary sources see James T. Todd & Edward K. Morris, *Case Histories in the Great Power of Steady Misrepresentation*, 47 AMER. PSYCHOLOGIST 1441 (1992); Kim J. Vicente & William F. Brewer, *Reconstructive remembering of the Scientific Literature*, 46 COGNITION 101 (1993); cf. Cass R. Sunstein, *On Academic Fads and Fashions*, 99 MICH. L. REV. 1251 (2001). Sunstein states:

The cautionary note is that in law and many other academic fields, ideas may spread and prosper, not because they are good, but because dozens, hundreds, or even thousands of imperfectly informed people have fortified the very signals by which they have been influenced. Whether bad ideas can prosper for a long time is another matter. Frequently good arguments and good evidence will puncture them, at least when there is agreement about the underlying criteria. But if the account here is correct, longevity, even for bad ideas, is hardly out of the question.

Id. at 1264.

[D]emonstrations of judgmental fallibility need careful interpretation lest they be seen as broad indictments of reasoning in general. In particular, it must be recognized that many of these demonstrations rely upon a narrow range of conditions that purposely are designed to highlight fallibility, and that their success in doing this does not necessarily mean that the demonstrated errors will occur in less contrived circumstances.¹²⁸

The leading proponents of the heuristic and bias view of cognitive error, Daniel Kahneman and Amos Tversky, forthrightly acknowledged that in their study of the conjunction fallacy, for example, they used word problems that they believed would elicit error regardless of whether the features of such problems were representative of problems in the real world: "Our problems, of course, were constructed to elicit conjunction errors, *and they do not provide an unbiased estimate of the prevalence of these errors.*"¹²⁹ This strategy is justifiable because Kahneman and Tversky, like many other behavioral decision theorists, sought to elicit errors to illuminate underlying psychological processes rather than determine the prevalence of irrational behavior.¹³⁰

128. Lee Roy Beach et al., *Assessing Human Judgment: Has it Been Done, Can it Be Done, Should it Be Done?*, in JUDGMENTAL FORECASTING 49 (George Wright & Peter Ayton eds., 1987). Jepson, Nisbett, and their colleagues make a similar point:

It seems likely . . . that the universality of the errors demonstrated to date is simply a result of the fact that subjects have been presented with problems for which the investigators themselves felt drawn to the erroneous line of reasoning. If one is simply seeking to demonstrate errors, this is an efficient strategy. But normatively correct principles of reasoning apply not only to hard problems, that is, those that investigators themselves have difficulty with, but also to easy problems.

Christopher Jepson et al., *Inductive Reasoning: Competence or Skill?*, 3 BEHAV. & BRAIN SCI. 494, 495 (1983).

129. Tversky & Kahneman, *supra* note 124, at 311 (emphasis added); see also Daniel Kahneman, *supra* note 60, at 142 (characterizing many of the heuristic and biases problems as "challenging brain teasers").

130. See Kahneman & Tversky, *supra* note 78, at 582 ("[T]he study of systematic error can illuminate the psychological processes that underlie perception and judgment."); Kahneman & Tversky, *supra* note 56, at 124 ("There are three related reasons for the focus on systematic errors and inferential biases in the study of reasoning. First, they expose some of our intellectual limitations and suggest ways of improving the quality of our thinking. Second, errors and biases often reveal the psychological processes and the heuristic procedures that govern judgment and inference. Third, mistakes and fallacies help the mapping of human intuitions by indicating which principles of statistics or logic are non-intuitive or counter-

One danger of a research strategy focused on error elicitation, of course, is that it may lead to a belief that normative errors are more common than they really are. Recall that the use of cognitive heuristics is adaptive, for heuristic processing often leads to the same result as does more systematic information-processing that requires greater effort or resources.¹³¹ Indeed, heuristic processing may lead to the normative result with a high degree of frequency.¹³² Because the use of a cognitive heuristic does not ensure normative error, researchers who seek to examine or demonstrate deviations from rationality must utilize experiments that sample those potentially unrepresentative situations in which the result given by a cognitive heuristic departs from the normative result.

A second danger is that the results from these more difficult or trickier problems may hold little or no relevance for other types of decision problems. The social psychologist and behavioral decision theorist Reid Hastie candidly commented on this point a few years ago:

I am struck by the unexceptional quality of the research that has been conducted by psychologists who focus on the goal of proving expected utility theory wrong. The economist's armchair "paradox" has been adopted as a replacement for the carefully controlled, labor-intensive experiment. The texts are filled with too many theoretical overgeneralizations and unconvincing interpretations of naturally-occurring behavior, based on

intuitive."); *id.* at 125 ("Errors and biases in judgment under uncertainty are the major source of data for the mapping of the boundaries of people's statistical intuitions."); Tversky & Kahneman, *supra* note 124, at 313 ("Our studies of inductive reasoning have focused on systematic errors because they are diagnostic of the heuristics that generally govern judgment and inference."); *see also* Kahneman, *supra* note 77, at 682 ("Contrary to a common perception, researchers working in the heuristics and biases . . . mode are less interested in demonstrating irrationality than in understanding the psychology of intuitive judgment and choice. The purpose of studying deviations of human judgments and choices from standard rational models . . . is to develop diagnostics for particular psychological processes."); *supra* text accompanying note 87.

131. *See, e.g.*, BAZERMAN, *supra* note 25, at 7 ("The availability heuristic can be a very useful managerial decision-making strategy, since instances of events of greater frequency are generally revealed more easily in our minds than events of less frequency. Consequently, this heuristic will often lead to accurate judgment."); Arkes, *supra* note 14, at 487-92 (discussing the adaptive nature of simplified information-processing approaches); *see also supra* note 26.

132. *See supra* note 92.

responses to 60-second brainteaser problems. (The brainteasers do serve a valid function in the rhetoric of scientific discourse; they have been remarkably effective at illustrating some of the cognitive processes that underlie decision making, and they have energized a field that was in the doldrums at the end of the sixties. But, I think it is time to move on to more substantial empirical analyses.)¹³³

The context-dependent nature of reasoning and choice argues against the assumption that performance on one type of problem will necessarily be the same on another type of problem:

It has been demonstrated that many adults do not have generally valid intuitions corresponding to the law of large numbers, the role of base rates in Bayesian inference, or the principles of regressive prediction. But it is simply not the case that every problem to which these rules are relevant will be answered incorrectly, or that the rules cannot appear compelling in particular contexts.

The properties that make formally equivalent problems easy or hard to solve appear to be related to the mental models, or schemas, that the problems evoke. . . . It appears that the actual reasoning process is schema-bound or content-bound so that different operations or inferential rules are available in different contexts. Consequently, human reasoning cannot be adequately described in terms of content-independent formal rules.¹³⁴

133. Hastie, *supra* note 88, at 138; cf. Ward Edwards, *Unfinished Tasks: A Research Agenda for Behavioral Decision Theory*, in *INSIGHTS IN DECISION MAKING: A TRIBUTE TO HILLEL J. EINHORN* 47 (Robin M. Hogarth ed., 1990) ("It does not seem reasonable to me to expect that the same processes and outputs will be found in personal decisions made casually, perhaps with trivial stakes, in highly sophisticated decisions made after intensive financial analysis and staff study, or in decisions made by committees.").

134. Kahneman & Tversky, *supra* note 56, at 129-30 (citations omitted); see also William M. Goldstein & Elke U. Weber, *Content and Discontent: Indications and Implications of Domain Specificity in Preferential Decision Making*, in *RESEARCH ON JUDGMENT AND DECISION MAKING*, *supra* note 20, at 566-617 (discussing the importance of content domain to preferential decision making); David A. Rettinger & Reid Hastie, *Content Effects on Decision Making*, 85 *ORGANIZATIONAL BEHAV. & HUM. DECISION PROCESSES* 336, 356 (2001) ("In summary, the content of a decision problem plays a major role in determining the information processing that participants used in making choices. Content domain affected judgments, self-reported decision strategy, mental representations of the decisions, and the ultimate choice.").

Therefore, even if all types of people responded identically across a variety of behavioral decision-making tasks (which they do not), we would still have to assess the "psychological realism" of this research to determine whether the processes operative in these tasks are likely to be operative on other experimental and real-world tasks. Psychological realism refers to "how well an experiment captures psychological processes like those occurring in everyday life,"¹³⁵ and is similar to the notion of construct validity. In social psychological research, construct validity concerns the extent to which the constructs of theoretical interest are validly operationalized in the research setting.¹³⁶

Given the interest in finding error rather than studying naturalistic decision making, often little effort is made to assess whether the mental operations primed in behavioral decision experiments are the same operations primed in real-world settings.¹³⁷ Consider, for instance, that if we force subjects to rate on a continuous scale their confidence in the accuracy of their answers on a general knowledge test, we can expect to find poor calibration of accuracy in a fair number of subjects on this continuous scale

135. ELLIOT ARONSON ET AL., *SOCIAL PSYCHOLOGY: THE HEART AND THE MIND* 58 (1994).

136. LOUISE H. KIDDER & CHARLES M. JUDD, *RESEARCH METHODS IN SOCIAL RELATIONS* 28 tbl.2.1 (5th ed. 1986) (construct validity directs one to the question of "[t]o what extent are the constructs of theoretical interest successfully operationalized in the research?"). For a more detailed discussion of construct validity see THOMAS D. COOK & DONALD T. CAMPBELL, *QUASI-EXPERIMENTATION: DESIGN & ANALYSIS ISSUES FOR FIELD SETTINGS* 59-70 (1979).

137. See Ebbe B. Ebbesen & Vladimir J. Konečni, *On the External Validity of Decision-Making Research: What Do We Know About Decisions in the Real World?*, in *COGNITIVE PROCESSES IN CHOICE AND DECISION BEHAVIOR* 21, 25-36 (Thomas S. Wallsten ed., 1980) (reviewing evidence that counsels against the assumption that simulations capture well real-world decision making); Hastie, *supra* note 52, at 664 ("Far more is known about the consequences of alternative decision problem representations (e.g. gain vs loss frames and summary vs unpacked event descriptions) than is known about the determinants of the representations. Thus, one key problem is understanding the determinants of decision frames and event descriptions and the impact of these differences on evaluations and judgments when a person is uncertain.") (citations omitted); George F. Loewenstein et al., *Risk as Feelings*, 127 *PSYCHOL. BULL.* 267, 267 (2001) ("Many choice theorists are deliberately agnostic about the psychological processes underlying the patterns of choice that their models predict."); see also Lola L. Lopes, *Between Hope and Fear: The Psychology of Risk*, in *RESEARCH ON JUDGMENT AND DECISION MAKING*, *supra* note 20, at 716 ("Theories that attempt to explain all of risky choice in the narrow terms of purely perceptual or purely cognitive or purely motivational mechanisms will necessarily miss much of what impels people toward or away from particular risks. The factors that influence human risk taking range from psychophysics to society and from fear to fun. So too should the psychology of risk.").

(with subjects tending to overestimate the percentage of correct answers). If people in the real world internally judge their confidence more categorically (e.g., high versus low confidence in their accuracy), however, then we may find that this less fine-grained calibration leads to a more realistic assessment of confidence for behavioral purposes.¹³⁸ Thus, something as simple as the response scale used in the research may drastically alter our view of the competence of the decision maker. Given the complexity of the real world and decision making in it, and particularly given behavioral decision theory's emphasis on finding deviations from rationality norms, we simply cannot assume that results in behavioral decision research easily translate to conduct outside the laboratory.

Behavioral decision theorists, and consumers of their work, must resist two curses described by Baruch Fischhoff. First is the "curse of context": "We would like to interpret [subjects'] responses as reflecting deep-seated values, of the sort that come from a lifetime of intense involvement with real-world decisions. Yet, we set

138. The overconfidence effect occurs when subjects provide answers to questions, typically general knowledge questions, and then provide estimates, or probability judgments, of the likelihood that their answers are correct; "over a wide range of conditions, subjects' average probability judgments exceed the proportions of items they answer correctly." J. Frank Yates et al., *Beliefs about Overconfidence, Including Its Cross-National Variation*, 65 ORGANIZATIONAL BEHAV. & HUM. DECISION PROCESSES 138, 138 (1996). Whether people are irrationally overconfident in their ability on judgment and decision-making tasks or whether this "overconfidence" is simply the artifact of an improper approach to the assessment of judgmental confidence is the subject of serious debate among behavioral decision theorists. See generally Lyle A. Brenner et al., *Overconfidence in Probability and Frequency Judgments: A Critical Examination*, 65 ORGANIZATIONAL BEHAV. & HUM. DECISION PROCESSES 212 (1996); Robyn M. Dawes & Matthew Mulford, *The False Consensus Effect and Overconfidence: Flaws in Judgment or Flaws in How We Study Judgment?*, 65 ORGANIZATIONAL BEHAV. & HUM. DECISION PROCESSES 201 (1996); Dale W. Griffin & Carol A. Varey, *Towards a Consensus on Overconfidence*, 65 ORGANIZATIONAL BEHAV. & HUM. DECISION PROCESSES 227 (1996); see also Thomas S. Wallsten, *An Analysis of Judgment Research Analyses*, 65 ORGANIZATIONAL BEHAV. & HUM. DECISION PROCESSES 220 (1996) (discussing different approaches to the study of overconfidence).

Consider, as another example, a study in the context of risk assessment that examined respondents' subjective meaning in choosing the fifty-percent option on response scales. Rather than an assessment that there is a fifty-percent probability that a risky event will occur, the choice of the fifty-percent option may mean that the respondents simply have no good idea of the odds. See Baruch Fischhoff & Wändi Bruine de Bruin, *Fifty-Fifty = 50%?*, 12 J. BEHAV. DECISION MAKING 149, 160 (1999) ("It appears, thus, that the availability of the common phrase 'fifty-fifty' affects the use of the 50 response as an expression of epistemic uncertainty.").

minimalist problems before subjects and expect them to resist the temptation to impute a context."¹³⁹ Second is the "curse of cleverness": "We pride ourselves on devising just the right tasks for evaluating competing theories of human behavior, tasks that our predecessors were unable to concoct. Then, we expect subjects immediately to discern the structure of these tasks and decide what response is right for them."¹⁴⁰ When students of judgment and decision making fall prey to these two curses, they see themselves as too smart and their subjects as too dumb.

2. Lack of Feedback and Learning Opportunities

One of the clearest ways in which many behavioral decision experiments differ from many decision settings in the real world is in the lack of feedback or learning opportunities provided to subjects. This lack of feedback and learning is a natural consequence of the frequent use of between-subjects designs in behavioral decision experiments, in which subjects participate in only one experimental condition expressing typically only one choice or judgment. This lack of feedback and learning opportunities increases the likelihood that irrational responses will be found, for it denies subjects the opportunity to understand the larger context in which their decisions are being made (contrary to subjects who participate in repeated-measures designs) and to determine the success of a course of behavior.¹⁴¹

Experimental economists have been particularly critical of behavioral decision theory's emphasis on one-shot decision situations that fail to test for feedback, learning, and market interaction effects. Experimental economists argue that, while studies using between-subjects designs may "have relevance to measuring people's preference attitudes," such studies "provide no basis for extrapolation to behavior in markets."¹⁴² This failing is serious, at least when behavioral decision theorists seek to describe

139. Baruch Fischhoff, *The Real World: What Good Is It?*, 65 ORGANIZATIONAL BEHAV. & HUM. DECISION PROCESSES 232, 245 (1996).

