Vanderbilt Law Review

Volume 56
Issue 6 *Issue 6 - November 2003*

Article 2

11-2003

Chicago Man, K-T Man, and the Future of Behavioral Law and Economics

Robert A. Prentice

CHICAGO MAN, K-T MAN, AND THE FUTURE OF BEHAVIORAL LAW AND ECONOMICS

*Robert A. Prentice**

* University Distinguished Teaching Professor and Ed & Molly Smith Centennial Professor of Business Law, McCombs School of Business, University of Texas at Austin. The author thanks Jonathan J. Koehler, Jeffrey Rachlinksi, and Frank Cross for their helpful comments. All views expressed herein are mine alone.

I. INTRODUCTION

Most law is aimed at shaping human behavior, encouraging that which is good for society and discouraging that which is bad.[1]

1. *See* Jeremy A. Blumenthal, *Law and Social Science in the Twenty-First Century*, 12 S. CAL. INTERDISC. L.J. 1, 52 (2002) ("[T]he legal system is fundamentally based on assumptions about human behavior."); Chris Guthrie, *Prospect Theory, Risk Preference, & the Law*, 97 NW. U. L. REV. 1115, 1115 (2003) ("Only with an understanding of how people are likely to respond to legal rules can legal scholars, judges, legislators, and regulators craft rules that encourage

Nonetheless, for most of the history of our legal system, laws were passed, cases were decided, and academics pontificated about the law based on nothing more than common sense assumptions about how people make decisions.[2] A quarter century or more ago, the law and economics movement[3] replaced these common sense assumptions with a well-considered and expressly stated assumption—that man is a rational maximizer of his expected utilities.[4] Based on this premise,

---

desirable behavior and discourage undesirable behavior."); Donald C. Langevoort, *Behavioral Theories of Judgment and Decision Making in Legal Scholarship: A Literature Review*, 51 VAND. L. REV. 1499, 1499 (1998) ("Nearly all interesting legal issues require accurate predictions about human behavior to be resolved satisfactorily.").

2. *See* HUGO MUNSTERBERG, ON THE WITNESS STAND 10-11 (1908) (complaining that "[t]he lawyer and the judge and the juryman are sure that they do not need the experimental psychologist. . . . They go on thinking that their legal instinct and their common sense supplies [sic] them with all that is needed and somewhat more.").

3. The importance of economic analysis to law is usually traced to two seminal articles by Calabresi and Coase. *See* Guido Calabresi, *Some Thoughts on Risk Distribution and the Law of Torts*, 70 YALE L.J. 499 (1961) (analyzing the economic logic of tort law); Ronald Coase, *The Problem of Social Cost*, 3 J.L. & ECON. 1 (1960) (using economic principles to analyze nuisance law). However, the true founder of the law and economics movement is Judge Richard Posner. *See* RICHARD A. POSNER, ECONOMIC ANALYSIS OF LAW (1st ed. 1973).

4. POSNER, *supra* note 3, at 2-6; *see also* PAUL HEYNE, THE ECONOMIC WAY OF THINKING 2 (2d ed. 1976) (noting that the rational man model is "basically a way of thinking" and that economics assumes that everyone "acts in accordance with that rule: miser or spendthrift, saint or sinner, consumer or seller, politician or business executive, cautious calculator or spontaneous improviser"); Gregory S. Crespi, *Does the Chicago School Need to Expand Its Curriculum?*, 22 LAW & SOC. INQUIRY 149, 150-51 (1997) (observing that the two core assumptions of law and economics are that decision makers are rational actors and that the appropriate goal of policy is wealth maximization); Roger G. Noll & James E. Krier, *Some Implications of Cognitive Psychology for Risk Regulation*, 19 J. LEGAL STUD. 747, 750-51 (1990) (summarizing core assumptions of the standard model); Richard A. Posner, *Values and Consequences: An Introduction to Economic Analysis of Law*, *in* CHICAGO LECTURES IN LAW AND ECONOMICS 189, 191 (Eric A. Posner ed., 2000) ("Most economic analysis consists of tracing out the consequences of assuming that people are more or less rational in their social interactions."); W. Kip Viscusi, *Individual Rationality, Hazard Warnings, and the Foundations of Tort Law*, 48 RUTGERS L. REV. 625, 636 (1996) (noting that the "foundation of economic analysis of choice is based on the rationality of individual decision making").

According to this viewpoint:

> Individuals are assumed to act *as if* they maximize expected utility. That is, an individual's preferences are taken as given, consistent, and representable in the form of a utility function. An individual knows a priori the set of alternative actions and chooses the action with the highest utility or expectation thereof. When uncertainty exists as to the actions' consequences, an individual can assess the probability distribution corresponding to his or her knowledge. When new information may be collected from the environment, an individual knows the information's possible content and can assess, in accord with Bayes' theorem, the probability distribution conditioned on the conjunction of such content and his or her prior knowledge.

William S. Waller, *Decision-Making Research in Managerial Accounting: Return to Behavioral-Economics Foundations*, *in* JUDGMENT AND DECISION-MAKING RESEARCH IN ACCOUNTING AND AUDITING 29, 32 (Robert H. Ashton & Alison H. Ashton eds., 1995); *see also* Thomas S. Ulen, *Cognitive Imperfections and the Economic Analysis of Law*, 12 HAMLINE L. REV. 385, 386 (1989) (providing a similar summary).

law and economics has dominated interdisciplinary thought in the legal academy for the past thirty years.[5]

In the past decade it has become clear, however, that people simply do not make decisions as modeled by traditional law and economics.[6] A "mountain of experiments"[7] performed in psychology

---

5. *See* Anne C. Dailey, *The Hidden Economy of the Unconscious*, 74 CHI.-KENT L. REV. 1599, 1600 (1999) ("Economic analysis has without question enjoyed a powerful and widespread influence within the legal academy over the last few decades."); W. Bradley Wendel, *Mixed Signals: Rational-Choice Theories of Social Norms and the Pragmatics of Explanation*, 77 IND. L.J. 1, 8 (2002) ("[T]he rational-choice vision of the human predicament has achieved unparalleled dominance in the legal academy in thinking about individual and social behavior . . . .").

This simplifying approach has produced useful insights in other fields as well. *See* Bruno S. Frey & Matthias Benz, *From Imperialism to Inspiration: A Survey of Economics and Psychology* 3 (Univ. of Zurich, Inst. For Empirical Research in Econ., Working Paper No. 118, May 2002) ("Often termed 'economic imperialism,' the economic approach has produced fruitful insights in such areas as politics ('Public Choice'), law ('Law and Economics'), history ('New Economic History'), the arts ('Cultural Economics'), or family ('Economics of the Family')"), http://papers.ssrn.com/sol3/papers.cfm?ahstract_id=316681.

6. *See* HERBERT SIMON, REASON IN HUMAN AFFAIRS 13 (1983) ("Conceptually, the SEU [Subjective Expected Utility] Model is a beautiful object deserving a prominent place in Plato's heaven of ideas. But vast difficulties make it impossible to employ it in any literal way in making actual human decisions."); Kenneth G. Dau-Schmidt, *Economics and Sociology: The Prospects for an Interdisciplinary Discourse on Law*, 1997 WIS. L. REV. 389, 397 ("The assumptions of the neoclassical model are clearly unrealistic, and the importance of this lack of realism has been a matter of some debate both within and outside the discipline."); R. Duncan Luce & Detlof von Winderfeldt, *What Common Ground Exists for Descriptive, Prescriptive, and Normative Utility Theories?*, 40 MGMT. SCI. 263, 263 (1994) (defending the normative properties of the classic view, but admitting that "[t]here is no doubt that [subjective utility theory] is descriptively wrong"); Daniel McFadden, *Rationality for Economists?*, 19 J. RISK & UNCERTAINTY 73, 97 (1999) ("Confronted with the accumulated evidence, economists must recognize that the Chicago-man model does not apply universally, or even regularly, to choices made in non-market contexts."); Charles R.P. Pouncy, *The Rational Rogue: Neoclassical Economic Ideology in the Regulation of the Financial Profession*, 26 VT. L. REV. 263, 264 (2002) ("Economic rationality as it is currently deployed is a grossly inadequate approximation of the factors motivating human conduct."); Paul J. H. Schoemaker, *The Expected Utility Model: Its Variants, Purposes, Evidence, and Limitations*, 20 J. ECON. LITERATURE 529, 530 (1982) ("[M]ost of the empirical evidence is difficult to reconcile with the principle of [expected utility] maximization."); Frey & Benz, *supra* note 5, at 11 ("Over the last two decades, however, a large literature has accumulated that shows botb experimentally and theoretically that the theory of expected utility maximization can explain only a limited part of observed behavior.").

7. Conlisk notes:

> There is a mountain of experiments in which people: display intransitivity; misunderstand statistical independence; mistake random data for patterned data and vice versa; fail to appreciate law of large number effects; fail to recognize statistical dominance; make errors in updating probabilities on the basis of new information; understate the significance of given sample sizes; fail to understand covariation for even the simplest 2X2 contingency tables; make false inferences about causality; ignore relevant information; use irrelevant information (as in sunk cost fallacies); exaggerate the importance of vivid over pallid evidence; exaggerate the importance of fallible predictors; exaggerate the ex ante probability of a random event which has already occurred; display overconfidence in judgment relative to evidence; exaggerate confirming over disconfirming evidence relative to initial beliefs; give answers that are highly sensitive to logically irrelevant changes in questions; do redundant and

and related disciplines, much of it in the "heuristics and biases" tradition founded by psychologists Daniel Kahneman and Amos Tversky,[8] demonstrate that people tend to deviate systematically from rational norms when they make decisions.

The implications of a fundamental inaccuracy in a foundational pillar of the legal academy's leading theoretical construct are obviously substantial. The essential inaccuracy of the rational man model has minimized the capacity of law and economics to generate useful insights in many areas of the law.[9]

Dissatisfaction with this state of affairs gave rise to a movement, variously called Behavioral Law and Economics (BLE), Behavioral Decision Theory (BDT), and Legal Decision Theory (LDT), that seeks to provide a more descriptively and predictively accurate account of human behavior; this is done by replacing the law and economics movement's stylized rational man model with a more accurate model based on empirical research arising from psychology, cognitive science, behavioral biology, decision theory, and related fields.

In a relatively brief period, a raft of legal decision theorists have authored scores of articles that make up a growing body of behavioralist literature.[10] The new movement's momentum has not

---

> ambiguous tests to confirm an hypothesis at the expense of decisive tests to disconfirm; make frequent errors in deductive reasoning tasks such as syllogisms; place higher value on an opportunity if an experimenter rigs it to be the "status quo" opportunity; fail to discount the future consistently; fail to adjust repeated choices to accommodate intertemporal connections; and more.

John Conlisk, *Why Bounded Rationality?*, 34 J. ECON. LITERATURE 669, 670 (1996); *see also* Larry T. Garvin, *Adequate Assurance of Performance: Of Risk, Duress, and Cognition*, 69 U. COLO. L. REV. 71, 145 (1998) ("Cognitive psychology and experimental economics have found a smorgasbord of cognitive errors, which collectively falsify most of the axioms of rational choice theory."); Schoemaker, *supra* note 6, at 552 ("[A]t the individual level EU [expected utility] maximization is more the exception than the rule.").

8. Heuristics are mental shortcuts, rules of thumb that often depart from normative standards of rationality. *See* Jonathan St. B.T. Evans, *Heuristic and Analytic Processes in Reasoning*, 75 BRIT. J. PSYCHOL. 451, 462 (1984) (noting that a bias is "a source of error wbich is systematic rather than random, and consists of either failure to take account of a normatively relevant feature or else a tendency to respond to a normatively irrelevant feature"). *See generally* THOMAS GILOVICH ET AL., HEURISTICS AND BIASES: THE PSYCHOLOGY OF INTUITIVE JUDGMENT (2002) [hereinafter PSYCHOLOGY OF INTUITIVE JUDGMENT]; DANIEL KAHNEMAN & AMOS TVERSKY, CHOICES, VALUES, AND FRAMES (2000) [hereinafter CHOICES, VALUES, AND FRAMES]; DANIEL KAHNEMAN ET AL., JUDGMENT UNDER UNCERTAINTY: HEURISTICS AND BIASES (1982) [hereinafter JUDGMENT UNDER UNCERTAINTY].

9. *See supra* notes 1-8 and accompanying text.

10. *See, e.g.*, Hal R. Arkes & Cindy Schipani, *Medical Malpractice v. the Business Judgment Rule: Differences in Hindsight Bias*, 73 OR. L. REV. 587 (1994) (studying the psychology literature's hindsight bias regarding the contrasting rules for reviewing decisions of doctors and directors); Larry A. DiMatteo, *A Theory of Efficient Penalty: Eliminating the Law of Liquidated Damages*, 38 AM. BUS. L.J. 633 (2001) (analyzing the liquidated damages doctrine through a behavioral lens); Melvin A. Eisenberg, *The Limits of Cognition and the Limits of Contract*, 47

been blunted by various published critiques,[11] but in two new articles

---

STAN. L. REV. 211 (1995) (arguing that psychological constrains on cognition explain many aspects of contract law); 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) (suggesting that insights about actual human behavior can improve economic analysis of law); Garvin, *supra* note 7 (arguing that the behavioral literature's puncturing of the rational economic actor myth should affect the content of specified provisions of the Uniform Commercial Code); Chris Guthrie et al., *Inside the Judicial Mind,* 86 CORNELL L. REV. 777 (2001) (reporting the results of a study indicating that judges are subject to various heuristics and biases inconsistent with the rational man model); Russell B. Korobkin & Thomas S. Ulen, *Law and Behavioral Science: Removing the Rationality Assumption from Law and Economics,* 88 CAL. L. REV. 1051 (2000) (noting significant flaws in law and economics' theory's core behavioral assumption that people act rationally while suggesting a more nuanced approach); Russell Korobkin & Chris Guthrie, *Psychological Barriers to Litigation Settlement: An Experimental Approach,* 93 MICH. L. REV. 107 (1994) (arguing that economics-based rational-actor models, used to analyze why parties fail to settle litigation, be replaced with richer behavioral-based models); Donald C. Langevoort, *Monitoring: The Behavioral Economics of Corporate Compliance with Law,* 2002 COLUM. BUS. L. REV. 71 (2002) [hereinafter Langevoort, *Behavioral Economics*] (analyzing corporate governance issues through a behavioral lens); Donald C. Langevoort, *Organized Illusions: A Behavioral Theory of Why Corporations Mislead Stock Market Investors (And Cause Other Social Harms),* 146 U. PA. L. REV. 101 (1997) [hereinafter Langevoort, *Organized Illusions*] (using behavioral analysis to discuss why corporations commit securities fraud); Donald C. Langevoort, *Taming the Animal Spirits of the Stock Markets: A Behavioral Approach to Securities Regulation,* 97 NW. U. L. REV. 135 (2002) [hereinafter Langevoort, *Taming the Animal Spirits*] (providing behavioral insights useful to securities market regulation); Donald C. Langevoort, *Where Were the Lawyers? A Behavioral Inquiry Into Lawyers' Responsibility for Clients' Fraud,* 46 VAND. L. REV. 75 (1993) (providing an analysis of why securities lawyers sometimes get involved in their clients' fraud); Jeffrey J. Rachlinski, *A Positive Psychological Theory of Judging in Hindsight,* 65 U. CHI. L. REV. 571 (1998) (arguing that the law is already coping well with the hindsight bias); Jeffrey J. Rachlinski, *Heuristics and Biases in the Courts: Ignorance or Adaptation?,* 79 OR. L. REV. 61 (2000) (analyzing how courts variously ignore or adapt to behavioral considerations) [hereinafter Rachlinski, *Heuristics and Biases*]; Jeffrey J. Rachlinski & Forest Jourden, *Remedies and the Psychology of Ownership,* 51 VAND. L. REV. 1541 (1998) (presenting behavioral evidence with implications for damage and liability rules); Lynn A. Stout, *In Praise of Procedure: An Economic and Behavioral Defense of* Smith v. Van Gorkom *and the Business Judgment Rule,* 96 NW. U. L. REV. 675 (2002) (using both behavioral and economic analysis to defend the controversial holding in *Smith v. Van Gorkom*); Cass R. Sunstein et al., *Assessing Punitive Damages (with Notes on Cognition and Valuation in Law),* 107 YALE L.J. 2071 (1998) (reporting the results of a behavioral study with implications for punitive damages doctrine); Cass R. Sunstein, *Behavioral Analysis of Law,* 64 U. CHI. L. REV. 1175 (1997) [hereinafter Sunstein, *Behavioral Analysis*] (providing an introduction to the behavioral field and its relevance to law); Cass R. Sunstein et al., *Predictably Incoherent Judgments,* 54 STAN. L. REV. 1153 (2002) (giving a behavioral explanation for the wide variance in jury verdicts); Cass R. Sunstein, *The Laws of Fear,* 115 HARV. L. REV. 1119 (2000) (reviewing PAUL SLOVIC, THE PERCEPTION OF RISK (2000)).