140. *Id.*

141. See *supra* notes 78-81 and accompanying text.

142. Peter Knez et al., *Individual Rationality, Market Rationality, and Value Estimation*, 75 AMER. ECON. REV. 397, 398 (1985).

behavior in repetitive markets, because "[m]ost (but not all) experimental markets show some learning effects over time with equilibrium behavior quite different from start-up behavior."¹⁴³ Accordingly, experimental economists emphasize the study of behavior in experimental markets rather than responses to simple gambles or single word problems as so often found in psychological studies of judgment and choice.¹⁴⁴

143. *Id.*; see also *id.* at 400-01 (reporting decreased disparities in willing-to-pay and willing-to-accept prices over time in repeated-measures design); Vernon L. Smith, *Economics in the Laboratory*, 8 J. ECON. PERSPECTIVES 113, 118 (1994) (noting "tendency for rational behavior to emerge in the context of a repetitive market institution" and that "[i]n many experimental markets, poorly informed, error-prone, and uncomprehending human agents interact through the trading rules to produce social algorithms which demonstrably approximate the wealth maximizing outcomes traditionally thought to require complete information and cognitively rational actors") (footnote omitted); Vernon L. Smith, *Theory, Experiment and Economics*, 3 J. ECON. PERSPECTIVES 151, 165 (1989). Smith claims:

I think we economists need to accept these replicable empirical results [from behavioral decision theory] as providing meaningful measures of *how people think about economic questions*. For their part, psychologists need to accept the dominating message in experimental research on the performance of a wide variety of bidding, auctioning and customer (posted price) markets: markets quite often "work" in the sense that over time they converge to the predictions of the economist's paradigm.

Id. (footnote omitted). But see George A. Akerlof & Janet L. Yellen, *Can Small Deviations from Rationality Make Significant Differences to Economic Equilibria?*, 75 AMER. ECON. REV. 708, 708 (1985) ("This paper constructs a number of examples which show that small deviations from rationality can have first-order consequences in microeconomic analysis."); Thomas Russell & Richard Thaler, *The Relevance of Quasi Rationality in Competitive Markets*, 75 AMER. ECON. REV. 1071, 1081 (1985) ("The notion that individual irrationalities will disappear in the aggregate must be rejected."); Thomas Russell & Richard Thaler, *The Relevance of Quasi Rationality in Competitive Markets: Reply*, 77 AMER. ECON. REV. 499 (1987) (reporting correction to their 1985 paper).

144. See Richard H. McAdams, *Experimental Law and Economics*, in 1 ENCYCLOPEDIA OF LAW AND ECONOMICS, *supra* note 55, at 540-41 ("The idea of creating a market in the laboratory helps to distinguish experimental economics from a similar experimental literature in social and cognitive psychology. . . . Psychologists tend to ask their subjects questions. . . . By contrast, economic experimenters employ conditions that create different incentives for behavior and then observe how subjects react."); Vernon L. Smith, *Experimental Economics: Reply*, 75 AMER. ECON. REV. 265, 267 (1985) (Advancing the hypothesis that "subjects are more rational (in the sense of received decision theory under uncertainty) in the context of laboratory markets, than when responding to questionnaire choices among prospects, because of Heiner's conjecture that 'Exchange environments also enable agents to interact with each other in an organized and often repeated fashion.'"). But see Medin & Bazerman, *supra* note 23, at 536 ("Experimental economics is filled with demonstrations of performance improvement through multiple trials—with convergence on the economically rational solutions. In contrast, we suggest that much of experimental economics consists of contrived experiments created in order to show convergence. In fact, many [behavioral decision

Legal decision theorists have recognized this problem but have sought to head it off as a criticism by arguing that "effective" feedback and learning opportunities are probably rare or hard to obtain.¹⁴⁵ Yet even this qualified defense is based on mixed evidence, and its advocates must acknowledge that "learning can overcome some cognitive biases, at least for some tasks."¹⁴⁶ The evidence showing the beneficial effects of feedback and training supports the view that single-decision settings differ in important ways from repeated-decision settings and thus counsels against an assumption that behavior is the same in both settings.

3. *Conversational Cues and Demand Characteristics*

Implicit conversational cues and demands in experimental settings also lead subjects to provide normatively incorrect responses that they otherwise would not provide. A great deal of research in the last ten years has shown that subjects often provide non-normative responses in behavioral decision-making experiments not because the subjects are incapable of acting rationally but because the experimental situation indicates, or communicates, that the non-normative response is the correct or desired answer under the circumstances.

As early as 1982, Kahneman and Tversky noted the importance of the interaction between subject and experimenter in evaluating the behavior of subjects in behavioral decision research:

The presence of an error of judgment is demonstrated by comparing people's responses either to an established fact (e.g., that the two lines are equal in length) or to an accepted rule of arithmetic, logic or statistics. However, not every response that appears to contradict an established fact or an accepted rule is a judgmental error. The contradiction could also arise from the

research] results do live through shockingly high levels of feedback and multiple trials.") (citations omitted).

145. For instance, Professor Garvin acknowledges that "some studies have shown quite substantial learning effects." Garvin, *Of Risk, Duress, and Cognition*, *supra* note 18, at 168. But then he asserts that "[s]till, effective learning is far from common." *Id.* Professor Garvin at least acknowledges the evidence on both sides of the issue and closes on an equivocal note, which is rather uncommon for legal decision theory scholarship. *See id.* at 168-70.

146. *Id.* at 170.

subject's misunderstanding of the question, or from the investigator's misinterpretation of the answer. The description of a particular response as an error of judgment therefore involves assumptions about the communication between the experimenter and the subject. ... The student of judgment should avoid overly strict interpretations, which treat reasonable answers as errors, as well as overly charitable interpretations, which attempt to rationalize every response.¹⁴⁷

Later in the same article, Kahneman and Tversky emphasized that

we cannot safely assume that "experimental conversations" in which subjects receive messages and answer questions will simulate the inferences that people make in their normal interaction with the environment. Although some judgments in everyday life are made in response to explicit questions, many are not. Furthermore, conversational experiments differ in many ways from normal social interaction.¹⁴⁸

To address these concerns, Kahneman and Tversky suggested a research program focused on the investigator-participant interaction:

We conclude that the conversational aspect of judgment studies deserves more careful consideration than it has received in past research, our own included. We cannot always assume that people will or should make the same inferences from observing a fact and from being told the same fact, because the conversational rules that regulate communication between people do not apply to the information that is obtained by observing nature. It is often difficult to ask questions without giving (useful or misleading) clues regarding the correct answer, and without conveying information about the expected response.¹⁴⁹

This "experimental conversations" research program focused on the heuristics and biases research did not begin in earnest until several

147. Kahneman & Tversky, *supra* note 56, at 124.

148. *Id.* at 132.

149. *Id.* at 135.

years after the appearance of Kahneman and Tversky's article,¹⁵⁰ but this research program has now revealed that many of the experimental subjects' responses labeled as evidence of cognitive bias or error were actually reasonable responses to conversational cues contained in the experimental stimuli.¹⁵¹

Norbert Schwarz concludes his important earlier review of this experimental conversations research program in the following way:

In essence, social cognition researchers have conceptualized the judge as doing his or her cognitive work in social isolation. In many instances, the obtained findings provided a less than flattering portrait of our respondents. Apparently, people are happy to offer meaningless opinions on nonexistent issues and

150. Study of conversational influences in attribution and person perception processes preceded the study of conversational influences in the heuristics and biases literature. See Norbert Schwarz, *Judgment in a Social Context: Biases, Shortcomings, and the Logic of Conversation*, 26 *ADVANCED EXPERIMENTAL SOC. PSYCHOL.* 123, 123-24 (1994). Schwarz states:

On close inspection, it seems that much of what we consider to reflect biases in human judgment, artifacts in attitude measurement, and so on, may actually reflect researchers' ignorance regarding the conversational context of human judgment, rather than serious shortcomings on the side of our subjects. Accordingly, social cognition research may greatly benefit from a fuller consideration of the social context in which humans conduct much of their thinking about social as well as non-social stimuli. . . .

At present, social cognition researchers have primarily paid attention to conversational aspects of human judgment in exploring attribution processes and the impact of different audiences on the encoding and recall of person information.

Id. (citations omitted). By the mid-1990s, the research program remained nascent. See Denis J. Hilton, *The Social Context of Reasoning: Conversational Inference and Rational Judgment*, 118 *PSYCHOL. BULL.* 248, 249 (1995) ("Although many researchers have noted that widely shared assumptions about cooperative communication may license interpretations of the experimental task that the experimenter may not have intended, work on this topic remains fragmentary and underdeveloped in nature.") (citations omitted).

151. See Norbert Schwarz, *Social Judgment and Attitudes: Warmer, More Social, and Less Conscious*, 30 *EUR. J. SOC. PSYCHOL.* 149, 152 (2000) (noting that "a growing body of research demonstrates that many of the apparent biases and shortcomings of human judgment are attenuated or eliminated when this basic misunderstanding [about the nature of communication in a research setting] is taken care of") (citations omitted); Ben R. Slugoski & Anne E. Wilson, *Contribution of Conversation Skills to the Production of Judgmental Errors*, 28 *EUR. J. SOC. PSYCHOL.* 575, 576 (1998) (noting that "a sizeable body of research has now accumulated to demonstrate that much of what we once considered to reflect 'hard-wired' information processing mechanisms in human judgment (e.g., cognitive biases, attitudinal shifts, causal attributions) actually reflects the operation of communicative goals and strategies that people habitually rely upon in order to comprehend and generate meaningful conversation") (citations omitted).

are biased by irrelevant material such as the numeric values of a rating scale or the response alternatives of a frequency question. Moreover, they are apparently more willing to use worthless personality information and to ignore meaningful base rates, or to draw strong dispositional inferences despite obvious situational pressures, and so on. As soon as we conceptualize the assessed judgments as part of an ongoing conversation, however, the often dramatic findings seem less surprising. Rather, research participants seem to do what they would rightly be expected to do in any other conversation: They assume that *our* utterances as researchers are meaningful, that we do not ask questions about things that don't exist, that we do construct meaningful rather than arbitrary scales, and so on. Moreover, they try to make sense of our utterances and of our research instruments on the basis of these assumptions.

....

The apparent biases and errors that subjects commit by relying on conversational maxims are less likely to result in mistakes in everyday contexts where communicators try to conform to conversational norms, provide information that is relevant to the judgment at hand, and make the task one that is clear rather than ambiguous—and in which recipients are indeed expected to use contextual cues to disambiguate the communication, should the communicator not live up to the ideal. Thus, the behavior that may lead to errors in the experimental context may be adaptive in everyday settings.¹⁵²

In other words, when we examine cognitive processes in artificial or contrived settings in which ordinary rules of conversation do not apply *unbeknownst to the experimental subjects*, we should not be surprised to find that the operation of these cognitive processes appears biased or appears to lead to error.¹⁵³ However, when we

152. Schwarz, *supra* note 150, at 154-56.

153. The rules of communication that behavioral decision researchers are typically accused of violating are the conversational maxims of quantity, quality, relation, and manner, as explicated by Paul Grice. See H. P. Grice, *Logic and Conversation*, in *THE LOGIC OF GRAMMAR* 64, 67 (Donald Davidson & Gilbert Harman eds., 1975); see also Slugoski & Wilson, *supra* note 151, at 576 (stating that "researchers have come to rely on Grice's (1975) well known model of cooperative conversational exchanges delineating the obligations falling on speakers and expectations held by hearers in order to achieve common understanding") (citation omitted). The maxim of quantity prescribes: "(1) 'Make your contribution as informative as is required (for the current purposes of the exchange),' and possibly (2) 'Do not make your contribution

restore conversational sense to the questions and possible answers or when we make clear to the subjects that ordinary norms and conventions of communication do not apply, we find improvement in subjects' performance on tests of rational thinking—sometimes substantial improvement.¹⁵⁴

Thus, "[w]hen participants are aware that the usual 'guarantee of [conversational] relevance' does not hold, they prefer base-rate information over nondiagnostic individuating information; are less likely to show the fundamental attribution error; and are less likely to be influenced by misleading questions in eyewitness testimony."¹⁵⁵ Similarly, conversational cues paired with ambiguities in the questions may explain the behavior of a significant percentage of experimental subjects who ostensibly commit

more informative than is required." Grice, *supra*, at 67. The maxim of quality prescribes: "Try to make your contribution one that is true," and two more specific maxims: (1) 'Do not say what you believe to be false' [and] (2) 'Do not say that for which you lack adequate evidence.'" *Id.* The maxim of relation simply prescribes: "Be relevant." *Id.* The maxim of manner includes the supermaxim: "Be perspicuous" and various maxims such as (1) 'Avoid obscurity of expression'[,] (2) 'Avoid ambiguity'[,] (3) 'Be brief (avoid unnecessary prolixity)'[,] and] (4) 'Be orderly.'" *Id.* These maxims flow from the overarching "Cooperative Principle," which prescribes: "Make your conversational contribution such as is required, at the stage at which it occurs, by the accepted purpose or direction of the talk exchange in which you are engaged." *Id.* Norbert Schwarz discusses these maxims and how they may be violated in research settings. See Schwarz, *supra* note 150, at 124-27; see also Hilton, *supra* note 150, at 249-54 (discussing "[p]roperties of conversational inference"); Robert M. Krauss & Chi-Yue Chiu, *Language and Social Behavior*, in 2 THE HANDBOOK OF SOCIAL PSYCHOLOGY, *supra* note 9, at 41, 42-48 (contrasting four different paradigms or conceptions of interpersonal communication, including the intentionalist paradigm into which Grice's approach falls).

154. See Hilton, *supra* note 150, at 264. Hilton notes:

[U]ndermining the assumption of intentionality [in communication in experimental settings] reduces biases in such diverse tasks as Piagetian conservation tasks, the engineers-and-lawyers task of Tversky and Kahneman, and the leading questions paradigm of Loftus and Palmer. Another example is the similar kinds of experimental demands created by juxtaposing questions comparing specific and general quantities in tasks involving conservation, object naming, probability judgments, and survey judgments of life satisfaction. Conversational inference thus has general features that emerge in a wide range of tasks.

Id. (citations omitted).

155. Schwarz, *supra* note 151, at 152 (citations omitted); see also Schwarz, *supra* note 150, at 128-34 (discussing conversational influences in base rate and causal attribution studies); Slugoski & Wilson, *supra* note 151, at 579 ("Taken together, these studies strongly suggest that participants' pragmatic assumptions play an important role in the interpretation of 'relevance' in the experimental setting and hence their likelihood of producing the so-called 'base-rate error.'") (citation omitted).

conjunction errors, commit the fundamental attribution error (a bias to favor dispositional explanations of behavior and ignore situational explanations), show primacy effects (giving greater weight to early items in a sequence), show framing effects, perform poorly on the Wason selection task (a test of conditional reasoning), show overconfidence in judgments about the accuracy of their own performance, and exhibit the "dilution effect" (in which providing nondiagnostic information "dilutes" the impact of diagnostic information in person perception tasks).¹⁵⁶ Hence, we see participants responding quite reasonably to information given to them as they try to figure out exactly what the investigator wants of them: Surely the investigator would not suggest that Linda is a feminist bank teller unless Linda is a feminist bank teller, as unlikely as that may otherwise seem.¹⁵⁷

156. See Anton Khberger, *The Framing of Decisions: A New Look at Old Problems*, 62 ORGANIZATIONAL BEHAV. & HUM. DECISION PROCESSES 230, 231-36 (1995) (discussing the possible role of missing information in framing effect problems and presenting evidence that more complete descriptions of problems may reduce or eliminate framing effects); Slugoski & Wilson, *supra* note 151, at 579-81 (discussing research showing that conversational influences in the experimental setting may lead subjects to exhibit conjunction errors, dilution effects, primacy effects, and errors on the Wason selection task); Edward F. Wright & Gary L. Wells, *Is the Attitude-Attribution Paradigm Suitable for Investigating the Dispositional Bias?*, 14 PERS. & SOC. PSYCHOL. BULL. 183, 188 (1988) ("The results of this study supported the hypothesis that subjects' attributions in experiments employing the attitude-attribution paradigm are influenced by procedure-generated pressures. . . . The modified procedure mitigated, but did not eliminate the dispositional bias. . . . [T]he magnitude of the bias effect obtained with the traditional attitude-attribution paradigm has probably been exaggerated by the violation of the cooperativeness rule."); Ilan Yaniv & Dean P. Foster, *Precision and Accuracy of Judgmental Estimation*, 10 J. BEHAV. DECISION MAKING 21, 30-31 (1997) (discussing how expressions of overconfidence on some scales may be the result of subjects seeking to resolve a trade-off between judgmental accuracy and communicative informativeness); see also Deborah Frisch, *Reasons for Framing Effects*, 54 ORGANIZATIONAL BEHAV. & HUM. DECISION PROCESSES 399, 421 (1993) (discussing reasons behind different responses to different versions of the framing problems, including that some framing effects arise "because changes in the way a situation is described affect subjects' responses even though the subjects themselves agree that the two versions are equivalent"); Hilton, *supra* note 150, at 255-57, 259-61 (discussing conversational processes in studies of the use of base rate information, studies of the dilution effect, and studies of the conjunction fallacy).

157. See Norbert Schwarz, *Communication in Standardized Research Situations: A Gricean Perspective*, in SOCIAL AND COGNITIVE APPROACHES TO INTERPERSONAL COMMUNICATION 39, 40 (Susan R. Fussell & Roger J. Kreuz eds., 1998). Schwarz notes:

Having no reason to assume that a researcher would ask a question about a nonexistent issue, would present a scale that is haphazardly designed, or would provide useless information, research participants are likely to search for meaning in a researcher's contributions. Little do they know that researchers

4. Use of Ecologically Suspect Formats

An alternative critical perspective on behavioral decision research methodology likewise sees overestimation of the frequency of cognitive error and bias in many behavioral decision experiments, but attributes such error inflation not to demand characteristics but to the decontextualized, abstract, or unnatural nature of the research setting and research task. This critical perspective predicts that representations of problems in ecologically valid terms will lead to improved performance on a range of heuristic and bias tests.¹⁵⁸ In

violate every norm of conversational conduct that they would be likely to observe outside the research setting by providing information that is not relevant, informative, truthful, or clear.

Id.; see also Schwarz, *supra* note 150, at 134 ("In fact, [when explicitly asked, participants usually seem] aware that the normatively irrelevant information is of little informational value. Nevertheless, they typically proceed to use it in making a judgment" because the sheer fact that it has been presented renders it conversationally relevant in the given context.) (citation omitted).

The reference in the text to "Linda the feminist bank teller" alludes to Kahneman and Tversky's famous "Linda problem" used to test for the conjunction fallacy. See Tversky & Kahneman, *supra* note 124. After reading a description of a "Linda," some subjects will judge it more likely that Linda is both a feminist and a bank teller rather than just a bank teller, ostensibly in violation of the conjunction rule that the frequency of an item in a subset cannot be more likely than the frequency of an item in the larger set. Hertwig and Gigerenzer discuss how the standard phrasing of the Linda problem and other word problems used by Kahneman and Tversky may lead to a number of responses that look like conjunction errors but may be in actuality logical responses in light of the language used. See Ralph Hertwig & Gerd Gigerenzer, *The 'Conjunction Fallacy' Revisited: How Intelligent Inferences Look Like Reasoning Errors*, 12 J. BEHAV. DECISION MAKING 275 (1999); see also Gerd Gigerenzer, *On Narrow Norms and Vague Heuristics: A Reply to Kahneman and Tversky* (1996), 103 PSYCHOL. REV. 592, 593 (1996). But see Kahneman & Tversky, *supra* note 78, at 585-87. Hertwig and Kahneman engaged in an "adversarial collaboration" designed to arbitrate this dispute at least in part, but the collaboration did not yield definitive results. See Mellers et al., *supra* note 86, at 273-75 (discussion of results by Hertwig and Kahneman); *id.* at 275 ("In our case, Hertwig acknowledges that the semantic ambiguity of 'and' is only one factor that contributes to conjunction effects in the frequency format. Kahneman acknowledges that the experiments have identified conditions under which the representativeness heuristic is overridden, and that there may be more to be discovered."). Of course, as a simple matter of probability, it must be more likely that Linda is a bank teller rather than both a feminist and a bank teller. The real question is whether there is a reasonable explanation for why the mathematically less probable answer is chosen: are people simply poor or lazy mathematicians, are they falling prey to a bias by-product of the representativeness heuristic, or are they being understandably duped into the mathematically wrong answer by conversational cues?

158. In explicating this view, evolutionary psychologists Brase, Cosmides, and Tooby note a paradox: Why have so many studies of human judgment and decision making over the last several years presented "human reasoning faculties [as] riddled with crippling defects,"

support of this view, we find that performance on several decision tasks does improve across subjects, often dramatically, when these tasks are presented in less abstract and more realistic terms or when the response options are given in formats that are more familiar to the subjects.

To first illustrate the ecological validity critique, consider performance on abstract versus contextualized versions of the Wason four-card selection task, which is used to test conditional reasoning performance.

The four-card selection task is one of the more widely researched deductive reasoning tasks, primarily because of the surprising nature of the results. The task is deceptively simple. [In abstract versions of the task, the subject] is given a rule, "If a card has a vowel on its letter side, then it has an even number on its number side," and is shown four cards face up: A, K, 18, 5. The subject has to decide which card(s) must be turned over to prove the truth or falsity of the rule. This task corresponds to the material-implication rule of the form, "If p then q ." The cards A and K represent the antecedents (p and $\neg p$, respectively), and the cards 18 and 5 represent the consequents (q and $\neg q$). The solution is to turn over A and 5 (p and $\neg q$), because the rule is violated only by cards pairing a vowel with an odd number. Most studies have found that only about 10% of the subjects can solve abstract forms of the problem.¹⁵⁹

whereas, during the same time period, studies of nonhuman animal behavior show that "animals with truly minuscule nervous systems, such as bumblebees, make judgments under uncertainty during foraging that manifest exactly the kind of well-calibrated statistical induction that the human brain was widely thought of as 'too limited' to perform"? Gary L. Brase et al., *Individuation, Counting, and Statistical Inference: The Role of Frequency and Whole-Object Representations in Judgment Under Uncertainty*, 127 J. OF EXPERIMENTAL PSYCHOL. 3, 4 (1998) (citations omitted). Surely it cannot be that bumblebees are "more rational" than humans. Brase, Cosmides, and Tooby believe the answer lies in the methodology: "Bumblebees appeared rational while humans did not because they were tested under ecologically valid conditions. When one does the same for human subjects, they too perform like good intuitive statisticians." *Id.* (citation omitted); see also Jou et al., *supra* note 103, at 10 ("Our conclusion is in agreement with other reasoning and judgment studies in which it has been concluded that people are more likely to make judgments and decisions in accord with the normative rules when a problem is presented in a familiar and realistic context.") (citations omitted).

159. Stephen J. Hoch & Judith E. Tschirgi, *Logical Knowledge and Cue Redundancy in Deductive Reasoning*, 13 MEMORY & COGNITION 453, 453 (1985) (citations omitted); see also Jonathan St. B.T. Evans & David E. Over, *Rationality in the Selection Task: Epistemic Utility*

Whereas performance on abstract versions of the selection task is usually quite poor across subjects, when the selection task is presented in a format in which subjects are asked to test the conditional relation between stimuli whose relations are more natural or familiar, a majority of subjects usually make the correct selections. For instance, the great majority of participants asked to assume the role of an authority figure monitoring for violation of a drinking rule select the proper cards to test the conditional relation in the following version of the selection task:

Rule: If a person is drinking beer, then the person must be over 18 years of age.

The cards below have information about four patrons in a restaurant. Each card represents one person. One side of the card tells you the age of the person, and the other side of the card tells you what that person is drinking in the bar.

Indicate only the card(s) you definitely need to turn over to see if any of these people are violating the rule. [Cards with the following information face-up are presented to subjects:]

Coke	Beer	16	22
------	------	----	----

Thus, most subjects properly indicate that the "Beer" and "16" cards should be turned over in this version of the task.¹⁶⁰ Many other

Versus Uncertainty Reduction, 103 PSYCHOL. REV. 356, 356 (1996) ("The Wason selection task has become the single most investigated problem in the psychology of reasoning. The psychological interest in it reflects the fact that although the problem is apparently simple to present and understand, the great majority of adults fail to solve it logically in its standard form.") (citation omitted).

The task is often called the "Wason" selection task after the first researcher to report its use, Peter Wason. See P. C. Wason, *Reasoning*, in NEW HORIZONS IN PSYCHOLOGY 135 (Brian M. Foss ed., 1966). Wason interpreted errors on his task as evidence of a "confirmation bias": "Wason described the hypothesis testers as 'seeking confirmation' because they looked predominantly at cases that fit their hypothesized rule for targets . . ." Joshua Klayman & Young-Won Ha, *Confirmation, Disconfirmation, and Information in Hypothesis Testing*, 94 PSYCHOL. REV. 211, 213 (1987). Klayman and Ha provide an excellent discussion of the broader positive test strategy that humans seem to employ (and that encompasses the "confirmation bias") and of why this may well be the most appropriate general strategy. See *id.* at 225 ("Our review suggests that people use the +test strategy as a general default heuristic. . . . Our theoretical analyses indicate that, as an all-purpose heuristic, +testing often serves the hypothesis tester well.").

160. See Richard A. Griggs & James R. Cox, *The Elusive Thematic-Materials Effect in Wason's Selection Task*, 73 BRIT. J. PSYCHOL. 407, 414-17 (1982) (describing Experiment 3);

variations on the selection task lead to large improvements in performance as well.¹⁶¹ One prominent explanation of this improvement is that conditional reasoning comes more naturally on deontic versions of the task (such as the drinking rule version above), where participants are testing for violations of social contracts, but exactly why these changes in the presentation of the selection task lead to such improvement is still under debate.¹⁶² However, the demonstrated improvement for thematic versions of the task rather than abstract versions of the task is not disputed.¹⁶³

id. at 415 ("Seventy-three per cent of the subjects made the correct selection for the drinking-age problem while no one did so for the abstract problem."). As compared to the abstract version of the test where cards representing *p* and *~q* should be selected to test the if-then relation, the "Beer" card represents the "*p*" card and the "16" card represents the "*~q*" card.

161. See Mike Oaksford & Nick Chater, *A Rational Analysis of the Selection Task as Optimal Data Selection*, 101 PSYCHOL. REV. 608, 612-24 (1994) (reviewing results across a wide variety of versions of the selection task).

162. Almor and Sloman, for instance, argue that deontic contexts are sufficient to improve conditional reasoning but not necessary. See, e.g., Amit Almor & Steven A. Sloman, *Is Deontic Reasoning Special?*, 103 PSYCHOL. REV. 374, 378-79 (1996) (showing improvement by altering the thematic context of the rule tested but using nondeontic domains and suggesting that it is beliefs about dependence relations rather than deontic relations that necessarily lead to improvement on the task). For an array of views on why performance improves on different versions of the task see Woo-kyoung Ahn & Loranel M. Graham, *The Impact of Necessity and Sufficiency in the Wason Four-Card Selection Task*, 10 PSYCHOL. SCI. 237 (1999); Almor & Sloman, *supra*; Amit Almor & Steven A. Sloman, *Reasoning Versus Text Processing in the Wason Selection Task: A Nondeontic Perspective on Perspective Effects*, 28 MEMORY & COGNITION 1060 (2000); Evans & Over, *supra* note 159; Jerry Fodor, *Why We Are so Good at Catching Cheaters*, 75 COGNITION 29 (2000); Donald Laming, *On the Analysis of Irrational Data Selection: A Critique of Oaksford and Chater (1994)*, 103 PSYCHOL. REV. 364 (1996); Mike Oaksford & Nick Chater, *Rational Explanation of the Selection Task*, 103 PSYCHOL. REV. 381 (1996); Alexander Staller et al., *Perspective Effects in Nondeontic Versions of the Wason Selection Task*, 28 MEMORY & COGNITION 396 (2000).

163. Gigerenzer and Hug argue that, regardless of which theory is chosen to explain improvement on the selection task, the evidence of such "content effects" on the task is strong evidence of the domain-specific nature of reasoning. See Gerd Gigerenzer & Klaus Hug, *Domain-Specific Reasoning: Social Contracts, Cheating, and Perspective Change*, 43 COGNITION 127, 128 n.1 (1992). Gigerenzer and Hug note:

The term "content effect" refers to a percentage of *P* & not-*Q* responses in a "thematic" rule that is substantially larger than the corresponding percentage in an "abstract", alphanumerical rule such as "If there is a *D* on one side of any card, then there is a 3 on its other side." The percentage of *P* & not-*Q* responses in "abstract" rules is typically below 20%. Both thematic and abstract rules are conditionals of the form "if *P* then *Q*."

Id. (citation omitted); *id.* at 166 ("It is the pragmatics of who reasons from what perspective to what end (e.g., cheating detection) that seems to be sufficient.").

The general point is that there is no simple and unique division line between structure and content, or between information relevant and irrelevant to

Furthermore, performance on tests of statistical reasoning and calibrations of confidence often improve significantly when decision problems are presented using frequency formats (e.g., fifty out of one hundred contract the disease) instead of presenting problems in terms of single-event probabilities (e.g., fifty percent probability of contracting the disease).¹⁶⁴ A dispute between Bayesians and

rational reasoning. What counts as the relevant structure for reasoning about a domain therefore seems to need a domain-specific theory. Thus we need to define what the relevant structural properties are—modal operators, prior probabilities, likelihoods, perspectives, cheating detection, and the like—rather than to leave this job to one out of many possible logics, usually selected by convention.

Id. at 168.

164. See, e.g., Gerd Gigerenzer & Ulrich Hoffrage, *How to Improve Bayesian Reasoning Without Instruction: Frequency Formats*, 102 PSYCHOL. REV. 684, 697-98 (1995). Gigerenzer and Hoffrage assert:

Mathematically equivalent representations of information can entail computationally different Bayesian algorithms. We have argued that information representation affects cognitive algorithms in the same way. We deduced four novel predictions concerning when information formats and menus make a difference and when they do not. Data from more than 2,800 individual problem solutions are consistent with the predictions. Frequency formats made many participants' inferences strictly conform (in terms of outcome and process) to Bayes' theorem without any teaching or instruction. These results were found for a number of inferential problems, including classic demonstrations of non-Bayesian inference such as the cab problem and the mammography problem.

Id. (citations omitted). Mellers and McGraw conclude:

[I]f one had to select a single form in which to present information, we would recommend frequencies. With rare events, frequencies facilitate Bayesian reasoning. If one had to select a single type of information, we would recommend the short task or, more generally, any type of information that allows people to visualize nested sets. These types and forms of information help people make better inferences, although neither one is a silver bullet. Bayesian reasoning is hard for both experts and nonexperts, and there is plenty of room for additional improvement.