11. *See, e.g.*, Jennifer Arlen, *The Future of Behavioral Economic Analysis of Law,* 51 VAND. L. REV. 1765 (1998) (arguing that behavioral research is still too messy to offer a neat basis for resolving problems and that law and economics, with its simplifying assumptions, should remain the dominant interdisciplinary paradigm); Mark Kelman, *Behavioral Economics as Part of a Rhetorical Duet: A Response to Jolls, Sunstein, and Thaler,* 50 STAN. L. REV. 1577 (1998) (noting limitations on behavioral theory's contributions to law); Richard A. Posner, *Rational Choice, Behavioral Economics, and the Law,* 50 STAN. L. REV. 1551 (1998) (criticizing, on several dimensions, the application of behavioral theory to legal analysis); Tanina Rostain, *Educating Homo Economicus: Cautionary Notes on the New Behavioral Law and Economics Movement,* 34 LAW & SOC'Y REV. 973 (2000) (noting difficulties in using laboratory-based psychology

Professor Gregory Mitchell seeks to change that.[12] Because Mitchell has a Ph.D in psychology and uses controversies within the psychology discipline itself to launch a broadside attack on what he terms the Legal Decision Theory movement, his articles pose a credible threat to this new interdisciplinary scholarship.[13] In other words, Mitchell presents the proverbial "threat from within."

In both articles, Mitchell challenges the competence, motives, methods and claims of legal decision theorists. In his first article, he focuses his attacks on the legitimacy of most social science research, especially the heuristics and biases literature launched by Kahneman and Tversky that many legal decision theorists favor.[14] In his second article, Mitchell argues that there is much greater variation in human reasoning than legal decision theorists have allowed and that this variability prevents legal decision theory from offering useful policy prescriptions.[15]

---

experiments as the basis for legal reform); Robert E. Scott, *The Limits of Behavioral Theories of Law and Social Norms*, 86 VA. L. REV. 1603 (2000) (noting the difficulty of generalizing appropriate legal norms from particular behavioral studies).

12. *See* Gregory Mitchell, *Taking Behavioralism Too Seriously? The Unwarranted Pessimism of the New Behavioral Analysis of Law*, 43 WM. & MARY L. REV. 1907 (2002) [hereinafter Mitchell, *Pessimism*]; Gregory Mitchell, *Why Law and Economics' Perfect Rationality Should Not Be Traded for Behavioral Law and Economics' Equal Incompetence*, 91 GEO. L.J. 67 (2002) [hereinafter Mitchell, *Incompetence*].

13. Although I do not necessarily prefer LDT as a label for the new movement, I adopt it for purposes of this article since I am responding to two articles in which Mitchell uses the term and believe that Mitchell makes a credible case for use of the term. *See* Mitchell, *Incompetence*, *supra* note 12, at 78-83.

Like Mitchell, I think that the following quotation from Dawes captures the field in a general 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.

Robyn M. Dawes, *Behavioral Decision Making and Judgment*, *in* 1 HANDBOOK OF SOCIAL PSYCHOLOGY 497, 497 (Daniel T. Gilbert et al. eds., 4th ed. 1998).

14. Mitchell, *Pessimism*, *supra* note 12.

15. Mitchell, *Incompetence*, *supra* note 12. In arguing that legal decision theorists have ignored controversies within the psychology establishment regarding the validity and consistency of the Kahneman and Tversky heuristics and biases literature, Mitchell mirrors earlier criticisms of the law and economics scholarship. *See, e.g.*, Martha C. Nussbaum, *Flawed Foundations: The Philosophical Critique of (a Particular Type of) Economics*, 64 U. CHI. L. REV. 1197, 1197 (1997) (noting that the law and economics "movement has virtually ignored criticisms of its foundations that are increasingly influential in mainstream economics").

In some sense, Mitchell reframes the debate between two leading views of how people make decisions. On the one hand is the law and economics movement's rational man model—what Nobel Prize-winning economist Daniel McFadden terms "Chicago Man."[16] On the other hand is the leading model for decision making research,[17] the behavioral model described in the Kahneman and Tversky heuristics and biases line of research that McFadden terms "K-T Man."[18] Hence, my title.

If Mitchell is correct, the new field of legal decision theory (or behavioral law and economics or behavioral decision theory) holds very little promise. Believing that Mitchell is wrong, in Part II of this article, I briefly compare and contrast the relatively mature law and economics field with its more youthful counterpart, legal decision theory, in order to give to any readers unfamiliar with this new scholarship a flavor for what it seeks to accomplish.[19]

In Part III, I assess the attack Mitchell makes on the validity of social science research.[20] I shall show that there is substance to his arguments, but that he has failed to significantly undermine that body of research.

In Part IV, I address Mitchell's claim that individual and situational factors cause great variation in decision making, thereby destroying any uniformity needed to predict human behavior under the K-T Man model.[21] I will demonstrate that a large number of straw men were born and killed in the construction of Mitchell's arguments. In defending most legal decision theorists against Mitchell's claims, I will make clear that I believe the pretensions of such theorists to be not nearly so grandiose as those Mitchell ascribes to them.

In Part V, I briefly assess the future of legal decision theory in the aftermath of Mitchell's attacks.[22] I argue that the movement retains great potential to add valuable insights to legal scholarship, despite Mitchell's withering attack.

---

16. McFadden, *supra* note 6, at 76.

17. Anton Kuhberger et al., *Framing Decisions: Hypothetical and Real,* 89 ORGANIZATIONAL BEHAV. & HUM. DECISION PROCESSES 1162, 1162 (2002) (noting that Kahneman and Tversky's "heuristics and biases approach is probably the dominant tradition in decision research").

18. McFadden, *supra* note 6, at 83. K-T Man's dimensions are shaped by both Kahneman and Tverksy's heuristics and biases literature dealing with mental shortcuts under conditions of uncertainty and their prospect theory literature, focusing on the nature of people's utility functions under conditions of risk. *See* CASS R. SUNSTEIN, MORAL HEURISTICS 2 n.2 (Univ. of Chi., Law & Econ. Olin Working Paper No. 180, Mar. 2003, http://ssrn.com/abstract_id=387941.

19. *See infra* notes 23-59 and accompanying text.

20. *See infra* notes 60-307 and accompanying text.

21. *See infra* notes 308-528 and accompanying text.

22. *See infra* notes 529-578 and accompanying text.

II. A BRIEF OVERVIEW OF LEGAL DECISION THEORY

Despite its dominance, law and economics has been controversial. Critical Legal Studies (CLS) arose, in large part, as a response to the methods and values of law and economics,[23] but it has since largely died away. Dissatisfaction with law and economics has remained, however, because efficiency simply does not explain why the law is as it is,[24] despite the redoubtable Judge Posner's claims to the contrary.[25] Although economic analysis has improved the rigor of some legal analysis and has shed valuable light on some topics, such

---

23. *See* Daniel T. Ostas, *Postmodern Economic Analysis of Law: Extending the Pragmatic Visions of Richard A. Posner*, 36 AM. BUS. L.J. 193, 194 (1998) (noting that the Critical Legal Studies movement began "at least in part, as an attack on the perceived short-comings of the method and values of" law and economics).

24. It is difficult to accept the law and economics position that the common law is best explained by efficiency criteria when judges (and jurors for that matter) do not typically emphasize efficiency grounds when making decisions, as Baron and Ritov discovered in empirical tests.

> Our results create a puzzle for positive economic theories of law, particularly that of Landes and Posner (1987). If the system can be understood in terms of the consequentialist rationale, as they claim it can, what human judgments maintain it? Note that our main findings held even for judges, and most of our other subjects are potential jury members. Perhaps the present system is not so close to be[ing] the "best of all possible consequentialist worlds," as Landes and Posner would suggest.

Jonathan Baron & Ilana Ritov, *Intuitions About Penalties and Compensation in the Context of Tort Law*, 7 J. RISK & UNCERTAINTY 17, 32 (1993); *see also* Jonathan Baron, *Heuristics and Biases in Equity Judgments: A Utilitarian Approach*, *in* PSYCHOLOGICAL PERSPECTIVES ON JUSTICE 109, 111 (Barbara Mellers & Jonathan Baron eds., 1993) ("Utilitarianism often conflicts with our intuitive beliefs about what is morally right."); Kevin M. Carlsmith et al., *Why Do We Punish? Deterrence and Just Deserts as Motives for Punishment*, 83 J. PERSONALITY & SOC. PSYCHOL. 284, 295 (2002) (finding, inconsistent with economic reasoning, that "despite strongly stated preferences for deterrence theory, [subjects'] individual sentencing decisions seemed driven exclusively by just deserts concerns"); John M. Darley et al., *Incapacitation and Just Deserts as Motives for Punishment*, 24 LAW & HUM. BEHAV. 659, 676 (2000) (finding that a person's desire to punish is based primarily upon a just deserts motive rather than a deterrence rationale); Heidi Li Feldman, *Prudence, Benevolence, and Negligence: Virtue Ethics and Tort Law*, 74 CHI.-KENT L. REV. 1431, 1434 (2000) ("Lay jurors possess no particular expertise in economic analysis . . . [and c]ivil negligence actions do not ask jurors to apply a standard of care that even refers to these matters."); Jonathan J. Koehler & Andrew D. Gershoff, *Betrayal Aversion: When Agents Cause the Very Harm They Are Supposed to Prevent*, ORGANIZATIONAL BEHAV. & HUM. DECISION PROCESSES (forthcoming 2004) (finding, contrary to economic theory, that mock jurors did not assign greater punishment to a thief whose crime was harder to detect); Cass R. Sunstein et al., *Do People Want Optimal Deterrence?*, 29 J. LEGAL STUD. 237, 248 (2000) (finding that an economic approach to deterrence in the legal system is broadly rejected by the public).

25. WILLIAM M. LANDES & RICHARD A. POSNER, THE ECONOMIC STRUCTURE OF TORT LAW 1 (1987) (arguing that the common law of torts is best explained as if the judges were trying to promote efficient resource allocation); *see* Wendel, *supra* note 5, at 4 (noting that law and economics "adherents claimed it explains everything from nuisance remedies (which it probably does) to sexual idiosyncracies [sic], to racial discrimination, to holiday customs, to the whole universe of social norms . . .").

as antitrust,[26] the rational man model, with no room for cognitive limitations, emotion, or altruism, describes neither how man does act nor how man should act.[27] Therefore, as noted earlier, law and economics has provided relatively few useful insights in most areas of the law.

- *Contract law.* As Eric Posner recently noted, thirty years of economic analysis of contract law "has failed to produce an 'economic theory' of contract law, and does not seem likely to be able to do so,"[28] and has not had any particular impact on judicial decision making,

---

26. *See* Gerald B. Wetlaufer, *Systems of Belief in Modern American Law: A View from Century's End,* 49 AM. U. L. REV. 1, 34-35 (1999) ("Whole bodies of law are now exclusively within [law and economics'] domain, including antitrust, economic regulation, and major portions of corporate law.").

Even in these areas, psychological evidence can add helpful insights. Gerla has observed that "[t]he nature of human information processing makes the dissemination of false information an almost ideal strategic tool for raising rivals' costs." Harry S. Gerla, *Federal Antitrust Law and the Flow of Consumer Information,* 42 SYRACUSE L. REV. 1029, 1063 (1991); *see also* Harry S. Gerla, *The Psychology of Predatory Pricing: Why Predatory Pricing Pays,* 39 SW. L.J. 755, 779 (1985) ("[P]sychology, as the science of human behavior, has relevance to any aspect of antitrust analysis that involves assumptions with respect to human behavior, whether the assumptions relate to the behavior of humans as consumers or as managers of business enterprises.").

Professor Langevoort has penned several behavioral articles that lend insight to corporate governance issues. *See, e.g.*, Langevoort, *Behavioral Economics, supra* note 10, at 71 (analyzing corporate governance issues through a behavioral lens); Langevoort, *Organized Illusions, supra* note 10, at 101 (using behavioral analysis to discuss question of why corporations commit securities fraud).

And several authors have shown that psychological research can give us a deeper understanding of regulatory issues than mere law and economics. *See, e.g.*, Mark Seidenfeld, *Cognitive Loafing, Social Conformity, and Judicial Review of Agency Rulemaking,* 87 CORNELL L. REV. 486, 547-48 (2002) (using psychological principles to examine the interplay of judicial review and agency decision making); Cass R. Sunstein, *Cognition and Cost-Benefit Analysis,* 29 J. LEGAL STUD. 1059, 1096 (2000) (suggesting cost-benefit analysis of regulations as a reasonable response to citizens' psychological misperceptions of risk).

27. Law and economics has been accused of having "an almost pathological aversion to explanations that appeal to values, commitments, loyalties, relationships, or emotions." Wendel, *supra* note 5, at 3. It has also been accused of crowding out considerations of law and morality. *See* David A. Hoffman & Michael P. O'Shea, *Can Law and Economics Be Both Practical and Principled?,* 53 ALA. L. REV. 335, 339, 420 (2002) (noting that "legal economists have generally proceeded without a well-articulated moral basis" and expressing doubts that the question posed in their title can be answered affirmatively); Pouncy, *supra* note 6, at 264, 281 ("[E]conomic rationality acts as a cultural contaminant, devaluing other moral and cultural considerations and obscuring the mechanisms through which business organizations create and exercise economic power. . . . The decisional structure envisioned by neoclassical economic theory leaves little room for the operation of ethics and morality in its model-building project.").

28. Eric Posner, *Economic Analysis of Contract Law After Three Decades: Success or Failure?,* 112 YALE L.J. 829, 830 (2003). Posner also argues that "[s]cholarship influenced by cognitive psychology has so far produced few insights." *Id.* at 829. I disagree with that assessment but, in any event, behavioral research still has twenty years or so before it can match law and economics' record of futility.

statutory law, or regulatory law in the contract field.[29] The failure can be traced in large part to the false premise that man is a rational actor.[30]

- *Tort law.* Izhak Englard notes that "the law and economics movement has had little but a rhetorical effect upon contemporary processes of tort adjudication."[31] Courts and legislatures seldom cite law and economics scholars, and infrequently adopt their positions.[32] The reason, it has been suggested, is that the theory is inaccurate.[33]
- *Criminal law.* Empirical studies demonstrate that the conventional economic model's assumption that "[i]ndividuals will comply with a legal prohibition if the expected penalty—the expected cost to them of the violation—will exceed the gain they expect to derive from the violation"[34] is typically inaccurate and the theories it generates unhelpful.[35] As I have suggested

---

29. *Id.* at 870.

30. As Posner notes:

> Fundamental assumptions, common to nearly all efforts at economic analysis, are that individuals have preferences over outcomes; these preferences obey basic consistency conditions; and individuals satisfy these preferences subject to an exogenous budget constraint. Contract scholars usually assume that individuals do not have preferences regarding the consumption or well-being of other individuals, nor regarding contract doctrine itself—there is no preference of expectation damages, for example.

Id. at 832.

31. William E. Nelson, *From Fairness to Efficiency: The Transformation of Tort Law in New York, 1920-1980,* 47 BUFF. L. REV. 117, 121 (1999) (citing Izhak Englard, *Law and Economics in American Tort Cases: A Critical Assessment of the Theory's Impact on Courts,* 41 U. TORONTO L.J. 359 (1991)).

32. *See* Ronald J. Allen & Ross M. Rosenberg, *The Actual Practice: Legal Phenomena, Knowledge, and Theory: A Cautionary Tale of Hedgehogs and Foxes,* 77 CHI.-KENT L. REV. 683, 690-93 (2002).

33. *Id.* at 694 (criticizing law and economics and other "top-down" theories of tort law).

34. David Dana, *Rethinking the Puzzle of Escalating Penalties for Repeat Offenders,* 110 YALE L.J. 733, 740 (2001); *see also* Richard Posner, *An Economic Approach to the Law of Evidence,* 51 STAN. L. REV. 1477, 1477 (1999) [hereinafter Posner, *Evidence*] (applying this simplifying presumption to an analysis of rules of evidence). *See generally* RICHARD A. POSNER, ECONOMIC ANALYSIS OF LAW 164 (2d ed. 1977) ("A person commits a crime because the expected benefits of the crime to him exceed the expected costs.").

35. George A. Akerlof, *Proscrastination and Obedience,* AM. ECON. REV., May 1991, at 1, 2 ("Economic theories of crime . . . are deficient and yield misleading conclusions when [psychological limitations on rational behavior] are ignored."). Economists who specialize in criminal law have tended to jettison the Chicago Man model. *See* Ronald L. Akers, *Rational Choice, Deterrence, and Social Learning in Criminology: The Path Not Taken,* 81 J. CRIM. L. & CRIMINOLOGY 653, 665 (1990) ("[R]ational choice theory [as developed in criminology] does not assume that all or even most criminal acts result from well-informed calculated choices. The

> elsewhere, a psychological explanation is much more descriptive of reality.[36] A failure to take into account psychological evidence explains why "[r]esearch on the deterrent effect of law enforcement activities shows the extremely limited value of economic analyses for policy purposes . . . ."[37]

By seeking to base policy prescriptions upon actual evidence regarding how people make decisions, rather than a simplified and consistently inaccurate model, legal decision theorists have already made many contributions to legal scholarship. In fact, the contributions have been so numerous that even tbe slightest attempt at a comprehensive summary would take this article on a detour so lengthy that I would soon lose sight of my goal, which is to respond to Professor Mitchell's specific arguments. Therefore, I offer just a few examples to introduce the uninitiated.

A mainstay of law and economics is the Coase Theorem, which provides that the initial assignment of legal rights does not determine which use will ultimately prevail because the parties will bargain to the most efficient state of affairs.[38] Unfortunately, when he wrote in 1960, Coase did not have the benefit of exposure to the literature on loss aversion and the endowment effect. Kahneman and Tversky have demonstrated that people are loss averse in that they fear losses roughly twice as much as they enjoy gains.[39] Relatedly, Kahneman

---

rational choice models in the literature leave room for all levels of rationality, except the most mindless, pathological, and irrational.").

36. In exploring why people commit cimes, I wrote:

> [T]he reality is that people usually slide into crime not as the result of a single rationally-weighed cost-benefit decision, but because of a series of small irrational decisions to experiment with drugs, join a gang, or the like. Numerous limitations on economist-defined rationality . . . (including overoptimism) . . . prevent potential criminals from acting as economists predict. One of the most significant factors in criminal behavior may well be time-delay traps. Criminal acts tend to involve short-term pleasures and benefits, but long-term costs. The criminal is unable to fully appreciate the long-term costs because of the tendency to disproportionately discount future consequences. This phenomenon causes additional years of imprisonment to carry less deterrent impact for the average person than for the hypothetical rational actor.

Robert A. Prentice, *The Case of the Irrational Auditor: A Behavioral Insight into Securities Fraud Litigation*, 95 NW. U. L. REV. 133, 177-78 (2000) (citations omitted).

37. MICHAEL R. GOTTFREDSON & TRAVIS HIRSCHI, A GENERAL THEORY OF CRIME 73, 119 (1990); *see also* Neal Kumar Katyal, *Deterrence's Difficulty*, 95 MICH. L. REV. 2385, 2412 (1997) ("[T]he standard law and economics view that reducing the probability of detection can be compensated by increasing [the sentence] may not be realistic.").

38. Coase, *supra* note 3.

39. *See* Amos Tversky & Daniel Kahneman, *Loss Aversion in Riskless Choice: A Reference-Dependent Model*, *in* CHOICES, VALUES AND FRAMES, *supra* note 8, at 143, 154; *see also* Richard Coughlan & Terry Connolly, *Predicting Affective Responses to Unexpected Outcomes*, 85 ORGANIZATIONAL BEHAV. & HUM. DECISION PROCESSES 211, 217 (2001) (finding that "losses loom

and Tversky have shown that people tend to perceive the value of items as much greater when those items become part of their endowment.[40] Therefore, people generally demand more to part with what they have than they would be willing to pay to acquire it in the first place.[41] Loss aversion and the endowment effect combine to undermine the Coase Theorem in this regard.[42] If a statute provides that employees will presumptively have certain types of benefits unless they agree to forfeit them, a much different world will result than from a regime where the statute presumes that such benefits will not be available unless the employer agrees to provide them.[43] Contrary to the Coase theorem, the initial endowment matters and it matters substantially.

Another doctrine economists commonly cite[44] derives from Mancur Olson's *The Logic of Collective Action*[45] and assumes that people, being wealth-maximizing, will, in the absence of externally imposed incentives, almost always "free ride" on the contributions of other group members. Kahan has clearly demonstrated that "Olson's

---

larger than gains" to study's subjects); Janet Landman, *Regret and Elation Following Action and Inaction: Affective Responses to Positive Versus Negative Outcomes,* 13 PERSONALITY & SOC. PSYCHOL. BULL. 524, 527 (1988) ("[W]hen people are making real decisions in betting or life dilemma situations, they weigh potential losses more heavily than potential gains.").

40. *See generally* Daniel L. Kahneman et al., *Experimental Tests of the Endowment Effect and the Coase Theorem,* 98 J. POL. ECON. 1325 (1990) (discussing the endowment effect generally); Jack L. Knetsch & J.A. Sinden, *Willingness to Pay and Compensation Demanded: Experimental Evidence of an Unexpected Disparity in Measures of Value,* 99 Q.J. ECON. 507, 512-13 (1984) (reporting results of one of the most famous studies of the endowment effect).

41. This large disparity between willingness-to-accept (WTA) and willingness-to-pay (WTP) is well documented in both laboratory and real world settings. *See* Raymond S. Hartman et al., *Consumer Rationality and the Status Quo,* 106 Q.J. ECON. 141, 158-60 (1991) (finding 3-to-1 WTP/WTA disparity in surveys of consumers of residential electrical service); Jack Knetsch, *The Endowment Effect and Evidence of Nonreversible Indifference Curves, in* CHOICES, VALUES AND FRAMES, *supra* note 8, at 171, 171 ("The minimum compensation people demand to give up a good has been found to be several times larger than the maximum amount they are willing to pay for a commensurate entitlement. For example, when questioned about the possible destruction of a duck habitat, hunters responded that they would be willing to pay an average of $247 to prevent its loss but would demand $1044 to accept it."); JOHN K. HOROWITZ & KENNETH E. MCCONNELL, A REVIEW OF WTA/WTP STUDIES 2-3 (Univ. of Md., Dep't of Agric. & Res. Econ., Oct 2000) (finding in meta study of forty-five willing to accept/willing to pay studies the average WTA/WTP ratio was approximately seven and that this held over a wide variety of experimental designs, subjects, and products), http://papers.ssrn.com/sol3/papers.cfm?abstract_id=257336.

42. Cass R. Sunstein, *Looking Forward: Behavioral Analysis of Law,* 64 U. CHI. L. REV. 1175, 1179 (1997) ("The theorem is wrong because the allocation of the legal entitlement may well matter, in the sense that those who are initially allocated an entitlement are likely to value it more than those without the legal entitlement.").

43. *See* Cass R. Sunstein, *Switching the Default Rule,* 77 N.Y.U. L. REV. 106, 112 (2002).

44. *See, e.g.,* Barry Friedman, *The Law and Economics of Federalism: Valuing Federalism,* 82 MINN. L. REV. 317, 406 n.373 (1997); Jonathan R. Macey & Geoffrey P. Miller, *Toward an Interest Group Theory of Delaware Corporate Law,* 65 TEX. L. REV. 469, 507 n.140 (1987).

45. MANCUR OLSON, THE LOGIC OF COLLECTIVE ACTION (1965).

Logic is false"[46] in that, in collective settings, people are often more influenced by their relations to others than by wealth maximization. Thus, to give just one of several examples that Kahan explores, economic theory suggests that the way to induce people to pay their taxes more regularly is to impose heavier penalties for noncompliance.[47] Yet, penalties turn out to have relatively little to do with levels of tax compliance.[48] A more nuanced psychological accounting tells us that people are more likely to pay their taxes if they believe that others are paying theirs.[49] "Auditing crack downs and other high-profile modes of enforcement risk backfiring, the evidence suggests, because they function as a *cue* that evasion is widespread."[50]

Another conventional law and economics argument is that legislatures and courts should eliminate products liability law and allow consumers to bargain for their desired level of risk, and thereby pay more if they want safer products and accept more risk if they want cheaper products.[51] Similarly, if workers are injured on the job, it is simply because they voluntarily chose to accept the risk of injury in exchange for higher wages.[52] In response to this argument, Roszkowski and I pointed out that a large number of the heuristics and biases identified by Kahneman and Tversky make it unlikely that workers rationally bargain for their accepted level of risk when taking jobs, or that consumers do so when buying products.[53]

Among these biases is the overconfidence bias. Ninety-four percent of college professors think that they are better than average

---

46. DAN M. KAHAN, THE LOGIC OF RECIPROCITY: TRUST, COLLECTIVE ACTION, AND LAW 1(Yale Law Sch., Pub. Law Research Paper No. 31, 2003), http://ssrn.com/abstract_id=361400.

47. *Id.* at 10-11.

48. *See, e.g.*, FRANK A. COWELL, CHEATING THE GOVERNMENT: THE ECONOMICS OF EVASION 74 (1990); James Andreoni et al., *Tax Compliance*, 36 J. ECON. LITERATURE 818, 855 (1998) (noting that the "most significant discrepancy that has been documented between the standard economic model of compliance and real-world compliance behavior is that the theoretical model greatly overpredicts noncompliance" and calling for incorporation of psychological factors into economic models to increase accuracy of models); Steven Klepper & Daniel Nagin, *The Criminal Deterrence Literature: Implications for Research on Taxpayer Compliance*, *in* 2 TAXPAYER COMPLIANCE 126, 142 (J. Roth & J.T. Scholz eds., 1989).

49. *See* Steven M. Sheffrin & Robert K. Triest, *Can Brute Deterrence Backfire? Perceptions and Attitudes in Taxpayer Compliance*, *in* WHY PEOPLE PAY TAXES 193, 194-95 (J. Slemrod ed., 1992).

50. KAHAN, *supra* note 46, at 16.

51. PETER W. HUBER, LIABILITY: THE LEGAL REVOLUTION AND ITS CONSEQUENCES 7 (1988).

52. *Id.* at 8.

53. Robert A. Prentice & Mark E. Roszkowski, *"Tort Reform" and the Liability "Revolution": Defending Strict Liability in Tort for Defective Products*, 27 GONZ. L. REV. 251, 286-87 (1992); Mark E. Roszkowski & Robert A. Prentice, *Reconciling Comparative Negligence and Strict Liability: A Public Policy Analysis*, 33 ST. LOUIS U. L.J. 19, 94-96 (1988).

teachers,[54] and a majority of consumers believe that they will be safer with machines than others will.[55] Whereas people thinking about marriage know that half of all couples divorce, virtually none of them think that they will.[56] This overoptimism bias also leads workers and consumers to believe that the accidents that happen to other people will not happen to them. Irrationally, people who throw dice wanting a low number tend to throw the dice softly; those who want a high number tend to throw the dice much more vigorously.[57] This illusion of control[58] leads people to conclude that their chances of avoiding injury are "inappropriately higher than the objective probability would warrant."[59] Overconfidence, overoptimism, and the illusion of control are just three reasons why it is inaccurate to characterize consumer purchases and employer-employee negotiations as involving a rational bargaining for a desired level of risk.