Barbara A. Mellers & A. Peter McGraw, *How to Improve Bayesian Reasoning: Comment on Gigerenzer and Hoffrage* (1995), 106 PSYCHOL. REV. 417, 422 (1999). But Lewis and Keren state:

We suspect that the frequency formats advocated by Gigerenzer and Hoffrage (1995) and the corresponding algorithms they proposed comprise a program based on operant principles of learning in which underlying cognitive processes that will lead to a better and deeper understanding remain untouched. Like previous attempts to overcome cognitive biases and eliminate fallacious reasoning, the proposals of Gigerenzer and Hoffrage can be characterized by what Fischhoff termed "mechanical" manipulations. The effectiveness of such attempts may be more apparent than real.

Charles Lewis & Gideon Keren, *On the Difficulties Underlying Bayesian Reasoning: A*

frequentists about the proper normative approach to probability assessment has been ongoing for some time now.¹⁶⁵ One need not resolve this normative debate, however, to recognize that single-event versus frequency formats can lead to significantly different cognitive behavior.¹⁶⁶ In particular, the incidence of conjunction errors, expressions of overconfidence, violations of the law of large numbers, and base-rate neglect have been shown to decrease

Comment on Gigerenzer and Hoffrage, 106 PSYCHOL. REV. 411, 414 (1999) (citations omitted). *Contra* Gerd Gigerenzer & Ulrich Hoffrage, *Overcoming Difficulties in Bayesian Reasoning: A Reply to Lewis and Keren (1999) and Mellers and McGraw (1999)*, 106 PSYCHOL. REV. 425, 427 (1999) ("[C]ontrary to Lewis and Keren's hypothesis, joint statements alone cannot account for the beneficial effect of frequency formats.").

Slovic, Monahan, and MacGregor point out that frequency formats may not always yield better judgments than probability formats. They found that the use of a frequency format as opposed to a probability format for the ratings by psychologists and psychiatrists for likelihood of future violence and assessed risk of violence did not remove effects based on the amount of discriminability allowed by the scale (i.e., allowing more discriminability among smaller probabilities led to lower judged probabilities of future harmful acts, but use of a frequency format in the response scale did not eliminate this format effect). See Paul Slovic et al., *Violence Risk Assessment and Risk Communication: The Effects of Using Actual Cases, Providing Instruction, and Employing Probability Versus Frequency Formats*, 24 LAW & HUM. BEHAV. 271, 289 (2000). Also, the use of a frequency format response scale "consistently led to lower estimates of the likelihood of harm for a given case" but also to higher perceived risk. *Id.* at 289-90. They conclude that their data does seem to support the proposition that "thinking with probabilities differs from thinking with frequencies." *Id.* at 292. They also argue, however, that frequency formats are not necessarily superior to probability formats because they may both be susceptible to format effects. *Id.* Because they see advantages and disadvantages of both formats for the communication of risk estimates, they sensibly suggest that multiple formats should be used if possible.

If multiple formats were used in violence risk communication, the biases associated with any given risk communication format might, at least to some (unknown) extent, cancel each other out. In addition, the possibility of strategic behavior in choosing a risk communication format that promoted a favored policy outcome would be reduced if the risk communicator was instructed to use multiple formats, rather than to select a single one.

Id. at 294 (footnote omitted).

165. See Edwards & von Winterfeldt, *supra* note 9, at 228 ("Bayesian and frequentistic approaches lead to quite different conclusions about the rules for statistical inference, and statisticians of these persuasions have engaged in running arguments since the early 1960's.").

166. Note more generally that the structure and form of the problem can also greatly affect performance. See Vittorio Girotto & Michel Gonzalez, *Solving Probabilistic and Statistical Problems: A Matter of Information Structure and Question Form*, 78 COGNITION 247, 267-69 (2001) (finding that the problem representation evoked by the structure and form of the problem, rather than phrasing in probabilistic versus frequentistic terms, was most important to performance, with naive reasoners being able to draw correct inferences in both probabilistic and frequentistic formats).

significantly when people are given frequencies as input data or are asked to estimate frequencies of events as output data.¹⁶⁷ For example, Cosmides and Tooby report that frequentist versions of a medical diagnosis base-rate problem elicited high levels of Bayesian performance over numerous tests—"an average of 76% correct for purely verbal frequentist versions and 92% correct for a problem that requires subjects to construct a concrete, visual frequentist representation"—but nonfrequentist versions of the problem elicited very low levels of Bayesian performance, with only two percent of the subjects correctly answering some nonfrequentist versions of the problem.¹⁶⁸ Humans, and animals in general, seem much more adept at computing and utilizing frequency information than single-event probabilities, perhaps because of the advantages attached to frequency representations rather than proportions or single-event probabilities.¹⁶⁹

167. Leda Cosmides & John Tooby, *Are Humans Good Intuitive Statisticians After All? Rethinking Some Conclusions From the Literature on Judgment Under Uncertainty*, 58 COGNITION 1, 19-20 (1996) (reviewing research on frequentist representations of conjunction and overconfidence problems); *id.* at 22-37 (presenting evidence from frequentist representations of base-rate problems); Gerd Gigerenzer et al., *Probabilistic Mental Models: A Brunswikian Theory of Confidence*, 98 PSYCHOL. REV. 506, 526 (1991) ("We provided experimental evidence ... showing how changes in the task (confidence vs. frequency judgment) and in the relationship between task and environment (selected vs. representative sampling) can make the two stable effects reported in the literature—overconfidence and the hard-easy effect—emerge, disappear, and invert at will."); Hertwig & Gigerenzer, *supra* note 157, at 279-92 (presenting evidence from frequentist representations of conjunction problems); Koehler, *supra* note 67, at 7-8 (discussing frequentist representations of base-rate problems); Peter Sedlmeier & Gerd Gigerenzer, *Intuitions About Sample Size: The Empirical Law of Large Numbers*, 10 J. BEHAV. DECISION MAKING 33, 45 (1997) ("[T]he evidence showed that (1) frequency distribution problems that are directly comparable to sampling distribution problems elicit substantially higher percentages of participants who take sample size into account, with almost no overlap between the distributions of percentages; and (2) frequency distribution problems not directly comparable to sampling distribution problems result in participants' generally taking sample size into account."); *see also* Brase et al., *supra* note 156, at 19-20 (summarizing evidence of the importance of frequentist representations, particularly of whole objects, in Bayesian reasoning problems).

168. Cosmides & Tooby, *supra* note 167, at 58.

169. Brase et al., *supra* note 156, at 5-6 (discussing why adaptation may have favored frequency representations). When frequency information is converted into proportions or probabilities,

the absolute frequencies of the two component events cannot be recovered. As a result, (a) it is difficult to update the database as one encounters new instances; (b) the sample size is lost, and with it a basis for indexing how reliable one's estimate is ...; (c) more data needs to be stored (including likelihoods and base rates); and (d) the original data cannot be recategorized to construct novel

When Leda Cosmides and her colleagues summarize the evidence from the frequency format experiments, they echo behavioral decision theorists who emphasize the constructive nature of choice, but the conclusions drawn differ:

It turns out that human performance in probabilistic reasoning tasks is remarkably sensitive to the format in which information is presented and answers asked for. Most experiments that elicited "nonnormative" performance asked subjects to judge the probability of a single event (e.g., "What is the chance that *a person* who tests positive for the disease actually has it?"). However, many purported biases and fallacies disappear when people are asked to judge a frequency instead (e.g., "*How many people* who test positive for the disease will actually have it?"). ... With a frequency format, subjects not only used the base-rate information, but they used it fully, producing answers that conform to the strictures of Bayes's rule.

....

These results suggest that humans, like other animals, have inductive-reasoning mechanisms that embody certain rational principles, but that the design of these mechanisms requires representations of event frequencies to operate properly.¹⁷⁰

In sum, the use of unnatural and unfamiliar formats in behavioral decision-making experiments is another way in which the likelihood of finding deviations from rational actor norms is increased. When judgment and decision making are examined using more naturalistic or more familiar formats or, in the more specific Brunswikian sense, using formats where cues have more potential utility for the organism in its environment, performance often moves closer to predictions derived from rational actor norms.¹⁷¹

reference classes after the fact, as they become useful.

Id. at 6.

170. *Id.* at 4 (citations omitted).

171. Note that Evans and his colleagues recently questioned the reason for improved performance on frequency format problems, and they provide "a number of examples where frequency versions are no easier than the probability versions with which they are compared." Jonathan St. B.T. Evans et al., *Frequency Versus Probability Formats in Statistical Word Problems*, 77 COGNITION 197, 211 (2000). They conclude that

our findings cast serious doubts upon the widely cited claim that frequency formats facilitate correct statistical reasoning in quantitative word problems. Instead we have shown that the manner of participants' responses to such tasks

5. *Deprivation of Decision Tools*

There is another simple way in which behavioral decision researchers may increase the likelihood of finding non-normative responses in their experiments—deprive subjects of decision aids that the experimenters themselves found necessary to derive the normative answer to the problem at hand. For instance, while some researchers and subjects may be capable of performing accurate Bayesian calculations in their heads, researchers no doubt check their calculations using decision aids. The availability of instructional texts, calculators, or computers undoubtedly would make finding the answers to Bayesian problems easier for at least some subjects—perhaps even a great number.¹⁷²

As Edwards and von Winterfeldt appropriately point out, it is the availability of adequate time and decision tools, combined with accuracy motivation, that at least partially explain why the behavioral decision researchers themselves arrive at normatively correct answers:

What are the differences between the researchers and their respondents? Two suggest themselves. One is that the researchers need to, and do, make the mental effort required to get the right answers. The other is that, in doing so, the researchers have access to, and use, whatever intellectual tools

is strongly influenced by subtle variations in the presentation of task information which make it more or less easy to form a mental representation helpful to finding the normative solution.

Id. at 212. Thus, as with so much behavioral decision research, the evidence is ambiguous as to how and why subjects perform better in some experimental settings than others, and it may ultimately be the case with respect to the frequency versus probability format debate that it is some factor other than frequency formats that improves performance. See Girotto & Gonzalez, *supra* note 166, at 267-72 (discussing possible explanations); Dale Griffin & Roger Buehler, *Frequency, Probability, and Prediction: Easy Solutions to Cognitive Illusions?*, 38 COGNITIVE PSYCHOL. 48, 74 (1999) (noting that many factors, and not just frequency versus probability formats, affect performance and doubting that an “aggregate frequency approach to probabilistic reasoning” will turn out to be a “magic pill”).

172. Girotto and Gonzalez tellingly refer to “Bayes’ cumbersome rule.” Girotto & Gonzalez, *supra* note 166, at 271. For a primer on Bayes’ Theorem see ROBYN M. DAWES, RATIONAL CHOICE IN AN UNCERTAIN WORLD 323-26 (Jerome Kagen ed., 1988). Behavioral decision theorist Michael Birnbaum provides a Bayesian calculator on-line at <http://psych.fullerton.edu/mbirnbaum/bayes/BayesCalc.htm> (last visited Mar. 28, 2002).

they need—including books, calculators, computers, and whatever else the intellectual task requires.¹⁷³

Too often, behavioral decision theory experiments fail to report exactly what decision tools, if any, were made available and often fail to provide much detail about the procedures used during the experiment (no doubt because of the frequent use of quick question-and-answer surveys used in large classroom settings). Indeed, we are rarely told whether experimenters will even answer questions posed by the subjects or how many subjects refused to answer the questions out of frustration or due to confusion.¹⁷⁴ Thus, we are often left in the dark about what conditions, including what disadvantages, subjects may be operating under during the studies.

Time pressures and resource limitations may restrict the ability of the real-world decision maker to rely on decision aids, but to the extent that decision aids are regularly used in a real-world decision setting, then performance should be assessed in experiments with the use of these aids rather than simply based on intuitive judgment

173. Edwards & von Winterfeldt, *supra* note 9, at 268; see also E.C. POULTON, BEHAVIORAL DECISION THEORY: A NEW APPROACH 266-67 (1994) (discussing how insufficient time and complex mental arithmetic contributes to non-normative responses on certain behavioral decision theory tasks).

174. Cf. Beach et al., *supra* note 128, at 52 ("Interpretations of results for word problems often are weakened further by the lack of unanimity among the subjects about the answers, correct or otherwise. Many subjects seem mostly to be confused. Indeed, in one study, around 20 per cent of the subjects refused to give an answer at all, claiming that they simply did not know—subjects such as these usually are not included in the data analyses, so it is difficult to know how common such non-responding is.") (citations omitted); *id.* at 54 ("Many of [the word problems used in behavioral decision theory] are difficult for even sophisticated people to understand.").

or mental calculations.¹⁷⁵ Otherwise, we may obtain an inaccurate picture of real-world performance.¹⁷⁶

C. Features of Behavioral Decision Theory That Limit the Importance of the Research for the Legal System

Finally, certain features of behavioral decision theory limit its relevance and importance for the law. Thus, even if it were the case that experiments in behavioral decision theory found uniform behavior across all subjects regardless of the experimental situation, there remain several features of this research that would make extrapolation to real-world legal settings perilous. These features may be considered limits on the "external validity" of behavioral decision research, for they limit the extent to which this research generalizes to legal settings and the extent to which this research can establish a connection between irrationality in the laboratory and irrationality in the legal world.¹⁷⁷

175. The literature on decision aids indicates that people may be reluctant to use mechanical decision tools in the place of their own judgment. See, e.g., Steven E. Kaplan et al., *The Effects of Predictive Ability Information, Locus of Control, and Decision Maker Involvement on Decision Aid Reliance*, 14 J. BEHAV. DECISION MAKING 35, 47 (2001) ("[A] common theme emerging from recent decision aid research is that decision makers are reluctant to relinquish their judgments in favor of decision aids."); Peter Todd & Izak Benbasat, *Inducing Compensatory Information Processing through Decision Aids that Facilitate Effort Reduction: An Experimental Assessment*, 13 J. BEHAV. DECISION MAKING 91, 103 (2000) ("[D]ecision aids can induce the use of normatively oriented strategies. The key to inducing these strategies is to make the normative strategy easier to execute than competing alternative strategies."). Much less work has been performed with respect to the role of simple decision tools that merely assist with calculations or evaluations and do not interfere with the decision maker's cognitive control. Use of these simpler decision aids is probably quite common, and their role in improving performance is worthy of investigation.

176. See W. Edwards & B. Fasolo, *Decision Technology*, 52 ANN. REV. PSYCHOL. 581 (2001) (reviewing decision analysis methods and decision technology that can be used to improve decision making, including web-based technology). While cost-benefit analysis is susceptible to errors in input and in the interpretation of outputs, it may help the decision maker focus on relevant attributes of a decision problem and more systematically identify and weigh options. See Cass R. Sunstein, *Cognition and Cost-Benefit Analysis*, 29 J. LEGAL STUD. 1059, 1096 (2000) ("Cost-benefit analysis, taken as an inquiry into the consequences of varying approaches to regulation, is a sensible response not only to interest group power but also to limited information and to predictable problems in the public demand for regulation. . . . [A]n effort to identify costs and benefits can properly inform analysis.") (footnote omitted).

177. "External validity" refers to "the extent to which one can generalize the results of the research to the populations and settings of interest in the hypothesis." KIDDER & JUDD, *supra* note 136, at 27. For an overview of external validity issues in behavioral decision-making

1. *Real-World Success Versus Normative Coherence*

Implicit in much legal decision theory is the assumption that errors in the ways in which judgments or decisions are made will result in sub-optimal outcomes for the decision maker or others (i.e., deviations from norms of rational decision making will presumably result in real-world failure or dissatisfaction). For instance, it has been suggested that if contracting parties fail to negotiate fully around initially proposed contract terms because of cognitive imperfections that give these initial terms a difficult-to-defeat inertia, then the party preparing the first contract draft may "gain a powerful advantage in negotiations."¹⁷⁸ This claim could well be true with respect to the goal of prevailing in the negotiation, so that cognitive biases give the original drafter an advantage, but, of course, it does not necessarily follow that failing to negotiate fully around the originally proposed terms will leave either of the contracting parties better or worse off in the long run. The original terms, supposedly accepted with the assistance of cognitive imperfections, could well lead to a fair and efficient transaction. The operation of cognitive imperfections does not guarantee sub-optimal outcomes in any particular case.

A recent empirical study by legal decision theorists illustrates well the difficulties in drawing conclusions about real-world success or failure from performance in the laboratory. The legal decision theorists Chris Guthrie and Jeffrey Rachlinski, along with Federal Magistrate Judge Andrew Wistrich, conducted a study designed to assess whether judges fall prey to certain cognitive biases.¹⁷⁹ Guthrie and his colleagues administered a five-page questionnaire to approximately 200 magistrate judges attending a federal judicial

research not limited to the typical critique of the subject population, see Ebbesen & Konečni, *supra* note 137, at 21.

In the companion paper, I address several individual and situational differences in rational behavior that further limit the generality of behavioral decision theory and thus can be seen as additional external validity bounds. See Mitchell, *supra* note 8.

178. Korobkin, *supra* note 4, at 1627.

179. Guthrie et al., *supra* note 18, at 778 ("We report the results of an empirical study designed to determine whether five common cognitive illusions (anchoring, framing, hindsight bias, the representativeness heuristic, and egocentric biases) would influence the decision-making processes of a sample of 167 federal magistrate judges.").

conference.¹⁸⁰ The questionnaire presented five legal scenarios with questions about each scenario for the judges to answer.¹⁸¹ One hundred and sixty-seven judges anonymously completed the questionnaire in approximately ten minutes or less and allowed their responses to be used in the study without compensation, though some of these judges failed to respond to all problems in the questionnaire.¹⁸² The fact patterns and questions were designed to test, in a legal context, five common biases and errors from behavioral decision theory: bias due to the anchoring and adjustment heuristic, framing effects, the hindsight bias, bias due to the representativeness heuristic, and an egocentric bias.¹⁸³ The authors of the study report that the judges showed susceptibility to all five of the cognitive fallacies tested.¹⁸⁴

180. *Id.* at 784-87.

181. *Id.* at 786-87.

182. *Id.* at 787; *id.* at 791 n.62 (31.1% of the respondents did not fully answer the anchoring question); *id.* at 797 n.89 (2.4% did not answer the framing question); *id.* at 802 n.120 (1.2% did not answer the hindsight question); *id.* at 809 n.148 (4.8% did not answer the representativeness heuristic question); *id.* at 814 n.181 (7.2% did not answer the egocentric bias question). Given the ten-minute duration of the study, judicial participants presumably spent approximately two minutes considering each question and its answer.

183. *Id.* at 784.

184. *Id.* ("We found that each of these cognitive illusions influenced the decision-making processes of the judges in our study. Although the judges displayed less vulnerability to two of the five illusions than other experts and laypersons, the results show that under certain circumstances judges rely on heuristics that can lead to systematically erroneous judgments."); *id.* at 816 ("Our study demonstrates that judicial decision making, like the decision making of other experts and laypeople, is influenced by the cognitive illusions we tested.").

These broad conclusions, drawn from very limited data, exemplify the push within legal decision theory to reach generalized conclusions about behavior from very limited, aggregate data. Consider more closely the manner in which the authors reach their conclusion that judges in general are susceptible to the "cognitive illusions" that they tested in their ten-minute question-and-answer survey: (1) The authors assume the improper influence of an anchor by finding different average damage awards between the anchor and no anchor conditions even though the anchor is admittedly of some informational value and there is no objective standard against which to assess the proper size of the award. *Id.* at 791-93. (2) They infer individual-level framing effects from between-subjects data (i.e., there is no demonstration that any individuals actually switched choices between frames) despite the fact that the majority of subjects made the same choice across conditions (in both framing conditions a clear majority of respondents favored the "don't settle" option, with 67.5% of all respondents favoring the "don't settle" choice.). *Id.* (3) They infer individual-level changes in probability judgments made in hindsight following exposure to outcome information from differences in average probability ratings across conditions (i.e., there is no demonstration that any individual's probability judgment in the absence of outcome information differed

The authors then contend that cognitive biases in judges "can compromise the quality of justice that the courts deliver."¹⁸⁵ Thus, the assumption is made that the violations of formal principles of rationality supposedly observed in the judges will likely interfere

from that judgment made in the presence of outcome information). *Id.* at 802-03. (4) They treat the failure of some of the judges to obtain the proper result on a seemingly difficult-to-compute conditional-probability problem without any decision aid as evidence of susceptibility to the inverse fallacy and biased operation of the representativeness heuristic *even though 40.9% of the respondents chose the correct choice and 8.8% chose the next best answer.* *Id.* at 809-11. *But see id.* at 816 (Qualifying the conclusion somewhat: "Though still susceptible to framing and the representativeness heuristic, the judges appear less susceptible than other decision makers to these effects."). (5) They infer a "strong egocentric bias concerning the likelihood that [the judges'] would be overturned on appeal" even though the judges were not asked about future chances of being overturned and their inference was based on a question that the authors concede could have been accurately answered by a majority of the respondents given the potentially confusing or ambiguous information being sought by the question. *See id.* at 814-15. Judicial participants were given the following statement and question:

"United States magistrate judges are rarely overturned on appeal, but it does occur. If we were to rank all of the magistrate judges currently in this room according to the rate at which their decisions have been overturned during their careers, [what] would your rate be?"

Id. at 813. The judges were then asked to place themselves into a quartile representing their relative rate of reversal (e.g., the highest quartile was defined as meaning that "the rate at which you have been overturned is higher than that of 75% of the magistrate judges in this room"). *Id.* at 813-14 & n.180. In other words, respondents were forced to differentiate between themselves and others even though there may have been very little difference in the reversal rates between them (indeed, the sentence before the question tells the judges they are rarely reversed), and it was possible that "the 56.1% of the judges who placed themselves in the lowest quartile have never been overturned on appeal, which suggests that they really do belong in the lowest quartile." *Id.* at 814.

Further, Guthrie and his colleagues refer in one table (table 4) to the "size of the effect" observed, and they offer a comparison between the supposed size of the effects found in their study with the size of effects found in other studies. *Id.* at 816-18 tbl.4. It should be noted that the reported information is not what is typically considered effect size data within social scientific research and further that no systematic comparison of effect sizes, as might be found in a meta-analysis, is presented. *See* Rosenthal & DiMatteo, *supra* note 93, at 68-70 (describing the basic steps in a meta-analysis); *id.* at 70-73 (discussing formal effect size measures); *see also* APA MANUAL, *supra* note 69, at 18 (identifying the common statistical measures of effect size reported within psychological research). The comparisons should be viewed with great caution because Guthrie and his colleagues do not present effect size measures, report no systematic procedure for collecting relevant studies, often provide very little information about the comparison studies, and otherwise did not follow meta-analytic procedure. Although Guthrie and his colleagues should be commended for reporting some descriptive statistics and information beyond simply null hypothesis significance testing results, their presentation of results is potentially misleading and confusing, particularly to the statistically unsophisticated or to persons unfamiliar with behavioral decision theory.

185. Guthrie et al., *supra* note 18, at 821.

with the achievement of justice or whatever other goals the judges may pursue. No survey was actually reported, however, of what goals the judges were trying to fulfill in answering the questions in the survey administered by Guthrie and his colleagues; therefore, no assessment can be made of whether the conduct of the judges effectively achieved their desired goals. Indeed, if a judge's goal was simply to complete the questionnaire as quickly as possible, then randomly answering the questions may have been the most efficient route to achieving this goal. Random responses arguably would be instrumentally, or means-end, rational,¹⁸⁶ but the random responses probably would result in very little systematic adherence to formal principles of logic or to any other norm used to derive the researchers' "correct" answers to the problems. In short, there is no necessary connection here between the normative errors supposedly committed by the judges and the success or failure of judges' in achieving their goals.

What we see happening here is slippage between different conceptions of rationality, or, more accurately, between different standards for evaluating the success of behavior—violations of coherence criteria are being equated with violations of correspondence criteria.¹⁸⁷ Or, stated alternatively, laboratory

186. Elster provides a means-to-the-end definition of rationality: "Rational choice is instrumental: it is guided by the outcome of action. Actions are valued and chosen not for themselves, but as more or less efficient means to a further end." JON ELSTER, *NUTS AND BOLTS FOR THE SOCIAL SCIENCES* 22 (1989) (footnote omitted); see also ROBERT NOZICK, *THE NATURE OF RATIONALITY* 64 (1993) ("On this instrumental conception, rationality consists in the effective and efficient achievement of goals, ends, and desires.").

187. Todd and Gigerenzer explain the differences between the two sets of criteria:

One set of criteria that is often used to assess judgments and decisions is the laws of logic and probability theory. These are often called *coherence* criteria because they are primarily concerned with the internal logical coherence of judgments rather than with how well they help us to make useful decisions in the real world. Most experimental research programs aimed at demonstrating the rationality or (usually) irrationality of humans and animals have used abstract coherence criteria. For instance, many claims that there are systematic irrational fallacies in human reasoning are based entirely on a violation of some rule or other of logic or probability.

... The function of heuristics is not to be coherent. Rather, their function is to make reasonable, adaptive inferences about the real social and physical world given limited time and knowledge. Hence, we should evaluate the performance of heuristics by criteria that reflect this function. Measures that relate decision-making strategies to the external world rather than to internal consistency—measures such as accuracy, frugality, and speed—are called

"errors" are being inappropriately equated with real-world "mistakes."

An "error" is a judgment of a laboratory stimulus that deviates from a model of how that judgment should be made. When this model is normative and rational, the error represents an incorrect judgment. A "mistake," by contrast, is an incorrect judgment in the real world, such as a misjudgment of a real person, and so must be determined by different criteria. Detection of an error implies the existence of a mistake *only* when the process that produces the error also produces incorrect judgments in real life. Unfortunately, this cannot be determined by merely demonstrating the error itself, because the same judgment that is wrong in relation to a laboratory stimulus, taken literally, may be right in terms of a wider, more broadly defined social context, and reflect processes that lead to accurate judgments under ordinary circumstances.¹⁸⁹

Guthrie and his colleagues used coherence criteria to evaluate the judges' responses and, from a lack of coherence by some, adjudged the group of judges systematically biased. They then posited or assumed that these violations of the coherence criteria are likely to lead to violations of correspondence criteria. It is easy to understand why they have done this: With hypothetical cases, there is no way to test for real-world success, but with real legal cases, there is often no objective measure of truth or accuracy.¹⁸⁹ "In the absence of an

correspondence criteria.

Peter M. Todd & Gerd Gigerenzer, *Précis of Simple Heuristics That Make Us Smart*, 23 BEHAV. & BRAIN SCI. 727, 737 (2000) (citations omitted); *see also supra* note 58.

188. David C. Funder, *Errors and Mistakes: Evaluating the Accuracy of Social Judgment*, 101 PSYCHOL. BULL. 75, 76 (1987). Peter Suedfeld provides a similar analysis:

The word "bias" is used as though it were a descriptor, but ... any decision shown to be wholly or partly influenced by bias is then treated as *ipso facto* flawed. "Bias" becomes synonymous with "error," the analyst typically forgetting that being biased does not necessarily mean that one is wrong either morally or factually.

Suedfeld, *supra* note 34, at 436.

189. *See* Funder, *supra* note 188, at 76 ("When the study is of social judgment, ... the criterion problem becomes even stickier. The complexity of most social situations makes any degree of certainty and precision in the establishment of 'truth' difficult to come by, and the necessary assumptions are difficult to formulate, much less confirm."); Robert J. MacCoun, *Epistemological Dilemmas in the Assessment of Legal Decision Making*, 23 LAW & HUM. BEHAV. 723, 725 (1999) ("[I]n social psychology ... we often lack objective measures of the

objective criterion of validity, the normative theory of judgment under uncertainty has treated the coherence of belief as the touchstone of human rationality."¹⁹⁰ The term "cognitive illusion" (used as a synonym for cognitive bias) is thus unfortunate, as it implies some "reality" against which to judge the illusion when no such reality may exist.

Absent a measure of the objective truth or accuracy for the legal decision or judgment at hand, the best alternative is to assess whether the participants' responses are likely to have achieved the participants' own immediate goals, whatever those goals may be. Guthrie and his colleagues failed to make this alternative assessment, but without this information it is perilous to assume that a deviation from the researchers' view of the correct answer is anything other than an unspoken disagreement between the researcher and subject about the best way to answer the question:

[T]here is some sensible reason for most behaviors. Even when people appear to be making systematically biased judgments or irrational decisions, it is likely that they are trying to solve some problem or achieve some goal to the best of their abilities. The behavioral researcher is well advised to look carefully at his or her research participant's behavior, beliefs, and goals to discern "the method in the apparent madness."¹⁹¹

When this closer scrutiny of behavior is performed, the researcher may find a goal that the researcher did not expect to find. The researcher may disagree with this goal, but the subject's attempt to reach her putative goal rather than the researcher's goal cannot be

'true' state of the sociopolitical environment."); Tversky & Kahneman, *supra* note 124, at 313 ("[O]bjective measurements of probability are often unavailable, and most significant choices under risk require an intuitive evaluation of probability.").

190. Tversky & Kahneman, *supra* note 124, at 313.

191. Hastie, *supra* note 52, at 659 (citations omitted); *see also* Koehler, *supra* note 67, at 13 ("In short, normative claims about data obtained from laboratory studies cannot be made without understanding how individual subjects represent the tasks and what informational assumptions they make."); Jack C. Wright & Matthew Drinkwater, *Rationality vs. Accuracy of Social Judgment*, 15 SOC. COGNITION 245, 271 (1997) ("Evidence of irrationality is often interpreted as implying that ecologically significant errors occur. Yet irrationality need not indicate inaccuracy, either conceptually or empirically. Our results indicate that the degree to which people's probability judgments obey normative principles (Bayes' Theorem) may reveal relatively little about the accuracy of those judgments.") (citation omitted).

classified as instrumentally irrational unless the researcher takes a normative stance on what specific end-states should be desired by her subjects. Such an evaluative normative stance, however, if phrased at any useful level of specificity, will almost certainly be controversial (e.g., we may all agree that judges should seek justice, but that simply begs the question of what is justice).¹⁹²

Laboratory errors repeated in the real world may well lead to results that the decision maker and others would consider costly or undesirable, but there is no a priori reason to believe that this connection will always exist. Thus, a judge in the study by Guthrie, Rachlinski, and Wistrich who makes judgments and decisions in a very "irrational" manner in these researchers' view, may reach results that are much better received by the litigants than a judge who conducts herself "rationally." In other words, normative coherence or incoherence alone says nothing about quality of result. This disjunction parallels the disjunction between procedural and distributive justice: Just as a court's apparent adherence to fair procedures does not ensure a distributively just outcome, a court's use of rational decision principles does not ensure the best result.¹⁹³

2. Individual Decision Making Versus Group Decision Making

The unit of analysis in the great majority of behavioral decision research is the individual judge or individual decision maker.¹⁹⁴ Yet many important legal judgments and decisions are made by groups

192. See, e.g., Gregory Mitchell et al., *Judgments of Social Justice: Compromises Between Equality and Efficiency*, 65 J. PERS. & SOC. PSYCHOL. 629 (1993) (empirical study of differing lay conceptions of social justice).