This survey could continue for many pages, but these examples should serve to indicate that it is at least arguable that K-T Man provides a more descriptive model of human behavior upon which to base legal policy prescriptions than does Chicago Man. People simply are not unboundedly rational. Their decision-making efficacy is often constrained by a variety of biases, guided by non-normative heuristics, and affected by non-rational factors such as emotion and altruism.

III. LIMITATIONS ON BEHAVIORAL DECISION THEORY: ARE SOCIAL SCIENTISTS SCIENTIFIC?

The essential point of Mitchell's first article is that legal decision theorists have failed to fully disclose huge flaws in the psychological research on which they rely, and thus paint what Mitchell calls an unwarrantedly pessimistic view of human decision making.[60] He suggests that legal decision theorists puff and exaggerate their portrait of human irrationality in order to produce

54. K. Patricia Cross, *Not* Can, *But* Will *College Teaching be Improved?*, 17 NEW DIR. FOR HIGHER ED.: RENEWING AND EVALUATING TEACHING, Spring 1977, at 1, 10.

55. PETER ASCH, CONSUMER SAFETY REGULATION 76 (1988).

56. Lynn A. Baker & Robert E. Emery, *When Every Relationship Is Above Average: Perceptions and Expectations of Divorce at the Time of Marriage*, 17 LAW & HUM. BEHAV. 439, 443 (1993).

57. *See* Ellen J. Langer, *The Illusion of Control*, *in* JUDGMENT UNDER UNCERTAINTY, *supra* note 8, at 231, 238.

58. MAX H. BAZERMAN, JUDGMENT IN MANAGERIAL DECISION MAKING 95-96 (4th ed. 1998) (discussing the "illusion of control").

59. ASCH, *supra* note 55, at 76.

60. Mitchell, *Pessimism*, *supra* note 12, at 2017 ("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.").

more marketable articles,[61] and that student editors at law reviews are ill-equipped to regulate such blatant flackery.[62]

Before examining Mitchell's individual indictments, it is important to concede that psychology research is not perfect and is not likely to be so any time soon. Like all other sciences, psychology studies complex phenomena through imperfect tests run by fallible human scientists. But, as Bronowski has noted, it is the checks and balances system of scientists critiquing and replicating the work of other scientists that provides the power that scientific research has to reveal truths about our world.[63] Indeed, it is "not so much the critical attitude that individual scientists have taken with respect to their own ideas that has given science its success . . . but more the fact that individual scientists have been highly motivated to demonstrate that hypotheses that are held by some other scientists are false."[64]

In his second article, Mitchell illustrates Bronowski's point by citing much of the literature aimed at pointing out the limitations of and problems with the widely accepted Kahneman and Tversky heuristics and biases research.[65] The rich literature of heuristics and biases does not consist of only twenty or thirty studies by Kahneman and Tversky, but rather includes thousands of studies seeking to confirm, rebut, and/or determine the limits and conditions of the K-T findings. These studies, and Mitchell's articles, help provide the checks and balances that Bronowski notes will help reveal the scientific truth about how people make judgments and decisions.

Not only is the psychology literature upon which legal decision theorists rely constantly tested in the psychology journals, but the legal decision theorists' own work is also constantly critiqued by articles such as Mitchell's and those of other leading legal scholars.[66]

---

61. *See id.* at 1925.

62. *Id.* at 1929-31.

63. JACOB BRONOWKSI, THE ORIGINS OF KNOWLEDGE AND IMAGINATION 112, 117-22 (1978).

64. KEITH E. STANOVICH, HOW TO THINK STRAIGHT ABOUT PSYCHOLOGY 33 (6th ed. 2001) (quoting Raymond S. Nickerson).

65. *See* MASSIMO PIATTELLI-PALMARINI, INEVITABLE ILLUSIONS: HOW MISTAKES OF REASON RULE OUR MINDS 40 (1994) (noting that the heuristics and biases discoveries "rank in importance alongside the discovery of the subconscious in psychoIanalysis").

66. Mitchell notes articles already published which critique not only the general application of psychological research to the legal arena, *see, e.g.*, Mitchell, *Pessimism, supra* note 12 at 1938-45, but also critique specific articles. *See, e.g.*, 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)); Robert J. MacCoun, *The Costs and Benefits of Letting Juries Punish Corporations: Comment on Viscusi*, 52 STAN. L. REV. 1821 (2000) (critiquing W. Kip Viscusi, *Corporate Risk Analysis: A Reckless Act?*, 52 STAN L. REV. 547, 547 (2000)); *see also* Robert A. Hillman, *The Limits of Behavioral Decision Theory in Legal Analysis: The Case of Liquidated Damages*, 85 CORNELL L. REV. 717, 737-38 (2000) (suggesting that in its

The good news, from my perspective, is that Mitchell's article shows how the debate has shifted. Instead of simply accepting the rational man assumption, many scholars now recognize its inaccuracy and seek to explore the validity and the limitations of alternative explanations of human behavior. The shift in focus from the traditional Chicago Man assumption to the behavioral literature is a laudable development. While the critiques of psychological research that Mitchell raises are substantive and important to address, they do not, as we shall see, counsel that we go back to ignoring the psychological and cognitive literature or its implications for legal analysis.[67]

A. *Does Behavioral Decision Theory Mask Individual and Situational Differences in Rational Behavior and Distort Perceptions of the Prevalence of Irrational Behavior?*

Mitchell's first attack on psychology research claims that the methodology used by psychologists allows them to claim findings that support the heuristics and biases camp, even though some or perhaps a majority of subjects in a study answered normatively.[68] His argument has several prongs.

1. Between-Subjects Designs

Mitchell begins by criticizing psychology's use of between-subjects designs[69] in empirical experiments on grounds that it tends to

---

current state of development, behavioral decision theory can offer some insights into proper judicial treatment of liquidated damage clauses, but cannot provide definite answers); Philip G. Peters, *Hindsight Bias and Tort Liability: Avoiding Premature Conclusions*, 31 ARIZ. ST. L.J. 1277, 1313-14 (1999) (concluding that behavioralists had done a service by bringing the hindsight bias to the attention of the legal community, but arguing that there are reasons to study the issues further before making any major reforms).

67. Certainly the methodology of the harder sciences is not beyond reproach either. A recent article criticized reports of clinical trials on grounds that they often fail to report absolute risk reduction, and thereby mislead readers. *See* Jim Nuovo et al., *Reporting Number Needed to Treat and Absolute Risk Reduction in Randomized Controlled Trials*, 287 JAMA 2813, 2813-14 (2002). In other words, a new drug might be reported as twice as effective as an existing drug, but the report will not note that the new drug is only 20% effective (versus 10% efficacy for the old drug).

68. Mitchell, *Pessimism*, *supra* note 12, at 1945-70.

69. In a between-subjects design, each subject is exposed to only one of the experimental treatments, so that each condition of the experiment is made up of a different group of individuals. *See* J. MERRILL CARLSMITH ET AL., METHODS OF RESEARCH IN SOCIAL PSYCHOLOGY 266 (1976). In a within-subjects design, each subject is exposed to all experimental treatments, so that each condition of the experiment is made up of the same individuals. *Id.*

As Mitchell notes,

produce insight only into statistically average decisions rather than the decisions of particular individuals.[70] In between-subjects experiments, some subjects are tested under condition A and their responses are compared to those of subjects tested under condition B. Mitchell prefers the within-subjects experimental design, where the same subject is tested under different conditions.[71] Both designs have strengths and weaknesses.

Within-subjects tests, for example, often suffer from demand effects,[72] where features of the experiment itself allow the subjects to surmise the goals of the experimenter, an occurrence that results in the skewing of the subjects' responses.[73]

> [P]articipants in a within-participants design see more than one condition and thus are in a better position to guess at the experimental hypotheses. The resulting *demand characteristics* are an important potential source of bias, as participants start wondering

---

> 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.

Mitchell, *Pessimism, supra* note 12, at 1948 n.76.

70. Mitchell, *Pessimism, supra* note 12, at 1946-54.

71. *See* CARLSMITH ET AL., *supra* note 69, at 266 (discussing how in a within-subjects design, "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").

72. Mitchell himself notes that demand characteristics are a problem that researchers must deal with, yet he criticizes use of the experimental form that is best suited to minimize its effects. *See* Mitchell, *Pessimism, supra* note 12, at 1979-984; *see also* Daniel Kahneman & Shane Frederick, *Representativeness Revisited: Attribute Substitution in Intuitive Judgment, in* PSYCHOLOGY OF INTUITIVE JUDGMENT, *supra* note 8, at 49, 70 ("The message that the [within-subjects] design conveys to the participants is that the experimenter expects to find effects of every factor that is manipulated.").

73. *See, e.g.*, Marilynn B. Brewer, *Research Design and Issues of Validity, in* HANDBOOK OF RESEARCH METHODS IN SOCIAL AND PERSONALITY PSYCHOLOGY 3, 8 (Harry T. Reis & Charles M. Judd eds., 2000) [hereinafter HANDBOOK OF RESEARCH METHODS] (explaining and illustrating demand effects); Gary Charness & David I. Levine, *When are Layoffs Acceptable? Evidence from a Quasi-Experiment,* 53 INDUS. & LAB. REL. REV. 381, 384 (2000) (noting that in a within-subjects design, "it sometimes happens that respondents, guessing that the experimenter expects certain treatments to matter, conform their responses to those perceived demands," so authors used a between-subjects design to reduce the "demand effect"); Susan Daicoff, *(Oxymoron?) Ethical Decisionmaking by Attorneys: An Empirical Study*, 48 FLA. L. REV. 197, 227 n.206 (1996) (explaining that author used between-subjects design in comparing what lawyers thought they *should* do to what they *would* do because previous studies using within-subjects design automatically alerted subjects to the possibility of a potential discrepancy between tbe two); Norman J. Finkel, *But It's Not Fair!: Commonsense Notions of Fairness,* 6 PSYCHOL. PUB. POL'Y & L. 898, 923 (2000) ("Although the within-subject design reduces more error variance than the between-subjects design, it also may introduce inadvertently a demand characteristic—making the key contrasts (i.e., unfairness vs. misfortune vs. injustice) too salient, which may suggest to the participants that they make sharper differentiations than they might have otherwise done.").

> "what are they getting at here" or "what am I supposed to do in this experiment" rather than simply performing the task.[74]

Within-subjects tests can also create carryover effects where "a participant's response in a given condition depends on conditions that participant experienced previously within the experiment."[75] Those effects can create bias by serving to anchor a subject's judgment, "create fatigue, prime specific cognitive representations, or influence participants' mood."[76]

For these and other reasons,[77] between-subjects designs "are more appropriate for the study of heuristics of judgment."[78] As mentioned above, a within-subjects model often allows subjects to play several rounds and slowly learn from their mistakes in ways that often are not available in real life.[79] "The between-subjects design in contrast, mimics the haphazard encounters in which most judgments are made and is more likely to evoke the casually intuitive mode of judgment that governs much of mental life in routine situations."[80] Mitchell concedes that others agree that within-subjects designs are often suboptimal for examining the validity of the theory that people are rational maximizers of their expected utilities.[81] As Daniel Kahneman has noted in comparing these two approaches:

> 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 . . . 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.

---

74. Eliot R. Smith, *Research Design, in* HANDBOOK OF RESEARCH METHODS, *supra* note 73, at 17, 23 (citations omitted).

75. *Id.* at 23; *see also* CARLSMITH ET AL, *supra* note 69, at 267 (discussing carryover effects).

76. Smith, *supra* note 74, at 23.

77. CARLSMITH ET AL., *supra* note 69, at 266-67 (comparing within-subjects designs with between-subjects designs).

78. Kahneman & Frederick, *supra* note 72, at 70.

79. *See* Daniel Kahneman & Amos Tversky, *On the Reality of Cognitive Illusions,* 103 PSYCHOL. REV. 582, 587 (1996) (noting that one-shot decision tests in between-subjects experiments can provide "a clean test of the hypothesis that subjects rely on a given heuristic").

In many studies heuristics survived several rounds of practice by subjects. *See, e.g.*, GEORGE LOEWENSTEIN ET AL., PAYING $1 TO LOSE $2: MISPERCEPTIONS OF THE VALUE OF INFORMATION IN PREDICTING THE PERFORMANCE OF OTHERS 26 (Carnegie Mellon Behavioral Decision Research, Working Paper No. 301, 2002) (finding that a substantial percentage of subjects remained subject to the curse of knowledge even after several rounds of experience), http://papers.ssrn.com/sol3/papers.cfm?abstract_id=323221.

80. Kahneman & Frederick, *supra* note 72, at 72-73.

81. *See* Mitchell, *Pessimism, supra* note 12, at 1949, n.78 (citing 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, 244 (1988), which argues that within-subjects tests are often not the best approach for testing utility theory, especially when subjects are presented with the same stimuli more than once).

> 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.[82]

Despite the fact that a between-subjects design is generally preferable to a within-subjects design in testing for heuristics and biases, it is not without its limitations.[83] As Mitchell points out, after testing separate groups, results are typically reported in terms of percentages and averages.[84] Mitchell suggests that this makes it possible for psychologists to report that they have found irrational behavior "*largely without documenting that any particular individuals actually acted irrationally in the experiments.*"[85]

Mitchell's point is misdirected. We know from studies of large groups (smokers versus nonsmokers) that cigarettes kill, even if the

---

82. Daniel Kahneman, *A Psychological Point of View: Violations of Rational Rules as a Diagnostic of Mental Processes,* 23 BEHAV. & BRAIN SCI. 681, 682 (2000) (emphasis added).

83. In a clever experiment, Birnbaum asked two different groups whether a number seemed large or small (on a 10-point scale). Michael H. Birnbaum, *How to Show that 9 > 221: Collect Judgments in a Between-Subjects Design,* 4 PSYCHOL. METHODS 243, 245 (1999). One group was asked about the number 9. *Id.* The question apparently evoked thoughts of small numbers and among such numbers, 9 is large. *Id.* at 246. The other group was asked about the number 221, which apparently evoked thoughts of a range of larger numbers and they rated 221 to be "smaller" on the 10-point scale than the other group had rated 9. *Id.* Although no individual concluded that 9 is larger than 221, that was the apparent result of the between-subjects design. *Id.* "The key to the result is that when judges are 'free' to choose their own contexts, they choose different contexts for different stimuli." *Id.* at 249. Birnbaum admits that there are methods of avoiding the problems that his study makes obvious, *id.* at 247-49, but it is unlikely that researchers always use them.

84. Mitchell, *Pessimism, supra* note 12, at 1946; *see* Keith E. Stanovich, *Individual Differences in Cognitive Biases: Commentary on Krueger on Social-Bias,* 9 PSYCOLOQUY 11, ¶ 7 (1998) ("[P]roponents 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)."), *at* http://psycprints.ecs.soton.ac.uk/archive/00000624/#html.

One worry with between-subjects studies is that differences between the responses of subjects under condition A and subjects under condition B may result simply from variations in how they perceive the numerical scale upon which they are to base their answers (a variation that would not arise in a within-subjects test). *See* Mitchell, *Pessimism, supra* note 12, at 1946-48 (citing 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)). However, there are mechanisms for coping with this problem, and many studies use them. *See, e.g.,* Brenda Inman Rowe, Note, *A Possible Solution for the Problem of Juries Slighting Nonscientific Evidence: A Bayesian-Like Judicial Instruction,* 24 AM. J. CRIM. L. 541, 549 (1997) ("One criticism of experiments that employ between-subjects designs is that between-subjects designs may yield results that suggest the subjects were insensitive to variations in the evidence when the results are actually due to random variation in use of numerical response scales, differences in how people attach numerical values to subjective beliefs. The present experiment defuses this criticism by using a verdict as one of the response measures.").

85. Mitchell, *Pessimism, supra* note 12, at 1946. Mitchell also argues that because between-subjects tests concentrate on averages, individual variations are ignored or minimized. *See infra* Part IV.

studies do not give us the names of individual smokers who died of cancer.[86] Broad conclusions can be established statistically without naming individual names. For example, when a particular feature was offered as part of the default automobile insurance package in Pennsylvania but not in New Jersey (where it could easily be opted into), 75% of Pennsylvania residents chose it but only 20% of New Jersey residents did so. These results illustrate a healthy status quo bias.[87] We do not need to know the names of the Pennsylvania and New Jersey residents to detect a significant impact of the status quo bias.

Furthermore, many of the heuristics and biases have been shown in studies that were not between-subject studies. Rather, they were studies where every subject was given a task or problem, and a substantial majority acted in a manner inconsistent with an objective standard of rationality. There are numerous examples of such studies, including some that demonstrate that subjects violate rational standards of dominance[88] and of intransitivity.[89] And, as noted earlier, 94% of college professors believe that they do above average work,[90] most consumers believe that they possess either average or above average ability to avoid accidents from bicycles and power mowers,[91] and between 85% and 90% of individuals surveyed believe that their future will be better than the future of an average peer.[92]

---

86. *See* Morton L. Levin, *Smoking and Cancer: Retrospective Studies and Epidemiological Evaluation*, *in* TOBACCO AND HEALTH 163, 164-66 (George James & Theodore Rosenthal eds., 1962) (reporting studies).

87. Eric J. Johnson et al., *Framing, Probability Distortions, and Insurance Decisions*, 7 J. RISK & UNCERTAINTY 35, 48 (1993). The status quo bias is an irrational preference for the current state of affairs. *See generally* William F. Samuelson & Richard Zeckhauser, *Status Quo Bias in Decision Making*, 1 J. RISK & UNCERTAINTY 7, 26-33 (1988) (reporting evidence from one of the most influential studies of this bias).

88. A principle of rational choice is dominance, in that if choice A is at least as good as choice B in every respect and better than B in at least one respect, then choice A should dominate choice B. However, when presented with certain pairs of choices, most individuals prefer B, the rationally less desirable option. *See* Daniel Kahneman & Amos Tversky, *Choices, Values, and Frames*, 39 AM. PSYCHOLOGIST 341, 344 (1984).

89. Another principle of rational choice is intransitivity, in that if Mr. X prefers A to B and B to C, then Mr. X should also prefer A to C. However, again, when faced with certain pairs of choices Mr. X will often prefer C to A. *See id.*

90. Cross, *supra* note 54, at 10. (finding that 94% of college professors rate themselves as above average teachers and fully 68% rank themselves in the top quarter of effective teachers). *But see* David Dunning et al., *Ambiguity and Self-Evaluation: The Role of Idiosyncratic Trait Definitions in Self-Serving Assessments of Ability*, *in* PSYCHOLOGY OF INTUITIVE JUDGMENT, *supra* note 8, at 324, 332-33 (suggesting that peoples' apparently over-confident self-assessments may be a product of idiosyncratic definitions of the traits under assessment).

91. ASCH, *supra* note 55, at 76.

92. David A. Armor & Shelley E. Taylor, *When Predictions Fail: The Dilemma of Unrealistic Optimism*, *in* PSYCHOLOGY OF INTUITIVE JUDGMENT, *supra* note 8, at 334, 336.

In one set of studies, ninety-one percent of subjects, including those with substantive expertise, were induced by the representativeness heuristic to commit the conjunction fallacy.[93] Ninety-five of ninety-seven professional toxicologists showed the affect heuristic in rating the benefits and risks of exposures to various chemicals.[94] Ninety-one percent of basketball fans believed that a player has a better chance of making a shot after having just made his last few shots than having just missed his last few shots, although statistical studies conclusively disprove the "hot hand" theory.[95] Such studies demonstrate mathematically the existence of these effects, biases, and cognitive limitations, even though no single individual is identified.

In his critiques, Mitchell suggests that there are too many between-subjects studies and too few within-subjects studies.[96] While conducting as many studies as possible using both methodologies is likely a good idea,[97] psychologists are aware of the competing

---

93. Amos Tversky & Daniel Kahneman, *Extensional Versus Intuitive Reasoning: The Conjunction Fallacy in Probability Judgment, in* PSYCHOLOGY OF INTUITIVE JUDGMENT, *supra* note 8, at 19, 30 ("The incidence of violations of the conjunction rule in direct tests ranged from 73% to 100%, with an average of 91%.").

When using the representativeness heuristic, people tend to judge probabilities by flouting numerous rules of statistics and focusing instead upon the degree of similarity that an item seems to bear to a category or parent population. Daniel Kahneman & Amos Tversky, *Subjective Probability: A Judgment of Representativeness, in* JUDGMENT UNDER UNCERTAINTY, *supra* note 8, at 32. If "Linda" is described with adjectives that seem to fit a common stereotype of a feminist (including being single, outspoken, and concerned with issues of discrimination and social justice), almost 90% of people queried will answer that it is more likely that Linda is a bank teller *and* active in the feminist movement than that she is simply a bank teller. Amos Tversky & Daniel Kahneman, *Judgments of and by Representativeness, in* JUDGMENT UNDER UNCERTAINTY, *supra* note 8, at 84, 91-96. The similarity of the description to the stereotype of a feminist overwhelms the statistical fact that it must be more likely that Linda is only "a" than that she is "a" *and* "b." *Id.* This statistical error is called the conjunction fallacy.

94. Paul Slovic et al., *The Affect Heuristic, in* PSYCHOLOGY OF INTUITIVE JUDGMENT, *supra* note 8, at 397, 412 (citing a 1999 survey of members of the British Toxicological Society). When they rely on the affect heuristic, people make decisions rapidly and automatically based upon feelings of "goodness" or "badness." *Id.* at 410-13. Thus, scientists who perceive that the benefits of nuclear power are high are likely to rate its risks as low, whereas those who perceive its benefits as low are likely, based on the same evidence, to rate its risks as high. *Id.* at 411-12.

95. Thomas Gilovich et al., *The Hot Hand in Basketball: On the Misperception of Random Sequences, in* PSYCHOLOGY OF INTUITIVE JUDGMENT, *supra* note 8, at 601, 602, 613; Jonathan J. Koehler & C.A. Conley, *The Hot Hand Myth in Professional Basketball,* 26 J. SPORT & EXERCISE PSYCHOL. (forthcoming 2004).

96. Mitchell, *Pessimism, supra* note 12, at 1948.

97. For example, Birnbaum and Mellers found little evidence for the base-rate fallacy (the tendency to ignore the relative frequency with which an event occurs) when they repeated classic experiments using, not the typical one problem scenario, but a scenario involving several judgments. 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, 801 (1983). Without a doubt, the supposed base-rate fallacy is one of the most controversial

strengths of between-subjects and within-subjects tests and often plan lines of research accordingly.[98] For example, in a study of counterfactual thinking,[99] Shepperd and McNulty began with a survey that examined between-subjects differences, followed with a within-subjects survey, and concluded with a study of actual behavior.[100] In a study of jury decision making in sexual harassment cases, Wiener and his colleagues performed a between-subjects experiment[101] that Schoenfelt and colleagues followed with a within-subjects study of the same question[102] (finding similar results). In short, psychologists often study the same phenomena with both between-subjects and within-subjects studies and usually (but not invariably) find similar results.[103]

---

of the heuristics and biases identified to date. *See* Jonathan J. Koehler, *The Base Rate Fallacy Reconsidered: Descriptive, Normative, and Methodological Challenges*, 19 BEHAV. & BRAIN SCI. 1, 1 (1996) (arguing that researchers "have been oversold on the base rate fallacy from an empirical, normative, and methodological standpoint").

98. *See generally* Gideon Keren, *Between- or Within-Subjects Design: A Methodological Dilemma*, *in* A HANDBOOK FOR DATA ANALYSIS IN THE BEHAVIORAL SCIENCES: METHODOLOGICAL ISSUES 257, 271 (Gideon Keren & Charles Lewis eds., 1993) [hereinafter HANDBOOK FOR DATA ANALYSIS] (discussing advantages and disadvantages in use of the different experimental designs for different purposes, and recommending that use of both designs be considered where practicable).

99. Counterfactual thinking is "what if" thinking. *See* Vittorio Girotto et al., *Event Controllability in Counterfactual Thinking*, 78 ACTA PSYCHOLOGICA 111, 112 (1991) (defining counterfactual thinking as "the mental construction of alternatives to factual events").

100. James A. Shepperd & James K. McNulty, *The Affective Consequences of Expected and Unexpected Outcomes*, 13 PSYCHOL. SCI. 85, 86 (2002). (concluding that "[b]ad outcomes felt worse when unexpected than when expected, whereas good outcomes felt better when unexpected than when expected").

101. Richard L. Wiener et al., *Social Analytic Investigation of Hostile Work Environments: A Test of the Reasonable Woman Standard*, 19 LAW & HUM. BEHAV. 263, 266-67 (1995) (finding in a between-subjects study that differences in stated judicial standard did not affect mock jurors' findings of sexual harassment).

102. Elizabeth L. Shoenfelt et al., *Reasonable Person Versus Reasonable Woman: Does It Matter?*, 10 AM. U. J. GENDER SOC. POL'Y & L. 633, 666 (2002) (finding in a within-subjects study that differences in the stated judicial standard did not affect mock jurors' findings of sexual harassment).

103. *See, e.g.*, CARLSMITH ET AL., *supra* note 69, at 269-70 (giving examples of studies combining between-subjects designs with within-subjects designs); Norman J. Finkel & Jennifer L. Groscup, *When Mistakes Happen: Commonsense Rules of Culpability*, 3 PSYCHOL. PUB. POL'Y & L. 65, 93 (1997) (in study of how people assign culpability, finding "general consistency whether we tested using the between-subjects design or the within-subject design"); Peter K. Isquith et al., *Blaming the Child: Attribution of Responsibility to Victims of Child Sexual Abuse*, *in* CHILD VICTIMS, CHILD WITNESSES 203, 204-05 (Gail S. Goodman & Bette L. Bottoms eds., 1993) (finding in both within-subjects and between-subjects designs that subjects were more likely to assess some causality to older victims of child molestation than to younger victims); Kahneman & Tversky, *supra* note 79, at 582 (studying the conjunction effect with between-subjects and within-subject studies and finding little effect in within-subjects studies which they attribute to the fact that participants are more likely to detect set inclusion in that design).

In studies where within-subjects experiments do yield different results than between-subjects experiments, it is often because subjects can learn from mistakes if their errors are pointed out to them or they get several chances at the task.[104] These within-subjects experiments must be conducted, because it is important to know which heuristics and biases can be minimized by repeated trials. However, it is also important to remember that in the real world decision makers often do not have anyone to correct their errors and have only once chance to make a rational decision.

2. Null Hypothesis Significance Testing

Mitchell's second attack on psychology research focuses on the weaknesses of its statistical methodology.[105] Statistical analysis is very important in the psychology discipline.[106] As in the field of psychology in general, most studies in the heuristics and biases vein use null hypothesis significance testing (NHST) in their research design.[107] Mitchell points out that many people believe that use of NHST has a tendency to exaggerate the irrationality of study subjects. For example, Reid Hastie writes:

> 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

---

104. Mitchell, *Pessimism, supra* note 12, at 1951 n.81 ("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.").

105. *Id.* at 1954-65.

106. Zeno G. Swijtink, *A Plea for Popperian Significance Testing,* 21 BEHAV. & BRAIN SCI. 220, 220 (1998) ("Of all the human sciences, psychology must have the closest interest in statistics. This is no doubt because psychology has enough experimental control to stabilize variability, but not enough control to eliminate variability altogether.").

107. According to De Long and Lang:

> In classical hypothesis testing, a null hypothesis is posed against an alternative, and the null hypothesis is considered "rejected" or "not rejected" on the basis of whether a single test statistic exceeds some critical value (e.g., whether a large-sample *t*-statistic exceeds 1.96) . . . . [I]f the null is "rejected," our confidence in it is reduced; if the null hypothesis "fails to be rejected," our confidence in the correctness of the null hypothesis is increased because the data do not speak strongly against it.

J. Bradford De Long & Kevin Lang, *Are All Economic Hypotheses False?,* 100 J. POL. ECON. 1257, 1257-58 (1992).

The purpose of NHST is "to provide a procedure for deciding whether the probability of getting sample results as extreme or more so than the null hypothesized value was small enough that it was less likely that it could be attributed to mere chance." Lisa Harlow, *Significance Testing Introduction and Overview, in* WHAT IF THERE WERE NO SIGNIFICANCE TESTS? 1, 1-2 (Lisa L. Harlow et al. eds., 1997).

null hypothesis has yielded a large harvest of "significant," but unimportant "proofs" that humans are irrational.[108]

This argument that a null hypothesis is always judged false is itself wrong,[109] and, more importantly, misses the point.[110] Researchers in psychology are well aware that a large N (number of subjects) increases the power of a hypothesis test, making it easier to reject a false null hypothesis. However, they are also aware that others will call into question the scientific importance of tiny differences between populations, differences that are detectable only with very large sample sizes. The effects reported in most psychology journals, and the effects that Kahneman and Tversky have generally reported, are not miniscule.[111]

---

108. 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). Some critics are even more vocal. *See* 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) (arguing that NHST "is a terrible mistake, is basically unsound, poor scientific strategy, and one of the worst things that ever happened in the history of psychology").

109. *See e.g.*, William F. Oakes, *On the Alleged Falsity of the Null Hypothesis*, 25 PSYCHOL. REC. 265, 265 (1975) (pointing to a federal government study with *23,000* subjects that failed to disprove the null hypothesis). Oakes made several other arguments, concluding "it shouldn't be assumed that the null hypothesis is generally false in an experiment." *Id.* at 272; *see also* Robert W. Frick, *Accepting the Null Hypothesis*, 23 MEMORY & COGNITION 132, 132 (1995) (giving examples to illustrate his argument that "the null hypothesis *is* possibly correct, as examples easily demonstrate"); Richard L. Hagen, *In Praise of the Null Hypothesis Statistical Test*, 52 AM. PSYCHOLOGIST 15, 21 (1997) (arguing that the claim that the null hypothesis is always false "has never been sustained by either statistical or logical arguments"); Joseph S. Rossi, *Meta-analysis, Power Analysis, and the Null-Hypothesis Significance-Test Procedure*, 21 BEHAV. & BRAIN SCI. 216, 216 (1998) ("[A]s I am frequently engaged in the conduct of large randomized clinical trials of behavioral interventions for health promotion and disease prevention, I could only wish that the null hypothesis was, in fact, never true!").