193. See Tom R. Tyler & Gregory Mitchell, *Legitimacy and the Empowerment of Discretionary Legal Authority: The United States Supreme Court and Abortion Rights*, 43 DUKE L.J. 703, 790-92 (1994) (discussing the possible manipulation of procedural justice perceptions to increase acceptance of unfair outcomes).

194. Likewise, psycho-legal research into jury behavior has focused mainly on the behavior of individual mock jurors. See Monica L. McCoy et al., *The Effect of Jury Deliberations on Jurors' Reasoning Skills*, 23 LAW & HUM. BEHAV. 557, 558 (1999) ("Traditionally, research on jurors' cognitive reasoning has focused on the decision-making processes of individual jurors."). David R. Shaffer & Shannon R. Wheatman, *Does Personality Influence Reactions to Judicial Instructions? Some Preliminary Findings and Possible Implications*, 6 PSYCHOL. PUB. POLY & LAW 655, 657 (2000) ("In our opinion, perhaps the greatest limitation of mock-trial simulations is that the vast majority of them attempt to draw inferences from decisions rendered by nondeliberating mock jurors rather than deliberating mock juries.").

or after group deliberation (e.g., jury verdicts, appellate decisions, legislative and administrative enactments, board of director decisions). Legal decision theorists are thus confronted with two crucial questions: (1) Are groups more or less prone than individuals to commit systematic errors in their judgment and decision making? (2) Does collective consideration of an issue improve individual decision making? Despite the importance of these questions, only a small body of research directly addresses them, and this research yields no clear answers.¹⁹⁵ Accordingly, individual-level research findings cannot simply be assumed to apply to group-type decisions.¹⁹⁶

The complexity of the issues that arise in a group decision-making setting counsel against the simple assumption that the quality of group information-processing will be comparable to that of single individuals or that group discussions and deliberations will have no impact on individual behavior.¹⁹⁷ For instance, whereas studies of

195. See Kerr et al., *supra* note 45, at 713. Kerr and his colleagues note:

The central question of this paper has been, "Which is more likely to make a biased judgment, individuals or groups?" Our overview of the relatively small and diverse empirical literature suggested that there was no simple empirical answer to this question. Even when we restrict our attention to particular bias phenomena (e.g., framing effects, preference reversals), there was frequently little consistency in the direction (i.e., sign) and magnitude of observed relative bias, *RB*.

Id.; see also *id.* at 691 ("We have been able to locate fewer than 30 different empirical studies that directly examined both individual- and group-level biases."); *id.* at 692 (classifying and summarizing the relevant empirical studies); Rostain, *supra* note 12, at 985 ("[W]hether collective deliberation tends to negate [cognitive bias] effects, which have been primarily investigated in individuals, is still an open question").

196. An example of legal decision theorists simply assuming that individual-level research applies to group decisionmaking is found in the works claiming that juries fall prey to the hindsight bias. See *supra* note 45.

197. For an example of a group setting affecting individual choices see Tatsuya Kameda & James H. Davis, *The Function of the Reference Point in Individual and Group Risk Decision Making*, 46 ORGANIZATIONAL BEHAV. & HUM. DECISION PROCESSES 55, 71 (1990) (in framing effect study, "subjects in the group conditions consistently showed riskier personal preferences than subjects in the individual conditions"). For an example of group judgments differing from individual judgments see Janet A. Sniezek & Rebecca A. Henry, *Revision, Weighting, and Commitment in Consensus Group Judgment*, 45 ORGANIZATIONAL BEHAV. & HUM. DECISION PROCESSES 66, 66 (1990) (in a price estimation task, "group judgments [were] more accurate than individual [or revised] judgments"). For an example of the complex pattern of results that may occur in group versus individual decision-making settings see Linda Argote et al., *The Base-Rate Fallacy: Contrasting Processes and Outcomes of Group and Individual Judgment*, 46 ORGANIZATIONAL BEHAV. & HUM. DECISION PROCESSES 296, 308 (1990) ("Thus,

non-deliberating mock jurors have generally found that such jurors have difficulty disregarding inadmissible evidence heard during a simulated trial when later making decisions, studies of *deliberating* mock jurors have found that these jurors do seem to disregard inadmissible evidence in their decisions made after the deliberations.¹⁹⁸ Furthermore, some research suggests that different individual jurors process trial information in very different ways, with some using a simpler approach that seeks to force evidence into their preferred theory of the case and others processing the information in a more open-minded way to consider the relative strength of competing theories or stories.¹⁹⁹ It is unwise, therefore, to assume that any single theory of information-processing can be applied to all persons in the jury context.

In some cases, group deliberations and collective decision making moderate bias, in some cases they have no apparent net effect, and in some cases they amplify bias. Norbert Kerr, Robert MacCoun, and Geoffrey Kramer, after a review of the extant literature and with the assistance of computer modeling, identify several factors

it seems that the probability estimates of groups will be more biased than individuals for cases where the individuating information is informative . . . and less biased than individuals when the individuating information is uninformative.”)

198. Kamala London & Narina Nunez, *The Effect of Jury Deliberations on Jurors' Propensity to Disregard Inadmissible Evidence*, 85 J. APPLIED PSYCHOL. 932, 937 (2000). London and Nunez report:

Empirical investigations of the influence of inadmissible evidence have generally focused on individual nondeliberating jurors. Generally, such studies find that individuals are biased when exposed to inadmissible evidence. However, the results from Experiments 1 and 2, as well as results from other research, suggest that jury deliberations may somehow moderate the effect of biasing factors. . . . Taken together, the research supports the notion that deliberations may play an important role in jurors' propensity to disregard inadmissible evidence.

Id. (citations omitted); see also Davis et al., *supra* note 45 (discussing differences in verdict sizes between juror and jury decisions); Schkade et al., *supra* note 45 (same).

199. See Deanna Kuhn et al., *How Well Do Jurors Reason? Competence Dimensions of Individual Variation in a Juror Reasoning Task*, 5 PSYCHOL. SCI. 289, 293 (1994) (“[T]he results indicate significant individual variation in the manner in which people approach the juror task. This variation, moreover, is readily conceptualized along a competence dimension; considering alternatives and reflecting on evidence is clearly more competent than not doing so.”); see also McCoy et al., *supra* note 194, at 571 (“Overall, the deliberation process does appear to encourage jurors to make more statements that discount theories (their own and alternative theories) and make more judgmental supportive statements. According to Kuhn et al.'s model, these are indicators of thinking in terms of a theory-evidence coordination model.”) (citation omitted).

that may interact to determine whether groups exhibit more or less bias relative to that shown by individuals, including: (1) the size of the group, with increasing group size tending to affect the magnitude of the bias but not whether bias is likely to appear; (2) the magnitude of individual biases brought to the group; (3) where in the decision process individual bias is likely to appear; (4) the type of bias at issue; and (5) the nature of the group process involved.²⁰⁰ To arrive at even this complex but manageable set of factors, Kerr and his colleagues had to assume that the group process was constant across group and task contexts. When this simplifying assumption is relaxed, even more possible interactions arise.²⁰¹ In short, the question of whether group judgments and decisions tend to be more or less biased than individual judgments and decisions is enormously complex and defies a simple answer.²⁰² Accordingly, legal decision theorists should refrain from simple conclusions about the similarity of group and individual decision-making processes and outcomes.

3. *Framing Research Confounds and Confusions*

We previously saw that framing effects are not as prevalent nor as robust as legal decision theorists suggest.²⁰³ The legal decision theorists' portrayal of framing effect research is misleading in another respect: whereas this research is typically summarized as revealing that people are risk averse on positively framed problems and risk seeking on negatively framed problems,²⁰⁴ there really is no single "framing effect." Within the framing effect research, we find confounds in the methodology of the studies that make the results

200. Kerr et al., *supra* note 45, at 713.

201. *Id.* ("Also note that all of these complex (but tractable) patterns assume that in any given group and task context, group process . . . is constant. When we relax this assumption, many other patterns are possible.")

202. Before leaping to conclusions in this area, one would do well to read the article by Kerr et al., *supra* note 45, to gain an appreciation of the great complexity of the relevant issues in this area. See also Norbert L. Kerr et al., "When Are *N* Heads Better (or Worse) Than One?": *Biased Judgment in Individuals versus Groups*, in 1 UNDERSTANDING GROUP BEHAVIOR: CONSENSUAL ACTION BY SMALL GROUPS 105-36 (Erich H. Witte & James H. Davis eds., 1996); Kerr et al., *supra* note 45, at 70-86 (illustrating the context-dependent nature of the answer to the question of whether individuals or groups appear more biased).

203. See *supra* section II.A.2.a.

204. See *supra* note 99.

difficult to interpret, and we see that the particular ways in which decisions are "framed" and the substance of these decisions may make important differences in the results observed.²⁰⁵ Therefore, the framing effect research resists simplification for policy purposes.

First, as noted above, many people fail to show the loss aversion effect in standard framing effect studies.²⁰⁶ Second, the framing research manipulates frames in a variety of ways that lead to different effects. Levin and his colleagues provide one typology for classifying the different types of framing manipulations. They show that framing manipulations may be distinguished based on:

(1) what is framed, (2) what is affected, and (3) how the framing effect is measured. All of these distinctions are likely to influence framing effects. In risky choice framing, the complete set of options differing in risk level is framed either positively or negatively, and the effect on risk preference is assessed by comparing the frequency of choice of the risky option in each framing condition. In attribute framing, a single attribute of an object or event is framed (labeled) positively or negatively, and the effect on item evaluation is assessed by comparing attractiveness ratings for the object or event in each framing condition. In goal framing, the consequences of a particular behavior are specified in either positive or negative terms, and the impact of alternative framing in persuading the decision maker to engage (or not engage) in that behavior is assessed by comparing the rate of adoption of the behavior in the two framing conditions.

Within each framing type, results show substantial consistency. In risky choice framing, a choice shift (but not necessarily a reversal) typically occurs such that positive frames

205. See Kühberger, *supra* note 101, at 46. Kühberger notes:

Our analysis shows that there is no uniform framing effect. The effect sizes in the data-base are very heterogeneous. . . . In evaluating framing research, one has to bear in mind that the framing effect can interact, may be dampened, or even may be made [to] disappear, by some characteristics other than those studied here. These may be individual difference dimensions relating to age, gender, culture, or language. These may also be more general cognitive dimensions like motivation, emotion, knowledge, perception, memory, and cognition.

Id.; see also Mandel, *supra* note 103, at 61-62 (discussing a confound between descriptor and expected outcome formulations in many gain/loss framing problems).

206. See *supra* notes 101-04 and accompanying text.

generally enhance risk averse responding relative to negative frames. In attribute framing, attributes are judged more favorably when labeled in positive terms rather than negative terms. And in goal framing, a negatively framed message emphasizing losses tends to have a greater impact on a given behavior than a comparable positively framed message emphasizing gains.²⁰⁷

Fagley and Kühberger alternatively classify the framing manipulation studies by distinguishing between “framing effect” and “reflection effect” studies. In framing effect studies, the final outcomes are identical but are simply phrased in positive or negative terms. For instance, in Kahneman and Tversky’s influential Asian disease problem, 400 people always die under the program to be evaluated—in the positive frame, the possibility of saving 200 of 600 people is emphasized, whereas in the negative frame, the possibility of losing 400 of 600 people is emphasized.²⁰⁸ In reflection effect studies, the outcomes in the gains condition are different from the outcomes in the losses condition, but the size of movement relative to the status quo stays constant. Hence, in reflection effect studies, the focus is on the relative utility or disutility of equal gains and equal losses (i.e., the research examines how changing the sign attached to the amount of money or other quantity at stake affects behavior).²⁰⁹

As these typologies illustrate, there are a variety of “framing effects” that may occur. Given these differences, “it should become apparent that the ... different types of framing effects cannot and should not be treated the same. Not recognizing the distinctions leads to unwarranted comparisons that may create unnecessary confusion.”²¹⁰ The underlying cognitive and motivational processes

207. Irwin P. Levin et al., *All Frames Are Not Created Equal: A Typology and Critical Analysis of Framing Effects*, 76 ORGANIZATIONAL BEHAV. & HUM. DECISION PROCESSES 149, 181 (1998).

208. See Zickar & Highhouse, *supra* note 82, at 75-76 (“Most framing items are a variation on the classic ‘Asian disease problem’ of Tversky and Kahneman ...”).

209. See N. S. Fagley, *A Note Concerning Reflection Effects Versus Framing Effects*, 113 PSYCHOL. BULL. 451, 451-52 (1993); Kühberger et al., *supra* note 101, at 206-07 (explaining the contrast between framing and reflection effect studies).

210. Levin et al., *supra* note 207, at 178; see also Parthasarathy Krishnamurthy et al., *Attribute Framing and Goal Framing Effects in Health Decisions*, 85 ORGANIZATIONAL BEHAV. & HUM. DECISION PROCESSES 382, 395 (2001) (reporting research “confirming that the effects

may differ across types of frame manipulations and comparisons of results across the different types of frames may lead to confusion.²¹¹ Thus, we need to use care when specifying how a law or legal setting is likely to “frame” choices for legal actors, and we need to use care in explaining why a certain framing effect supposedly occurred or is predicted.

Third, it is often difficult to isolate the cognitive or motivational processes underlying framing effects (as opposed to reflection effects) because of confounded factors in the framing effect studies. In these studies, “objective (final) and subjective (change) payoff conditions, levels of payoffs, and levels of probabilities are confounded.”²¹² It can be difficult to determine whether the level of payoff, the probabilities associated with choice options, or the objective/subjective payoff drives behavior. One recent meta-analysis found that greater efforts need to be made to disentangle the respective influence of these factors:

The general message of the present meta-analysis is that risk preferences in [framing effect] tasks depend on framing, size and type of payoffs, and the probabilities used. The task structure implies that probability levels, payoff magnitudes, and framing conditions cannot be independently varied. Disentangling these factors shows a relatively complicated picture, where different variables are important.

... [C]ontrolling for possible variables which are relevant to formal modeling (payoff, probability) does not make the framing condition superfluous as a predictor. Quite to the contrary,

of positive vs negative framing differ for *attribute framing* and *goal framing*”).

211. See, e.g., Frisch, *supra* note 156, at 423 (“The results of the present study suggest that such endeavors [examining how framing affects experience utility] should be based on a careful analysis of the different types of framing effects. It is likely that in some cases, although framing affects decisions, it does not have a corresponding effect on experience. However, it is quite plausible that other instances of framing effects arise because changes in frame have an impact on one’s experience of the consequences of a decision”); Kühberger et al., *supra* note 101, at 238-39 (“We consider our research to be a first step in distinguishing two effects which are still empirically and theoretically confounded: separating effects resulting from presenting gambles with differing signs of outcomes as stimuli (reflection effect) and effects resulting from phrasing information in a way that induces particular frames (framing effect).”).

212. Kühberger et al., *supra* note 101, at 208. The confounds are present only in the framing effect studies because “to make actual gains look like losses, one has to provide an initial endowment to lose from.” *Id.*

framing remains the most important predictor. Thus there is more to the framing effect than can be captured by the formal properties of tasks. This should be taken as an invitation to do more research on framing based on theories which incorporate factors such as cognitive and motivational processes. ... [C]haracterization of the framing phenomenon as a "cognitive illusion" is misleading.²¹³

Finally, the specific content or subject matter of the framing problems—as opposed simply to changes in the frame of reference—may exert significant effects on behavior. When human lives are at stake rather than simply money—particularly when those lives are within one's own family or one's "in-group" we see risk-seeking behavior even in the domain of gains.²¹⁴ The size of the target group may also matter.²¹⁵ Whose money is at stake may matter; for instance, financial planners have been found to be more careful (i.e., more risk averse in general) with other people's money than with their own money.²¹⁶ Whether one views the choice as a

213. *Id.* at 223 (citations omitted).