110. As Mulaik points out:

> The point is that it doesn't matter if the null hypothesis is always judged false at some sample size, as long as we regard this as an empirical phenomenon. What matters is whether *at the sample size we have* we can distinguish observed deviations from our hypothesized values to be sufficiently large and improbable under a hypothesis of chance that we can treat them reasonably but provisionally as not due to chance error. There is no a priori reason to believe that one will always reject the null hypothesis at any given sample size. On the other hand, accepting the null hypothesis does not mean the hypothesized value is true, but rather than the evidence observed is not distinguishable from what we would regard as due to chance if the null hypothesis were true and thus is not sufficient to disprove it. The remaining uncertainty regarding the truth of our null hypothesis is measured by the width of the region of acceptance or a function of the standard error. And this will be closely related to the power of the test, which also provides us with information about our uncertainty.

Stanley A. Mulaik et al., *There Is a Time and a Place for Significance Testing*, *in* WHAT IF THERE WERE NO SIGNIFICANCE TESTS?, *supra* note 107, at 65, 80-81.

111. In some of the famous studies in the K-T tradition, the researchers did not even bother to run significance tests because the proportional differences were so dramatic. My thanks to Jonathan Koehler for pointing this out to me.

Hastie's other point, that statistical significance does not automatically translate into practical significance, is well taken. However, many believe that the error rates NHST generates are typically not out of line with reality.[112] Mitchell argues that behavioral decision theory must be moved into the real world in order to focus on the actual impact of cognitive biases on behavior,[113] but as noted elsewhere in this article, many of the heuristics and biases noted in laboratory experiments have also been identified in the real world and can carry very real practical implications.[114]

Not only are laboratory findings often replicated in studies in the outside world, observed phenomena in the outside world are usually the inspiration for laboratory studies that can be performed under controlled conditions. For example, researchers noted that people in the real world make predictions that tend to resemble too strongly the previous year's results.[115] In the real world, it is generally impossible to control for other factors that might be causing the phenomenon. By setting up laboratory conditions to control for these other factors, researchers were able to identify the anchoring and adjustment phenomenon, and then study its features in a number of follow-up studies.[116] Others then exported those findings to the real world and found similar results in real world decision making.[117]

Mitchell also argues that one shortcoming of NHST is that "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

---

112. W.K. Estes, *Significance Testing in Psychological Research: Some Persisting Issues*, 8 PSYCHOL. SCI. 18, 18 (1997) ("Are error rates out of line with reality? Historically, most attention has been given to Type I errors (inappropriate rejections of null hypotheses) because the primary purpose of significance testing is avoidance of excessive rates of those errors. True rates can never be known for empirical situations, but for simulated data, rates specified in significance tests have been found to be quite well approximated even when there are appreciable deviations from the assumptions of the statistical model.").

113. Mitchell, *Pessimism*, *supra* note 12, at 1959.

114. *See infra* notes 169-195 and accompanying text.

115. I once suggested that the anchoring and adjustment phenomenon—the tendency to anchor upon a given number and then adjust insufficiently to new information—underlay the typical "same as last year" plan for audits. Prentice, *supra* note 36, at 165.

116. *See generally* Jeffery R. Boyll, *Psychological, Cognitive, Personality, and Interpersonal Factors in Jury Verdicts*, 15 LAW & PSYCHOL. REV. 163, 170 (1991) (citing studies indicating that juries anchor on *ad damnum* clauses in making damage awards); Amos Tversky & Daniel Kahneman, *Judgment Under Uncertainty: Heuristics and Biases*, *in* JUDGMENT UNDER UNCERTAINTY, *supra* note 8, at 3, 14-18 (reporting early study).

117. *See, e.g.*, Jean C. Bedard, *An Archival Investigation of Audit Program Planning*, AUDITING: J. PRAC. & THEORY, Fall 1989, at 57, 57 (finding evidence of anchoring and adjusting by auditors in study of workpapers of actual audits).

reject, the null hypothesis."[118] Mitchell cites as an example a study where inconsistent conduct by only 15% of the subjects led to a rejection of the null hypothesis of rational decision making.[119] Nonetheless, as also noted earlier, many of the landmark studies in the heuristics and biases literature involve "irrational" actions not by 15% of the subjects (although errors by 15% of a population could be very important and have a large practical effect), but often by 85% or 90% of the population.[120]

Surely the NHST has many severe critics;[121] even its defenders concede its limitations,[122] and the psychology discipline is continually searching for methods to improve, supplement,[123] or replace it. For example, several critics of NHST have supported greater use of confidence intervals,[124] model fitting,[125] and meta-analyses,[126] although these statistical methods have their own limitations.[127] The

118. Mitchell, *Pessimism, supra* note 12, at 1955.

119. *Id.* at 1951-52, 1955 (citing Dawes, *supra* note 13, at 503-04).

120. *See supra* notes 90-95 and accompanying text.

121. *See, e.g.*, John E. Hunter, *Needed: A Ban on the Significance Test,* 8 PSYCHOL. SCI. 3, 6 (1997) (proposing that NHST be abandoned); Roger E. Kirk, *Practical Significance: A Concept Whose Time Has Come,* 56 EDUC. & PSYCHOL. MEASUREMENT 746 (1996) (arguing that scientists should replace the objective, mechanical NHST with more subjective assessments of practical, rather than statistical, significance); Leonard G. Rorer, *Some Myths of Science in Psychology, in* 2 THINKING CLEARLY ABOUT PSYCHOLOGY 61, 61 (D. Cicchetti & W.M. Grove eds., 1991) (arguing that both theory testing and null-hypothesis significance testing should be abandoned in favor of Bayesian formulations).

122. Kathleen M. Dillon, *I Am 95% Confident That the Earth Is Round: An Interview About Statistics with Chris Spatz,* 26 TEACHING PSYCHOL. 232, 232 (1999) (quoting psychology Professor Chris Spatz as stating that all defenders of the NHST admit that objectors have several valid points).

123. *See, e.g.*, Richard J. Harris, *Reforming Significance Testing via Three-Valued Logic, in* WHAT IF THERE WERE NO SIGNIFICANCE TESTS?, *supra* note 107, at 145, 171 (recommending that NHST be used to test three-alternatives rather than the typical two-alternative presentation currently used).

124. *See, e.g.*, De Long & Lang, *supra* note 107, at 1269 (suggesting that economists rely less on NHST and more on confidence intervals); Hunter, *supra* note 121, at 6 (suggesting that social sciences use confidence intervals, the dominant technique used in the quantitative sciences).

125. *See, e.g.*, Michael M. Granaas, *Model Fitting: A Better Approach,* 53 AM. PSYCHOLOGIST 800, 800-01 (1998) (suggesting this method as preferable to NHST).

126. *See, e.g.*, Mark W. Lipsey & David B. Wilson, *The Efficacy of Psychological, Educational, and Behavioral Treatment: Confirmation from Meta-analysis,* 48 AM. PSYCHOLOGIST 1181, 1200 (1993) (performing a meta-analysis and touting its advantages); Frank L. Schmidt, *What Do Data Really Mean?,* 47 AM. PSYCHOLOGIST 1173, 1180 (1992) (suggesting that meta-analyses can remedy the problems critics perceive with NHST).

127. *See, e.g.*, Robert P. Abelson, *On the Surprising Longevity of Flogged Horses: Why There Is a Case for the Significance Test,* 8 PSYCHOL. SCI. 12, 12 (1997) (arguing that "all the foolishness associated with the null hypothesis might also infect confidence limits"); Siu L. Chow, *A Precis of "Statistical Significance: Rationale, Validity, and Utility,"* 21 BEHAV. & BRAIN SCI. 169, 190 (1998) (noting several difficulties with meta-analysis, including especially a lack of commensurability among the studies included in the meta-analysis); Domenic V. Cicchetti, *Role of Null Hypothesis Significance Testing (NHST) in the Design of Neuropsychologic Research,* 20

debate in psychology is mirrored in economics and other fields that often use NHST.[128]

Notwithstanding this debate, NHST is by far the most widely used statistical tool in psychology.[129] NHST is a valuable tool that many respected researchers strongly support,[130] although it is one

---

J. CLINICAL & EXPERIMENTAL NEUROPSYCHOLOGY 293, 294 (1998) (quoting leading NHST critic Jacob Cohen as advising, "don't look for a magic alternative to NHST, some other objective mechanical ritual to replace it. It doesn't exist."); Edward Erwin, *The Logic of Null Hypothesis Testing,* 21 BEHAV. & BRAIN SCI. 197, 198 (1998) (arguing that epistemological reviews of meta-analytic reviews have found them to be flawed); Richard J. Harris, *Significance Tests Have Their Plaee,* 8 AM. PSYCHOLOGIST 8, 9-10 (1997) (noting that confidence intervals often demonstrate the same limitations as NHST and arguing that meta-analyses also have more limitations than generally recognized).

128. *See, e.g.*, De Long & Lang, *supra* note 107, at 1257 (examining use of the null hypothesis in economics); Jonathan A.C. Sterne & George D. Smith, *Sifting the Evidence—What's Wrong with Significance Tests?,* 322 BRIT. MED. J. 226, 226 (2001) (criticizing use of NHST in medical research); *see also* De Long & Lang, *supra* note 107, at 1258 n.1 (noting that any "distinction between the 'standard' approach to testing in 'science' and that used in economics should not be exaggerated"); Stanley A. Mulaik et al., *supra* note 110, at 94 ("[P]hysicists *do* use procedures that are comparable to significance tests.").

129. Raymond S. Nickerson, *Null Hypothesis Significance Testing: A Review of an Old and Continuing Controversy,* 5 PSYCHOL. METHODS 241, 241 (2000) (noting that NHST "is arguably the most widely used method of analysis of data collected in psychological experiments and has been so for about 70 years").

130. *See, e.g.*, Abelson, *supra* note 127, at 14 ("Null-hypothesis tests are cogent in scrutinizing surprising results that critics doubt."); Robert P. Abelson, *A Retrospective on the Significance Test Ban of 1999 (If There Were No Significance Tests, They Would Be Invented), in* WHAT IF THERE WERE NO SIGNIFICANCE TESTS?, *supra* note 107, at 117, 129 ("Realistically, if the null hypothesis test did not exist, it would have to be (re)invented."); Galen L. Baril & J. Timothy Cannon, *What Is the Probability That Null Hypothesis Testing Is Meaningless?,* 50 AM. PSYCHOLOGIST 1098, 1099 (1995) (disputing major arguments against use of NHST); Chow, *supra* note 127, at 170 (arguing that "the resiliency of [NHST] is warranted" and criticisms against it are "debatable"); Jose M. Cortina & William P. Dunlap, *On the Logic and Purpose of Significance Testing,* 2 PSYCHOL. METHODS 161, 170 (1997) (noting that "the arguments against the use of NHST are built on faulty premises, misleading examples, and misunderstanding of certain critical concepts" and that "there are many cases in which drawing conclusions about hypotheses based on *p* values is perfectly reasonable"); Robert W. Frick, *Chow's Defense of Null-Hypothesis Testing: Too Traditional?,* 21 BEHAV. & BRAIN SCI. 199, 199 (1998) (agreeing with Chow that NHST "plays an essential and irreplaceable role in science"); Anthony G. Greenwald et al., *Effect Sizes and* p *Values: What Should Be Reported and What Should Be Replicated?,* 33 PSYCHOPHYSIOLOGY 175, 182 (1996) (noting that despite its limitations, NHST continues to be widely used because of its value in providing results in the form of a dichotomous hypothesis evaluation and providing an index that has (*p* values) that is informative and indicative regarding the likelihood of replicability); Hagen, *supra* note 109, at 22 (arguing that the NHST has been "unfairly maligned," and claiming that "[t]he logic of the NHST is elegant, extraordinarily creative, and deeply embedded in our methods of statistical inference"); Harlow, *supra* note 107, at 11 (noting that when properly used and supplemented, the NHST "can be very effective in highlighting hypotheses that are worthy of further investigation, as well as those that do not merit such efforts"); Harris, *supra* note 127, at 8 ("[A]s applied by most researchers and journal editors, NHST provides a very useful form of social control over researchers' understandable tendency to squander analytic effort 'explaining' effects whose sign in a given sample may not match the sign of the corresponding population effect."); John F. Kihlstrom, *If You've Got an Effect, Test Its Significance; If You've Got a Weak Effect, Do a Meta-analysis,* 21

with limitations that researchers must keep in mind. If they do not do so, journal editors will likely refresh their memories.[131]

Mitchell is right to remind psychologists of NHST's limitations. It can provide confidence that results did not stem from sampling errors; it cannot prove the validity of a theory.[132] But with or without Mitchell's reminder, psychologists already know that NHST is only a tool; a *p* of .05 is not the Holy Grail, and a *p* of .06 is still quite interesting even if it falls short of the traditional .05 threshhold of statistical significance. They know this just as baseball managers know that a .300 batting average is a magical number, but if the player can't run or field you'd be better off signing a shortstop with range who hits .294.

BEHAV. & BRAIN SCI. 205, 206 (1998) (concluding that significance tests "constitute a principled way for researchers to claim that their experimental results are worth knowing about, and for consumers to evaluate researchers' claims" and that "significance testing has kept the behavioral, cognitive, and social sciences from lapsing into solipsism"); Lester E. Krueger, *The Ego Has Landed! The .05 Level of Statistical Significance is Soft (Fisher) Rather than Hard (Neyman/Pearson),* 21 BEHAV. & BRAIN SCI. 207, 207 (1998) (arguing that NHST "has a highly circumscribed yet vital role as the initial gatekeeper in scientific research"); Joel R. Levin, *Statistical Significance Testing from Three Perspectives,* 61 J. EXPERIMENTAL EDUC. 378, 382 (1993) (arguing that researchers fully understand NHST's limitations and that alleged problems with NHST are "more apparent than real"); Mulaik et al., *supra* note 110, at 73 (noting that whatever its limitations, "significance testing contributes to the cumulative research enterprise in allowing one to assess whether differences from predicted values under an integrative hypothesis are more reasonably regarded as due to random measurement errors and sampling errors or not"); Nickerson, *supra* note 129, at 291 (noting that NHST can be applied inappropriately, but "can also be an effective aid to data interpretation when used appropriately as an adjunct to good experimental design and in conjunction with other methods of extracting information from noisy data"); Swijtink, *supra* note 106, at 221 ("[T]here remains an important role for significance testing . . . ."); John R. Vokey, *Statistics Without Probability: Significance Testing as Typicality and Exchangeability in Data Analysis,* 21 BEHAV. & BRAIN SCI. 225, 225 (1998) (noting that many criticisms of NHST are "either irrelevant or misplaced").