214. See N. S. Fagley & Paul M. Miller, *Framing Effects and Arenas of Choice: Your Money or Your Life?*, 71 ORGANIZATIONAL BEHAV. & HUM. DECISION PROCESSES 335, 367 (1997) ("[P]eople in the positive framing condition made more risky choices in the human life arena than in the monetary arena."); X. T. Wang et al., *Social Cues and Verbal Framing in Risky Choice*, 14 J. BEHAV. DECISION MAKING 1, 12 (2001) ("The framing effects found in Experiment 1 have shown three distinguishable risk-preference patterns. The risk preference of the participants was (1) clearly risk seeking under both framing conditions in a kinship context, (2) not significantly different from risk-neutral point under both framing conditions in a stranger-group context, and (3) risk averse under positive framing and risk seeking under negative framing in the heterogeneous group contexts where more than half of the endangered lives were strangers"); see also X. T. Wang, *Framing Effects: Dynamics and Task Domains*, 68 ORGANIZATIONAL BEHAV. & HUM. DECISION PROCESSES 145, 153 (1996) ("Subjects were significantly more risk seeking when facing life-death choice problems than facing their counterpart problems presented in a public property or personal money domain ... , and a risk-seeking augmenting unidirectional framing effect occurred in the family context ...").

215. X. T. Wang has found that both American and Chinese subjects show no framing effects when target groups are relatively small (of size six or sixty) but do show framing effects when target group size increases substantially, with Americans showing framing effects when the target group size increases to 600 (the target group size used in standard versions of the Asian disease problem) and with Chinese showing framing effects only when the target group size increased to 6000. See X. T. Wang, *Domain-Specific Rationality in Human Choices: Violations of Utility Axioms and Social Contexts*, 60 COGNITION 31, 35-42 (1996).

216. See Michael J. Roszkowski & Glenn E. Snelbecker, *Effects of "Framing" on Measures of Risk Tolerance: Financial Planners Are Not Immune*, 19 J. BEHAV. ECON. 237, 244-45 (1990) (finding little evidence of experience effects related to risk-taking patterns—with greater risk aversion in the domain of gains and greater risk seeking in the domain of losses—but finding

purchase of insurance against loss and, if so, the probabilities and consequences attached to the possible loss, matter.²¹⁷ Asking subjects to provide a rationale for their choices can eliminate framing effects, and attaching a probability of success greater than two-thirds to the risky option may eliminate risk aversion.²¹⁸ Changes in the "surface structure" of a framing problem can have important influences on the "deep structure" (or psychological representation and importance) of the problem and thus have important influences on the preferences expressed in response to the problem.²¹⁹ These examples merely illustrate the great content- and context-sensitivity of framing effects.²²⁰

In short, although the framing of choices in terms of losses or gains may impact the behavior of some people under some conditions, behavior in such decision settings involves much more than simply whether an option is phrased positively or negatively

that financial planners were more cautious with clients' money than their own).

217. See Paul J. H. Schoemaker & Howard C. Kunreuther, *An Experimental Study of Insurance Decisions*, 46 J. RISK & INSURANCE 603, 616 (1979) ("The influence of context was strong in the insurance question where the low deductible policy was perceived as more attractive (relative to other policies) than when the same choice was presented in a purely mathematical fashion without reference to insurance and deductibles. Format and context effects seem symptomatic of people's tendency to resort to simplifying strategies or to derive cues from the larger context of the problem."); Paul Slovic et al., *Preference for Insuring Against Probable Small Losses: Insurance Implications*, 44 J. RISK & INSURANCE 237, 253 (1977) ("The most striking result of the experiments just described is that people buy more insurance against moderate- or high-probability, low-loss events than against low-probability, high-loss events.").

218. See Paul M. Miller & N.S. Fagley, *The Effects of Framing, Problem Variations, and Providing Rationale on Choice*, 17 PERSONALITY & SOC. PSYCHOL. BULL. 517, 521 (1991) ("Regarding the conditions under which framing effects are observed, it appears that framing effects may be limited to situations where no rationale is requested and where the probability of success in the risky option is less than 2/3.").

219. See Willem A. Wagenaar et al., *Islanders and Hostages: Deep and Surface Structures of Decision Problems*, in RESEARCH ON JUDGMENT AND DECISION MAKING: CURRENTS, CONNECTIONS AND CONTROVERSIES, *supra* note 20, at 552-65 (finding that variations in the presentation and context of problems and variations within confounding variables, though changes in the cover story for framing problem, led to significant differences in risk preferences); see also Herbert Bless et al., *Framing the Framing Effect: The Impact of Context Cues on Solutions to the 'Asian Disease' Problem*, 28 EUR. J. SOC. PSYCHOL. 287, 290 (1998) ("[W]e observed a reflection effect when the context cue suggested a meaningful psychological situation, for example a medical decision problem. . . . This reflection effect was eliminated, however, when participants were provided with the context cue 'statistical research'.").

220. For additional qualifying examples see the studies summarized in tables 2, 3, and 4 of the paper by Levin and his colleagues. See Levin et al., *supra* note 207, at 154-57 tbl.2, 161-63 tbl.3, 169-71 tbl.4.

in relation to some psychological reference point. From the perspective of rational choice theory, formally irrelevant changes in the presentation of risky choices should not affect behavior so long as the expected utility stays constant, but we know that gain-loss framing changes may be psychologically relevant. From the perspective of behavioral decision theory, we thus should not be surprised that additional changes in the presentation of framing problems may be psychologically relevant.²²¹ Yet the legal decision theorists seem to treat all framing problems and resultant framing effects as uniform, indicating that they believe that the psychological reference point and manipulations of losses/gains around that reference point are the only psychologically relevant factors. The precise words used in a law or a contract and the subject matter of the law or contract, not to mention the characteristics of the legal actors and characteristics of the larger social context, may dramatically affect whether and to what extent different behavioral tendencies (most typically in the form of different risk preferences) are observed across positive and negative framings of formally equivalent judgment and choice options. In short, there is no universal framing effect that can be easily translated into legal doctrine.

4. Neglect of Systematic Information-Processing Modes

As discussed previously, legal decision theory is founded on a predominantly unimodal view of information-processing that emphasizes heuristic processing and transient preference construction and excludes a deliberative mode of thought that functions by careful application of normative rules.²²² Given this unimodal view of thought, it is easy to understand why legal decision theorists assume systematic cognitive bias and errors in all persons across all situations: legal decision theorists assume the universality of a standard set of imperfect mental mechanisms used in all judgment and decision making.

221. See, e.g., Kühberger et al., *supra* note 101, at 238 ("Seemingly inconsequential variations in problem wordings were shown to produce a framing effect, no framing effect, or a reversal of the framing effect.").

222. See *supra* notes 23-31 and accompanying text.

Whether legal decision theorists intend to reject the possibility or the importance of alternative, nonheuristic means of information-processing is unclear, because they do not clearly address this possibility. The unimodal view of thought subscribed to by the legal decision theorists seems to be inaccurate or at least incomplete, however, because it appears that individuals are capable of at least two fundamentally different forms of information-processing. In fact, Kahneman and Tversky—whose work serves as the basis for much of legal decision theory and legal decision theory's emphasis on cognitive heuristics—posited *two* forms of reasoning: a natural assessment, intuitive mode (that relies on cognitive heuristics) and an extensional, logical mode (that proceeds more algorithmically), with reasoning occurring most commonly but not only by means of natural assessments.²²³ The former mode refers to the heuristic form of information-processing presumed to lead to systematic biases and errors, whereas the latter mode refers to a more formalistic and deliberative form of information-processing less susceptible to the same kinds of cognitive biases and errors.

223. See Kahneman, *supra* note 77, at 682 ("Tversky and I always thought of the heuristics and biases approach as a two-process theory. We proposed that intuitive judgments of probability or frequency are based on 'natural assessments'—that is, of similarity, causality, affective valence, or past frequency—which are effortless and largely automatic. In the terminology of the target article, judgment by heuristics is a manifestation of System 1. However, we also believed that System 2 can override intuitive judgments that conflict with a general rule—but only if the relevant rule is evoked."); Tversky & Kahneman, *supra* note 124, at 308-09, 310 (discussing cues that may elicit extensional reasoning instead of natural assessment-based reasoning); cf. *id.* at 313-14 ("Our studies of the conjunction rule show that normatively inspired theories that assume coherence are descriptively inadequate, whereas psychological analyses that ignore the appeal of normative rules are, at best, incomplete. A comprehensive account of human judgment must reflect the tension between compelling logical rules and seductive nonextensional intuitions."). Tversky and Kahneman describe the natural assessment approach as follows:

Our conception of judgmental heuristics is based on *natural assessments* that are routinely carried out as part of the perception of events and the comprehension of messages. Such natural assessments include computations of similarity and representativeness, attributions of causality, and evaluations of the availability of associations and exemplars. . . .

The term *judgmental heuristic* refers to a strategy—whether deliberate or not—that relies on a natural assessment to produce an estimate or a prediction. One of the manifestations of a heuristic is the relative neglect of other considerations. . . . Hence, the use of judgmental heuristics gives rise to predictable biases.

Id. at 294.

Several psychologists have focused considerable attention on the issue of the extent to which a single individual may engage in different types of information-processing at different times (i.e., intra-individual variation in reasoning processes) and how exactly the different systems or modes of thought may interact and be consciously controlled.²²⁴ The work of these psychologists suggests that, instead of always relying on the same cognitive mechanisms to process information, people sometimes engage in thought processes more closely approximating normatively rational decision strategies (which may lead to random errors when people have insufficient processing resources), whereas at other times people engage in thought processes utilizing “arational” heuristics (which typically lead to “good” choices but may be more prone to nonrandom errors).²²⁵ Although the specific “dual-process” theories of cognition

224. See Eliot R. Smith & Jamie DeCoster, *Dual-Process Models in Social and Cognitive Psychology: Conceptual Integration and Links to Underlying Memory Systems*, 4 PERSONALITY & SOC. PSYCHOL. REV. 108, 108 (2000) (“In recent years, researchers working in numerous areas of social and cognitive psychology have developed models that follow these general lines: *dual-process models*, as we label them here. Such models contain three major components. They provide accounts of how people process in quick-and-dirty fashion, how they process when willing and able to engage in extensive thought, and what conditions encourage such effortful processing.”) (citation omitted). For introductions to dual-process models in social and cognitive psychology see Gordon B. Moskowitz et al., *The History of Dual-Process Notions, and the Future of Preconscious Control*, in DUAL PROCESS THEORIES IN SOCIAL PSYCHOLOGY 12 (Shelly Chaiken & Yaacov Trope eds., 1999); Steven A. Sloman, *The Empirical Case for Two Systems of Reasoning*, 119 PSYCHOL. BULL. 3 (1996). See also Keith E. Stanovich & Richard F. West, *Individual differences in reasoning: Implications for the rationality debate?*, 23 BEHAV. & BRAIN SCI. 645, 658-59 (2000) (summarizing the features common to the two modes of information-processing in several dual-process models).

225. Sloman describes the two modes of thought in terms of the two views of cognitive error mentioned earlier—error due to bounded rationality and error due to use of cognitive heuristics.

Each of these two views of human error suggests a different way to conceive of the place of normative rules in human thought. Bounded rationality suggests that people are able to follow rules derived from normative theory, although their ability to do so is limited by their limited cognitive resources. The natural assessment approach suggests that thought involves a process in which no rules of any kind are followed. Evidence can be found for both of these suggestions. Human thought seems to have two complementary aspects. These aspects have been distinguished in a variety of ways by a number of theorists over the years. To take only one example, Reason distinguishes an attentional control mode from an automatic one, which he takes to be the locus of much human error.

Sloman, *supra* note 23, at 576 (citation omitted). As Epstein and Pacini note, a long-standing debate in cognitive psychology concerns whether people process information primarily by association or by use of rules, and a dual-process approach such as that taken by Sloman

differ in important respects on their particulars,²²⁶ they all view information processing as capable of occurring through two different modes, with one mode of thought being more systematic, more effortful, and more rule-governed, and the other being more automatic, less effortful, and more schema-driven.²²⁷

Dual-process models allow both reasoned and heuristic forms of information-processing to operate in parallel—indeed, the two modes may interact with one another, at times working well together and at other times leading to internal conflict.²²⁸ These models likewise allow judgment and decision-making errors to arise from different sources and suggest a somewhat more flattering view of human judgment and decision making than has been dominant recently.²²⁹

offers a compromise view:

Cognitive psychologists are divided on whether people are parallel information processors who operate by associative connections, or whether they operate according to a more deliberative, rule-based, sequential mode of information processing. Noting that there is impressive evidence on both sides of the issue, Sloman has suggested a resolution based on the assumption that people operate both ways.

Seymour Epstein & Rosemary Pacini, *Some Basic Issues Regarding Dual-Process Theories from the Perspective of Cognitive-Experiential Self-Theory*, in DUAL-PROCESS THEORIES IN SOCIAL PSYCHOLOGY, *supra* note 224, at 465 (citation omitted).

226. For a sampling of the diverse dual-process models see DUAL-PROCESS MODELS IN SOCIAL PSYCHOLOGY, *supra* note 224.

227. See Seymour Epstein et al., *Individual Differences in Intuitive-Experiential and Analytical-Rational Thinking Styles*, 71 J. PERSONALITY & SOC. PSYCHOL. 390, 390 (1996) ("Psychologists from various persuasions have proposed two fundamentally different modes of processing information: one that has been variously referred to as intuitive, natural, automatic, heuristic, schematic, prototypical, narrative, implicit, imagistic-nonverbal, experiential, mythos, and first-signal system and the other as thinking-conceptual-logical, analytical-rational, deliberative-effortful-intentional-systematic, explicit, extensional, verbal, logos, and second-signal system.") (citations and emphasis omitted). To Epstein's listing of descriptors for the two modes of thought, we may add Sloman's distinction between associative and rule-based systems. See Sloman, *supra* note 224, at 4-6.

228. See, e.g., JOHN ST. B.T. EVANS & DAVID E. OVER, RATIONALITY AND REASONING 146 (1996) ("The dual process theory that we now favour is neither a sequential model nor a conflict model, but rather an interactive model. . . . The interactive nature of the two processes lies in the fact that our conscious thinking is always shaped, directed, and limited by tacit, pre-attentive processes.").

229. Consider, for example, the comments of the social psychologist Philip Tetlock: Experimental research on judgment and choice casts us, human beings, in a less-than-flattering light. We fall prey, it has been claimed, to a wide assortment of errors and biases. We are too quick to draw conclusions about others, too slow to change our minds, excessively confident in our predictions, and prone to give

Unfortunately, given the popularity of studying heuristics and biases, behavioral decision theorists have devoted relatively little effort to studying the systematic mode of processing and how it interacts with the heuristic mode. Rather, there has been an emphasis on the study of the heuristic mode of information-processing that may be more common in laboratory settings—in which quick, intuitive judgments or choices, often on unfamiliar, difficult, or abstract tasks, are requested, typically in the absence of feedback or decision aids and without serious consequences for the actor.²³⁰ The result has been the derogation of both the study of a more systematic mode of thought that all competent adults are capable of to some extent and the role of motives and affect in reasoning and choice.²³¹

The most immediate impact of this evidence of dual modes of thought, and the relative neglect of study of the more systematic mode of thought, is to further bring into question the validity of

too much weight to irrelevant cues (such as sunk costs) and too little weight to relevant ones (such as opportunity costs). Although this grim portrait has been qualified by the recent proliferation of dual-process models of judgment and choice that, in the spirit of Simon, bestow on people some limited capacity to decide how to decide . . . , the dominant emphasis in the last quarter century of experimental work has clearly been on judgmental shortcomings.