131. Estes notes:

> In the course of some 20 years of editing psychological journals, I found reports of significance levels and effect sizes to be useful aids in the task of screening out from an enormous input of manuscripts those whose results were not likely to prove robust or replicable. However, the use of these indicators by me and my consultants was not mechanical or constrained by rigid criteria. When results of a study were accompanied by recommendations for changes of public policy (not an infrequent occurrence in the case of *Psychological Science*), we required significance levels to be stricter than the norm and effects sizes larger. But when studies involved large amounts of data collected on very few individuals, often from special populations (an increasingly common occurrence in research on long-term memory in natural environments, extremely deviant abilities, and effects of specific kinds of brain damage on mental functions), we often advised contributors to dispense with reports of statistical tests and concentrate on other kinds of evidence bearing on the soundness of conclusions

Estes, *supra* note 112, at 19.

132. Robert W. Frick, *The Appropriate Use of Null Hypothesis Testing,* 1 PSYCHOL. METHODS 379, 380 (1996) ("It is well agreed that null hypothesis testing by itself does not provide sufficient evidence for accepting the null hypothesis.").

Ultimately, we must rely on the research design and execution of the psychologists. Fortunately, perhaps the most extreme critic of NHST, Professor Hunter, reminds us:

> I have served on hundreds of graduate student committees, I am close friends with several hundred other researchers, and I have reviewed hundreds of manuscripts for publication. Every person that I have ever known worked hard to make his or her study the best study it could be. Although scientists do make errors, they work very hard and very intelligently at their research. There are almost no "garbage studies."[133]

Mitchell's discussion of the limits of NHST leads him to tout the benefits of meta-analyses.[134] Beginning with framing effects,[135] Mitchell argues that meta-analyses by Kuhberger minimize the impact of such effects.[136] One of Kuhberger's studies concludes that research has done just what Mitchell called for—taken the concept outside the laboratory. According to Kuhberger, "framing research has stepped outside the lab to a considerable degree."[137] Furthermore, Kuhberger reports, "experts are also influenced by framing, but maybe to a lesser degree than students."[138] After surveying 136 studies involving 30,000 participants, Kuhberger concludes that "framing is a phenomenon now in its teenage years,"[139] althougb the effect, as Mitchell points out, is in the small to moderate range in most studies.[140]

In another meta-analysis, Kuhberger, writing with colleagues, found that Kahneman and Tversky's prospect theory (with attendant

---

133. Hunter, *supra* note 121, at 4 (using this point to argue that the variation in results among studies performed in psychology is not due to the fact that some studies are well designed and others are poorly designed); *see also* Chow, *supra* note 127, at 178 (stating that "experimental psychologists are meticulous about the internal validity of experiments" and "are aware that a statistically significant result may be ambiguous at the conceptual level as a result of various features found in the data collection procedure or situation").

134. Mitchell, *Pessimism, supra* note 12, at 1959-60.

135. Framing effects are a family of complex effects, but the essential notion is that, contrary to the Chicago Man model, people's preferences for risk and other choices are not invariant; they often change with how a problem is presented or framed. *See generally* Amos Tversky & Daniel Kahneman, *The Framing of Decisions and the Psychology of Choice,* 211 SCI. 453 (1981). Thus, in one scenario a decision maker might choose Alternative A over Alternative B, whereas in another scenario the same decision maker might choose a mathematically identical version of Alternative B over a mathematically identical version of Alternative A. *See id.* at 455-56. Whether the alternatives are framed as gains or losses can make a definitive difference. *Id.* at 456. *See generally* SCOTT PLOUS, THE PSYCHOLOGY OF JUDGMENT AND DECISION MAKING 69-76 (1993) (providing an accessible explanation of the basics of framing effects).

136. *See* Anton Kuhberger, *The Influence of Framing on Risky Decisions: A Meta-analysis,* 75 ORGANIZATIONAL BEHAV. & HUM. DECISION PROCESSES 23, 42, 47 (1998) (discussing "small to moderate size" framing effects across 136 studies); *see also* Mitchell, *Pessimism, supra* note 12, at 1960-63.

137. *Id.* at 45.

138. *Id.*

139. *Id.* at 47.

140. Mitchell, *Pessimism, supra* note 12, at 1962.

framing effects) was "supported by our meta-analysis."[141] Regarding this study, Mitchell points out that on average only 60% of the subjects in studies are impacted by framing effects,[142] but this is more than enough to have a significant impact in the real world.[143]

Mitchell also argues that meta-analyses of hindsight bias studies demonstrate that the bias has relatively small effects.[144] Indeed, the study he cites looks at 122 studies, then finds that the hindsight bias clearly exists,[145] but that it is generally of relatively small magnitude.[146] The authors go on to point out that "[t]his 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."[147] If a jury in a medical malpractice case is split nearly equally regarding the foreseeability to the physician-defendant of the plaintiff-patient's complications, a small amount of hindsight bias could have a significant effect.[148] More meta-analyses are always a good idea, but those done so far do little to minimize the standard characterization of K-T Man.

3. The File Drawer Problem

Mitchell's next point is that journals are reluctant to publish, and therefore authors are reluctant to submit for publication, studies

---

141. Anton Kuhberger et al., *The Effects of Framing, Reflection, Probability, and Payoff on Risk Preference in Choice Tasks*, 78 ORGANIZATIONAL BEHAV. & HUM. DECISION PROCESSES 204, 216 (1999).

142. Mitchell, *Pessimism*, *supra* note 12, at 1962 n.101.

143. Kuhberger et al., *supra* note 141, at 219 (arguing that their study found "*significant* bidirectional framing effects" (emphasis added)). If, by using framing effects, sellers of products can impact the decisions of 60% of consumers or political candidates can impact the decisions of 60% of voters, framing can have a major impact on real world decision making.

144. *See* Jay J.J. Christensen-Szalanski & Cynthia F. Willham, *The Hindsight Bias: A Meta-analysis*, 48 ORGANIZATIONAL BEHAV. & HUM. DECISION PROCESSES 147, 162 (1991) ("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.").

145. *Id.* at 154.

146. *Id.* at 153-54 ("The average weighted effect size of all 122 studies was $r = .17$ with a 95% confidence interval . . . .").

147. *Id.* at 162 (going on to qualify this statement with the warning that we should be careful before issuing warnings about the hindsight bias because "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 world than would have occurred had the effect size been larger").

148. *See generally id.* at 158 (noting that if the threshold probability for choosing a particular alternative was 80%, "[i]f in foresight a person estimated the event to be 78% and in hindsight estimated it to be 81%, then *even though the effect of the bias on probability assessments was small, it would still be of practical importance* since it resulted in the person making a different decision").

that show insignificant results.[149] Therefore, many harbor the lurking suspicion that the published studies showing biased human judgment may be silently contradicted by unpublished studies showing no such bias that are sitting in file drawers in psychology professors' offices. As Rosenthal notes, the fear 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."[150]

The file drawer problem is a legitimate concern in disciplines as diverse as marketing[151] and oncology.[152] Fortunately, there is little firm evidence that the file drawer problem is significant.[153] As

149. Mitchell, *Pessimism, supra* note 12, at 1966-67.

150. Robert Rosenthal, *Cumulating Evidence, in* HANDBOOK FOR DATA ANALYSIS, *supra* note 98, at 519, 535. A Type I error (a false positive) is the rejection of a true null hypothesis, while a Type II error (a false negative) is the failure to reject a false null hypothesis. *See* Vincent Bauchau, *Is There a "File Drawer Problem" in Biological Research?*, 79 OIKOS 407, 408 (1997) (finding "no strong evidence so far of the existence of a file drawer problem in biology").

151. Raymond Hubbard & J. Scott Armstrong, *Are Null Results Becoming an Endangered Species in Marketing?*, 3 MARKETING LETTERS 127, 134 (1992) (finding in a review of marketing journals that few studies failing to reject the null hypothesis are published).

152. *See* Jesse A. Berlin et al., *An Assessment of Publication Bias Using a Sample of Published Clinical Trials*, 84 J. AM. STATISTICAL ASS'N 381, 391 (1989) (finding grounds to worry about publication bias in research on cancer treatment); Robert J. Simes, *Publication Bias: The Case for an International Registry of Clinical Trials*, 4 J. CLINICAL ONCOLOGY 1529, 1538-39 (1986) (finding that pooled results of published trials showed statistically significant benefits for a certain treatment, but pooled results of published and unpublished trials together did not). *See also* Mathias Egger et al., *Bias in Meta-analysis Detected by a Simple, Graphical Test*, 315 BRIT. MED. J. 629 (1997) (addressing potential biases in medical research, including nonpublication of negative trials).

153. As Rosenthal suggests

> [R]ecent research suggests that the magnitude of the file drawer problem may be somewhat less than had been feared. Although studies published at the time of a meta-analysis are more likely to yield significant results than are studies unpublished at the time of the meta-analysis, this bias may well shrink over time because a very large proportion of the originally unpublished studies may eventually be published. In a large meta-analysis, therefore, it may be useful to conduct a subanalysis with a cut-off date for study retrieval approximately 5 years earlier than the date of the actual meta-analysis. It is likely that the file drawer problem will be lessened appreciably at least for this subanalysis.

Rosenthal, *supra* note 150, at 537-38 (citations omitted); *see also* Jeff Gill & Kenneth J. Meier, *Public Administration Research and Practice: A Methodological Manifesto*, 10 J. PUB. ADMIN. RES. & THEORY 157, 167 (2000) (noting that data mining is related to the file drawer problem and that "there is evidence that the file drawer problem is not pervasive"); Nickerson, *supra* note 129, at 270-71 (doubting that the file drawer problem leads to an understatement of the probability of reporting chance effects as real); Harris Cooper, *Finding the Missing Science*, 30 APA MONITOR ONLINE (Sept. 1999) (noting that psychology researchers "are smart enough to spend little time studying pure chance phenomena [so g]enerally, the results that appear in our journals are indices of real, systematic relationships"), *at* http://www.apa.org/monitor/sep99/scispeak.html. *See generally* David Neumark & William Wascher, *Is the Time-Series Evidence on Minimum Wage Effects Contaminated by Publication Bias?*, 36 ECON. INQUIRY 458 (1998) (using sophisticated statistical approach to argue that file drawer and related publication biases do not undermine conclusions of published time-series studies of minimum wage effects).

Rosenthal, who has studied this problem more than any other scholar, points out:

> In the past there was very little we could do to assess the net effect of studies tucked away in file drawers that did not make the magic .05 level. Now, however, although no definitive solution to the problem is available, we can establish reasonable boundaries on the problem and estimate the degree of damage to any research conclusion that could be done by the file drawer problem. The fundamental idea in coping witb the file drawer problem is simply to calculate the number of studies averaging null results that must be in the file drawers before the overall probability of a Type I error can be just brought to any desired level of significance, say $p = .05$. This number of filed studies, or the tolerance for future null results, is then evaluated for whether such a tolerance level is small enough to threaten the overall conclusion drawn by the reviewer. If the overall level of significance of the research review will be brought down to the level of *just significant* by the addition of just a few more null results, the finding is not resistant to the file drawer threat.[154]

In a meta-analysis of framing studies, Kuhberger studied 136 papers and calculated that "66,388 studies finding null results would have to exist [in file drawers] somewhere before the overall results could reasonably be ascribed to sampling bias. This is not plausible."[155] Rosenthal found that 65,123 studies averaging null results would be needed to conclude that the 345 published studies examining the effects of interpersonal self-fulfilling prophecies were possibly due to sampling bias.[156]

Many of the important heuristics and biases in the Kahneman and Tverksy tradition have been demonstrated in literally hundreds of published studies.[157] Gilovich and Griffin recently observed that the major biases uncovered by Kahneman, Tverksy, and others, such as the availability bias, anchoring, the conjunction fallacy, and others "have all been demonstrated in countless contexts and with varied paradigms and dependent measures, and with domain experts as well as student volunteers."[158]

---

154. Rosenthal, *supra* note 150, at 535-36 (citations omitted).

155. Kuhberger, *supra* note 136, at 42.

156. Rosenthal, *supra* note 150, at 537; *see also* Robert Rosentbal & Donald B. Rubin, *Interpersonal Expectancy Effects: The First 345 Studies,* 7 BEHAV. & BRAIN SCI. 377, 381 (1978) (making the same point).

157. *See, e.g.*, Armor & Taylor, *supra* note 92, at 336 (noting regarding the overoptimism bias that "[r]esults from hundreds of empirical investigations have shown that, on average, people tend to view themselves as more likely to experience positive outcomes, and less likely to experience negative ones, than the average members of the group from which they have been drawn").

158. Thomas Gilovich & Dale Griffin, *Introduction—Heuristics and Biases: Then and Now, in* PSYCHOLOGY OF INTUITIVE JUDGMENT, *supra* note 8, at 1, 12. I have replicated these experiments and their results in classroom demonstrations, as have hundreds of psychology professors. It seems unlikely that the file drawer problem is significant for the major tenets of the heuristics and biases literature.

Rosenthal's "fail-safe file-drawer" (FSFD) method is controversial and Scargle recently published an interesting critique.[159] Still, there is little more than conjecture to support the notion that countless debunking studies are sitting in file drawers in psychology professors' offices. In addition, this seems unlikely in light of the many, many studies published to support most of the basic heuristics and biases discovered by Kahneman, Tversky, and their followers.

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

Next, Mitchell argues that certain features of psychological research increase the likelihood that subjects will give non-normative or "irrational" responses.[160] This argument also has several prongs.

1. Experiments Designed To Elicit Non-normative Responses

Initially Mitchell claims that psychological experiments too often produce results that have little relevance to the real world.[161] In other words, they lack "psychological realism," which refers to how well they capture thinking processes that occur in everyday life.[162] Psychologists do, of course, worry about achieving psychological realism, as opposed to "mundane realism" (replicating in the research setting the events of every day life).[163] And certainly laboratories are not the real world,[164] but as Brewer points out:

> Laboratory experiments are inherently artificial in the sense that causal variables are isolated from their normal contextual variation. This isolation and control is the essence of testing causal hypotheses with a high degree of internal validity. . . . [I]solation does not necessarily jeopardize external validity if the experimental situation

159. Jeffrey D. Scargle, *Publication Bias: The "File-Drawer" Problem in Scientific Inference*, 14 J. SCI. EXPLORATION 91, 102 (2000) (criticizing FSFD "because it treats the inherently biased file drawer as unbiased and gives grossly wrong estimates of the size of the file drawer").

160. Mitchell, *Pessimism*, *supra* note 12, at 1971-95.

161. *Id.* at 1971-77.