Philip E. Tetlock, *Cognitive Biases and Organizational Correctives: Do Both Disease and Cure Depend on the Politics of the Beholder?*, 45 ADMIN. SCI. Q. 293, 293 (2000) (citations omitted).

230. One example of cognitive bias that may be ameliorated by greater cognitive processing is the overconfidence phenomenon, in which subjects express unwarranted confidence in the accuracy of their answers to general knowledge or other types of questions. See, e.g., Janet A. Sniezek et al., *The Effect of Choosing on Confidence in Choice*, 46 ORGANIZATIONAL BEHAV. & HUM. DECISION PROCESSES 264, 279 (1990) (discussing data showing that more serious cognitive effort on decision tasks resulted in better calibration of accuracy judgments). Motivation to get the “right answer” may alone not be sufficient to overcome cognitive bias on mentally taxing tasks, but increased accuracy motivation is likely to improve accuracy on less-taxing tasks. See Brett W. Pelham & Efrat Neter, *The Effect of Motivation of Judgment Depends on the Difficulty of the Judgment*, 68 J. PERSONALITY & SOC. PSYCHOL. 581, 590 (1995) (“[T]he research presented in this article indicates that the motivation to judge accurately can either facilitate or debilitate judgment. In particular, when people’s cognitive resources are heavily taxed, their attempts to make especially accurate judgments may backfire. . . . The results of these studies add to a growing list of findings in support of contextual models of human judgment.”).

231. See, e.g., Kahneman, *supra* note 60, at 145 (discussing shortcomings of much heuristics and biases research and calling for greater investigation into the robustness of prior findings and the effects of many important features that may be operative in real-world settings, such as the effects in sequential rather than isolated choices, the effects of the social context of choice, and the effects of motivational and emotional factors).

legal decision theory's assumption that all people systematically rely on a heuristic and biased mode of thought when making judgments and decisions under conditions of risk or uncertainty.²³² As with the use of experiments designed to elicit error, the use of experiments designed to examine primarily the heuristic mode of thought may lead to an unrepresentative sample of findings regarding how judgment and decision making occur. The related but more general significance of dual-process models of cognition for legal decision theory is twofold: If legal actors are differentially

232. Dual-process models of information-processing also hold immediate importance for specific areas of legal decision theory. Legal scholars interested in how the law may encourage or discourage the use and existence of social stereotypes and in how social stereotypes are or should be taken into account in the law have recently drawn on psychological research on social stereotyping to inform their work. Some of these scholars treat social stereotypes as automatic and uncontrollable, occurring through a heuristic mode of thought. *See, e.g.,* Joseph W. Rand, *The Demeanor Gap: Race, Lie Detection, and the Jury*, 33 CONN. L. REV. 1, 41 (2000) ("Social science research has shown that the effect of racial stereotypes are so strong that they can influence the decision making of even the most well-meaning and consciously progressive individuals. This is not conscious racism, but the effect of stereotyping on the cognitive processes due to social conditioning that automatically forces the observer to see events in a certain fashion."); Amy L. Wax, *Discrimination as Accident*, 74 IND. L.J. 1129, 1158 (1999) ("Research in cognitive psychology suggests that biases in judgment stemming from categorical generalizations cannot be reliably manipulated or controlled either by the person harboring those biases or by outsiders seeking to redesign the decisionmaking process to reduce such bias. Inherent features of human cognition prevent individuals from detecting or effectively correcting all but the most egregious biases in their own judgments."); *see also* Michael Selmi, *Discrimination as Accident: Old Whine, New Bottle*, 74 IND. L.J. 1233, 1235-43 (1999) (responding to Professor Wax's article and discussing the concept of unconscious discrimination). Dual-process models of social cognition suggest that, while social stereotypes may serve as a default heuristic in much person perception, individuating information may also be taken into account via a less-automatic processing mode, and this mode may override the social stereotype arising from the heuristic mode under certain conditions. *See* Galen V. Bodenhausen et al., *On the Dialectics of Discrimination: Dual Processes in Social Stereotyping*, in DUAL-PROCESS THEORIES IN SOCIAL PSYCHOLOGY, *supra* note 224, at 271-90 (reviewing and evaluating dual-process models of person perception specifically in relation to social stereotypes). The body of research on person perception is quite complex, and questions remain over the degree of control that individuals may exert over the stereotyping process, if any. *See* Susan T. Fiske, *Stereotyping, Prejudice, and Discrimination*, in 2 HANDBOOK OF SOCIAL PSYCHOLOGY, *supra* note 9, at 364-75, 385-91 (reviewing evidence of automaticity in stereotyping and evidence that individuals may have some control over stereotyping). The dual-process theories of person perception should, at a minimum, be taken into account when working in this area in order to ensure proper qualifications to one's conclusions. *Cf.* Clark Freshman, *Whatever Happened to Anti-Semitism? How Social Science Theories Identify Discrimination and Promote Coalitions Between "Different" Minorities*, 85 CORNELL L. REV. 313, 442 (2000) (noting debate within psychology about the controllability of social stereotypes and biases).

capable of engaging in rational versus arational modes of thought, we need to (1) develop a theory of the conditions under which the rational versus the arational mode of thought is more likely to be triggered, and in what legal contexts, and for whom and (2) determine whether active interventions within the legal setting can trigger the rational mode of thought if the benefits of such interventions outweigh their costs.

CONCLUSION

Behavioral decision theory does not support the bold claim of legal decision theorists that all legal actors systematically fall prey to cognitive illusions when forming judgments and making decisions. Indeed, most behavioral decision research does not even address the prevalence or robustness of findings of cognitive bias and error, thus failing to provide the information that is most crucial to the application of this research to the legal system. Moreover, all of the cognitive biases and errors that the legal decision theorists tout in their attack on the rationality assumption are based on the responses of only subsets of experimental subjects under refined experimental conditions. Often a high percentage of subjects provide the rational response even under the experimental conditions designed to elicit error, and even higher percentages of subjects provide the rational response with changes in the experimental setting that make the decision problems less confusing, less misleading, or more ecologically valid.

We are left with the question of why the legal decision theorists have so misunderstood behavioral decision theory. Or, if they have not misunderstood behavioral decision theory, why they have presented behavioral decision theory's ambiguous evidence on the prevalence of irrationality as if this evidence were unambiguous. Legal decision theorists' misleading and partial portrayal of research into judgment and decision making may be due to one or more factors.²³³

233. With regard to the partial presentation of research, consider that recent surveys by Hanson & Kysar, *Behavioralism I*, *supra* note 1; Jolls et al., *supra* note 9; and Korobkin & Ulen, *supra* note 4, omit discussion of most of the research and issues discussed in this Article that do not comport with the legal decision theorists' view of judgment and decision making as pervasively and irremediably flawed.

First, market forces may be driving these scholars to puffery and exaggeration in their selling of the portrait of human irrationality. Whereas law and economics scholars market theories of law based on a model of rational and efficient decision making, the legal decision theorists market themselves as the law and economics realists who offer theories of law based on a model of irrational (but predictably irrational) decision making. In short, blanket irrationality probably sells better than a nuanced, contextualized picture of human behavior full of individual and situational differences in rationality and lacking in cognitive universals.

Second, behavioral decision theorists aid and abet the imprecision of the legal decision theorists by speaking as if all subjects in all experiments always commit the same errors, even though the subjects do not.²³⁴ Behavioral decision theorists often fail to specify clearly the boundary conditions on much of their own work, perhaps because they typically speak to a psychology audience that is expected to be aware of the limits on the research. Moreover, evidence exists of a possible citation bias within psychology that favors pessimistic accounts of decision making: "[I]t appears that reports of good performance virtually have been ignored and that it is possible that this has had an undue negative influence upon people's views about the quality of judgment and reasoning."²³⁵ If

234. See Jepson et al., *supra* note 128, at 495 ("Previous work on inductive heuristics has left many theorists with the impression that subjects in reasoning experiments behave quite uniformly."); see also David C. Funder, *Gone With the Wind: Individual Differences in Heuristics and Biases Undermine the Implication of Systematic Irrationality*, 23 BEHAV. & BRAIN SCI. 673, 673-74 (2000) (discussing the "attention-getting claim" of the heuristics and biases researchers that "human cognition [is] characterized by systematic irrationalities" and how this claim is arguably wrong) (footnote omitted); Lola L. Lopes, *The Rhetoric of Irrationality*, 1 THEORY & PSYCHOL. 65, 75-78 (1991) (discussing how an early, very influential article by Kahneman and Tversky presented an overly negative view of the use of cognitive heuristics); *id.* at 80 ("In the case of the biases and heuristics literature . . . one cannot criticize the message without criticizing the way it is packaged and presented. The view that people are irrational is real in the sense that people hold it to be true. But the reality is mostly in the rhetoric.").

235. Beach et al., *supra* note 128, at 51; see also Jay J. J. Christensen-Szalanski & Lee Roy Beach, *The Citation Bias: Fad and Fashion in the Judgment and Decision Literature*, 39 AM. PSYCHOLOGIST 75, 77 (1984) (discussing a disparity in scholarly citations that favors discussion of empirical studies showing "poor" judgment over studies showing "good" judgment and speculating that this "citation bias" may be due in part to a selective presentation of supportive evidence by proponents of bias research). But see Richard W. Robins & Kenneth H. Craik, *Is There a Citation Bias in the Judgment and Decision Literature?*, 54 ORGANIZATIONAL BEHAV. & HUM. DECISION PROCESSES 225, 242 (1993)

such a bias exists in the reports of the research, legal decision theorists, as consumers of this literature, would naturally be left with a pessimistic but unrepresentative account of human behavior.

Third, the legal decision theorists may themselves have fallen prey to a cognitive or motivational bias that causes them to ignore research that conflicts with their views. In fact, the legal decision theorists' own argument of pervasive, unavoidable biases would suggest that the partial presentation of behavioral decision theory is due to the operation of confirmation and disconfirmation biases (or more generally "myside bias") within the legal decision theorists themselves: these theorists selectively seek and attend to evidence supportive of their own views and exclude or discount evidence not supportive of their views.²³⁶ If this is true, then legal decision theorists are in good company, for psychologists have been found to exhibit a confirmatory bias.²³⁷ This bias explanation is possibly more flattering than two alternative explanations—ignorance or deliberate neglect of relevant research.

Fourth, it may be that the legal decision theorists have fully and fairly evaluated work that brings into question the error and bias

(reviewing Christensen-Szalanski & Beach's 1984 study and finding "empirical support for a citation-rate disparity [in favor of 'poor judgment' studies that] is modest at best and possibly unreliable").

236. See BARON, *supra* note 20, at 195 ("People tend not to look for evidence against what they favor, and, when they find it anyway, they tend to ignore it. David Perkins has named these two characteristics 'myside bias.'") (citation omitted); BAZERMAN, *supra* note 25, at 35 ("Most of us seek confirmatory evidence and exclude the search for disconfirming information from our decision process.") (emphasis omitted); Kari Edwards & Edward E. Smith, *A Disconfirmation Bias in the Evaluation of Arguments*, 71 J. PERSONALITY & SOC. PSYCHOL. 5, 6 (1996) ("When faced with evidence contrary to their beliefs, people try to undermine the evidence. That is, there is a bias to disconfirm arguments incompatible with one's position.").

237. Mahoney conducted an experiment to test for confirmatory bias in referees of psychological research, and he found that "(a) referee evaluations may be dramatically influenced by such factors as experimental outcome, and (b) interreferee agreement may be extremely low on factors relating to manuscript evaluation." Mahoney, *supra* note 115, at 173. Mahoney appropriately asks, "[w]ith our vast literatures on information processing and social psychology, have we assumed that scientists are somehow unaffected by the processes which appear to be so common in other members of the species?" *Id.*; see also Jonathan J. Koehler, *The Influence of Prior Beliefs on Scientific Judgments of Evidence Quality*, 56 ORGANIZATIONAL BEHAV. & HUM. DECISION PROCESSES 28, 47 (1993) ("In the present study, even scientists who had a good deal of methodological training and research experience differentially perceived the quality of scientific research as a function of how well the data supported their beliefs and how strongly the beliefs were held."). On the general topic of the biased use of psychological research see MacCoun, *supra* note 37.

research and found this critical work to be too lacking in merit to warrant mention. This explanation seems unlikely, however, given the quality of this work relative to the quality of the error and bias research, given its publication in leading peer-edited journals, and given the raising of cautions by leading behavioral decision theorists. A final, partial explanation is simply that some of the research discussed here is relatively new and did not exist at the time of publication of some of the early legal decision theory articles.

For whatever reason, many legal scholars use insufficient care and precision in their interpretations and uses of psychological research on judgment and decision making. Consumers of this growing literature should thus look very skeptically on the claims being made and should resist the contention that the cognitive-miser model being offered by these scholars is more complete and accurate than the rational-actor model. Legal decision theorists argue that the economic analysis of law is incomplete and incorrect as a descriptive account of judgment and choice and, therefore, legal policy based on standard economic prescriptions is suspect. Yet in the end, all that legal decision theory presently offers is a partial account of human behavior to compete with the partial account offered by law and economics.²³⁸ Some adherents of legal decision theory acknowledge this status but hope for more.²³⁹ Progress is

238. This conclusion accords with Professor Cross' appraisal:

Behavioral economics criticizes classical economics for failing to describe the reality of human behavior, which arguably leads to undesirable policy prescriptions. Yet one can criticize much behavioral economics for failing to make accurate descriptive predictions or to produce policy prescriptions that are demonstrably superior to those of classical economics. . . .

The prevailing battle between law and economics and behavioral economics is . . . misguided. The conventional rational decision-making presumptions of classic law and economics are amply demonstrated to a degree, but the theory does not explain all behavior completely. Behavioralism is not so much an alternative to law and economics as it is a complement. It supplements the classic model and explains why deviations may occur from the model, but it does not supplant that model. Both models are valuable only insofar as they explain actual behavior, and their descriptive validity can be tested empirically, yielding the knowledge necessary for policy.

Frank B. Cross, *In Praise of Irrational Plaintiffs*, 86 CORNELL L. REV. 1, 32 (2000); see also Hillman, *supra* note 47, at 718 ("[I]f legal theorists focus too narrowly on particular behavioral observations, their analysis will be no more realistic than predictions based on economic analysis's wealth-maximization precept.").

239. See, e.g., Jolls et al., *supra* note 9, at 1547 ("We hope that this article will encourage others to continue the inquiry and research, both theoretical and empirical, that will be

unlikely, however, so long as legal decision theorists fail to specify the boundary conditions on their empirical claims and instead settle for overly broad generalizations about nonrationality to compete with the economic theorists' overly broad generalizations about rationality.

Legal decision theorists address important questions. Do our rules of evidence truly serve the functions we intend? Are there predictable unwanted biases in how juries and judges award damages, and, if so, how might these biases be combated? Are bureaucrats capable of fairly applying cost-benefit analyses? Research into such questions should of course continue, but more cautiously and carefully, and the conclusions drawn from empirical research should be as precise as possible and should be accompanied by self-critical candor. Just as we should not base our legal rules and standards on faulty assumptions about the basic rationality of people, so should we not base reforms of our legal rules and standards on faulty claims about the basic irrationality of people.