162. ELLIOT ARONSON ET AL., SOCIAL PSYCHOLOGY 53 (2d ed. 1997).

163. *See* Brewer, *supra* note 73, at 12 ("An experimental setting may have little mundane realism but still capture processes that are highly representative of those that underlie events in the real world.").

164. Note that after criticizing between-subjects tests and calling for more within-subjects designs, Mitchell asks for ecological validity. Mitchell, *Pessismism*, *supra* note 12, at 1985-92. Obviously, in terms of human decision making, between-subjects tests usually have more ecological validity, because within-subjects tests involve the subject being asked to make the same decision over and over again in differing scenarios. *See* Finkel & Groscup, *supra* note 103, at 110 (making this point regarding jury research).

has psychological realism, that is, if the causal processes being represented in the lab setting are the same as those that operate in nonlaboratory contexts.[165]

The artificial nature of the laboratory setting certainly creates cause to question the generalizability of such studies' results to the complex real world,[166] but such generalizability (known as "ecological validity") is often attainable.[167] Recall that laboratory experiments are typically motivated by real world phenomena. After psychologists perform laboratory studies and derive results, they are often then able to produce hypotheses testable in the real world. Questions of ecological validity are on the minds of every psychologist who structures an experiment, and studies that lack it are often attacked.[168] Therefore, psychologists have tested their hypotheses

---

165. Brewer, *supra* note 73, at 14-15. Brewer goes on to illustrate:

> The issue here is one of the level[s] of abstraction at which constructs or principles are defined. Consider, for example, the construct of "threat to self-esteem." No one would seriously deny that being informed that one had failed a test of creative problem-solving would have more impact on self-esteem of a Harvard undergraduate than it would on a 50-year-old mineworker. Thus, if we were interested in the effects of lowered self-esteem on aggression, we might have to use different techniques to lower self-esteem in the two populations. Threats to self-esteem based on challenges to one's academic self-concept are certainly different in many ways from challenges that threaten one's sense of group belonging or of physical stamina. But if each of these, in their appropriate context, proves to have an impact on anger or aggressiveness, then we have gained confidence in a general principle that threats to areas of self-esteem that are important or central to one's sense of identity increase aggression.

Id. at 15.

Evans makes a similar point:

> The distinction between the laboratory and the real world is an odd one. Laboratory experiments are part of the real-world experience of the subjects and their behaviour in them must tell us sometbing. No-one suggests in the science of metallurgy, for example, that the properties of metals studied in the laboratory will have no relevance when the same substances are used in the "real world", for example, as a component in a machine or as a structural support on a bridge. The laboratory is part of the universe in which the laws of physics and chemistry apply. Similarly, subjects of psychological experiments use the same brain in the laboratory as tbey do elsewhere. Psychologists would have to be very clever indeed to succeed constantly in contriving situations wholly unrepresentative of those outside. If biases, errors, and mistakes are so easy to produce in laboratory reasoning tasks, it beggars belief to suppose that these are easily avoided at all other times. Moreover, we are surrounded by evidence of bias, error, and misjudgment in the real world.

J. St. B.T. Evans, *Bias and Rationality, in* RATIONALITY: PSYCHOLOGICAL AND PHILOSOPHICAL PERSPECTIVES 6, 24-25 (K.I. Manktelow & D.E. Over eds., 1993) [hereinafter RATIONALITY].

166. *See* Richard E. Redding, *How Common-Sense Psychology Can Inform Law and Psycholegal Research,* 5 U. CHI. L. SCH. ROUNDTABLE 107, 126 (1998) (criticizing some of Kahneman and Tversky's work as lacking ecological validity).

167. *See* John A. Bargh, *Losing Consciousness: Automatic Influences on Consumer Judgment, Behavior, and Motivation,* 29 J. CONSUMER RES. 280, 281 (2002) (noting that in many experiments, "the dependent measure is taken when the participant believes he or she is entirely outside of an experimental situation—wben arriving, when between different studies, or when leaving the lab").

168. For example, in the legal literature, there has been an intense debate over the ecological validity of studies involving memory accuracy of children who are alleged victims of child abuse. *See, e.g.*, Judith L. Alpert et al., *Symptomatic Clients and Memories of Childhood Abuse: What*

over and over both in the laboratory and in the field. The following examples illustrate this point:

- Consider jury research. In the past fifty years, psychologists have performed literally hundreds of studies, some in the field and more in the laboratory, of various aspects of jury decision making. The field studies are valuable for their ecological validity; the laboratory studies are valuable because "no other approach is capable of yielding the same degree of control over influential extraneous factors, particularly characteristics of the case."[169] Fortunately, mock jury research has improved so that there is little or no difference between the results of studies of mock jurors in university laboratories and the results of studies of real jurors in actual courtrooms.[170]
- Laboratory experiments regarding "public goods" show that subjects do not seek to maximize wealth as the Chicago Man model predicts, but often are willing to contribute to public goods rather than free ride if they believe others will contribute as well.[171] Studies of real

---

the Trauma and Child Sexual Abuse Literature Tells Us, 4 PSYCHOL. PUB. POL'Y & L. 941, 961 (1998) (criticizing some child memory research as lacking ecological validity); Lisa Manshel, *The Child Witness and the Presumption of Authenticity After* State v. Michaels, 26 SETON HALL L. REV. 685, 751-52 (1996) (same).

Scholars have also criticized laboratory studies of "unconscious transference," a memory difficulty that afflicts eyewitnesses to crimes, on grounds that the tension and fear one feels during the commission of a crime is not easily reproduced in a laboratory setting and that, therefore, the results in the laboratory might not be representative of what happens during a real crime. *See, e.g.,* Francis A. Gilligan et al., *The Theory of "Unconscious Transference": The Latest Threat to the Shield Laws Protecting the Privacy of Victims of Sex Offenses,* 38 B.C. L. REV. 107, 123-24 (1996).

169. Dennis J. Devine et al., *Jury Decision Making: 45 Years of Empirical Research on Deliberating Groups,* 7 PSYCH. PUB. POL'Y & L. 622, 698 (2001).

170. *See, e.g.,* Brian H. Bornstein, *The Ecological Validity of Jury Simulations: Is the Jury Still Out?,* 23 LAW & HUM. BEHAV. 75 (1999) (concluding, after reviewing the methodological trends in jury decision-making studies, that different methods of presentation—short case summaries, written transcripts, audiotaped transcripts, and more realistic videotaped presentations—produce very similar results, and that there are virtually no significant differences between results obtained from mock juries composed of college students and those composed of actual jurors); Jonathan J. Koehler, *The Psychology of Numbers in the Courtroom: How to Make DNA-Match Statistics Seem Impressive or Insufficient,* 74 S. CAL. L. REV. 1275, 1297 (2001) ("[A] large amount of empirical literature provides little reason to believe that patterns of data obtained from student-subjects fail to generalize to the jury population."); Michael J. Saks, *What Do Jury Experiments Tell Us About How Juries (Should) Make Decisions?,* 6 S. CAL. INTERDISC. L.J. 1, 7-8 (1997) (defending results of mock jury simulations).

171. *See, e.g.,* Ernst Fehr & Simon Gachter, *Reciprocity and Economics: The Economic Implications of Homo Reciprocans,* 42 EUR. ECON. REV. 845, 854-57 (1998) (arguing that

world behavior confirm the laboratory finding. "Individuals have been shown, for example, to reciprocate the disposition of others to give (or not) to charity, to refrain (or not) from littering, and to wait their turn (or not) in lines."[172]

- Kahneman and Tversky's prospect theory suggests that, inconsistent with traditional economic theory, people expecting a refund after tax withholding are less likely to try to cheat on their taxes than people who expect to make an additional payment.[173] Robben and his colleagues performed laboratory experiments that bore out this prediction.[174] Then they examined IRS analyses of actual taxpayer behavior and found the same effect.[175]
- Prospect theory also assumes that reference points are a key to risk preferences, and predicts that managers of a company falling short of "target points" will be more risk seeking than average managers.[176] Laughhunn and colleagues confirmed this prediction in laboratory experiments.[177] Fiegenbaum and Thomas confirmed the prediction with data from twenty years of corporate activity.[178]
- Griffin and Tversky note that laboratory studies showing irrational overconfidence have been tentatively

---

reciprocity is an important social norm that affects a wide range of human behavior, often in ways that seem inconsistent with utility maximization).

172. KAHAN, *supra* note 46, at 4 (citing studies).

173. Guthrie, *supra* note 1, at 1143-44.

174. Henry S.J. Robben et al., *Decision Frame and Opportunity as Determinants of Tax Cheating: An International Experimental Study*, 11 J. ECON. PSYCHOL. 341, 355 (1990) (finding that "[n]oncompliance was more likely to occur, occurred on more occasions, and involved larger amounts of money among subjects confronting the prospect of additional tax payment after withholding").

175. *Id.* at 345-46.

176. See Mary C. Daly, Panel: Integrity in the Practice of Law: Teaching Integrity in the Professional Responsibility Curriculum: A Modest Proposal for Change, 72 FORDHAM L. REV. 261, 273 (2003) (suggesting that prospect theory may help explain the risks Enron executives were willing to take).

177. Dan J. Laughhunn et al., *Managerial Risk Preferences for Below-Target Returns*, 26 MGMT. SCI. 1238, 1242, 1248 (1980) (reporting results from survey of executives that "suggest the need for new positive models of risky choice behavior, such as that developed by Kahneman and Tversky [in prospect theory], that allow for risk seeking for below target returns" (citation omitted)).

178. Avi Fiegenbaum & Howard Thomas, *Attitudes Toward Risk and the Risk-Return Paradox: Prospect Theory Explanations*, 31 ACAD. MGMT. J. 85, 97 (1988).

supported in studies involving the real world performance of experts.[179]

- Use of the availability heuristic (the tendency to answer hard questions about probability by use of examples that readily come to mind) has been repeatedly documented in laboratory studies,[180] and has also been found in both surveys and actual consumer behavior.[181]
- The endowment effect predicts that people will demand more to sell something they consider in their endowment than they would be willing to pay to obtain it in the first place.[182] Again, studies of real world behavior confirm laboratory experiments.[183]
- Theory predicts that the omission bias, which causes people to regret bad consequences stemming from their actions more than bad consequences stemming from their inaction,[184] will affect people's decisions regarding vaccinating their children with serums that might carry side effects. This has been confirmed first in the laboratory and then in the real world.[185]
- Derek Koehler and colleagues noted that few studies had evaluated how well descriptive theories of probabilistic reasoning captured the behavior of experts in their natural environments.[186] They took laboratory-generated theories and applied them to medical settings, weather forecasting, legal judgments, business

179. Dale Griffin & Amos Tversky, *The Weighing of Evidence and the Determinants of Confidence, in* PSYCHOLOGY OF INTUITIVE JUDGMENT, *supra* note 8, at 230, 230; *see also* Max Henrion & Baruch Fischhoff, *Assessing Uncertainty in Physical Constants, in* PSYCHOLOGY OF INTUITIVE JUDGMENT, *supra* note 8, at 666, 668 ("The few studies of judgments in such real-world contexts [such as medicine, toxicology, and nuclear safety] outside the psychologist's laboratory suggest that the laboratory findings of overconfidence may generalize to situations of practical importance. However, such evaluations have been rare." (citation omitted)).

180. *See generally* Tversky & Kahneman, *supra* note 116, at 11-14.

181. Jacob Gersen, *Strategy and Cognition: Regulating Catastrophic Risk* (2001) (unpublished manuscript 2001), *cited in* CASS R. SUNSTEIN, CONFORMITY AND DISSENT 39 (Univ. of Chi. Law & Econ. Olin Working Paper No. 164 (2d Series), October 2002) (finding that people often buy flood and earthquake insurance after such events occur, but as time passes and memories fade they are much less likely to buy such insurance), http://ssrn.com/abstract_id =341880.

182. JONATHAN BARON, THINKING AND DECIDING 289 (3d ed. 2000).

183. *See supra* note 41.

184. Jonathan Baron, Preferences and Rational Choice: New Perspectives and Legal Implications: Value Analysis of Political Behavior—Self-Interest: Moralistic:: Altruistic: Moral, 151 U. PA. L. REV. 1135, 1150-51 (2003) (explaining omission bias).

185. *See infra* note 467.

186. Derek J. Koehler et al., *The Calibration of Expert Judgment: Heuristics and Biases Beyond the Laboratory, in* PSYCHOLOGY OF INTUITIVE JUDGMENT, *supra* note 8, at 686, 710.

settings, and sports settings, concluding that "[i]n all domains of expert judgment surveyed, systematic miscalibration was observed. In each case, the observed patterns matched the qualitative predictions of the heuristics and biases perspective . . . ."[187]

- In the field of organizational behavior, Locke observes that "[i]n case after case and topic after topic, basically the same results were obtained in the field as in the laboratory. . . . The evidence indicates that a detailed, point-by-point similarity with respect to subjects, tasks, settings, and so forth is not necessarily required in order to achieve generalizability."[188] Thus, studies show very similar results between laboratory experiments and field studies in such organizational behavior and organizational psychology areas as goal setting,[189] feedback effects,[190] decision-making participation,[191] financial incentives,[192] and the relationship of job performance and job satisfaction.[193] All this led Professor Ilgen to conclude that "[t]ime and again, results of research conducted in the laboratory were found to generalize to organizational settings."[194]

---

187. *Id.*

188. Edwin A. Locke, *Generalizing from Laboratory to Field: Ecological Validity or Abstraction of Essential Elements?*, *in* GENERALIZING FROM LABORATORY TO FIELD SETTINGS 3, 6 (Edwin A. Locke ed., 1986) [hereinafter GENERALIZING FROM LABORATORY].

189. *See* Gary P. Latham & Thomas W. Lee, *Goal Setting*, *in* GENERALIZING FROM LABORATORY, *supra* note 188, at 101, 108 (noting that laboratory studies of goal setting and its impact on behavior "readily" generalize to field studies).

190. *See* Richard E. Kopelman, *Feedback*, *in* GENERALIZING FROM LABORATORY, *supra* note 188, at 119, 139-40 (concluding from a survey of thirty laboratory studies and forty-two field experiments that objective feedback "consistently has a positive effect in both the laboratory and the field" although its effect in the field is often stronger).

191. *See* David M. Schweiger & Carrie R. Leana, *Participation in Decision Making*, *in* GENERALIZING FROM LABORATORY, *supra* note 188, at 147, 161 (finding that "research results obtained in the laboratory generally agree with those obtained in the field" in the context of experiments regarding whether subordinate participation in decision making improves their goal acceptance).

192. *See* G. Douglas Jenkins, Jr., *Financial Incentives*, *in* GENERALIZING FROM LABORATORY, *supra* note 188, at 167, 177 (finding in survey of studies on impact of financial incentives on behavior that "laboratory findings about financial incentives do generalize to field settings, but only when certain specified conditions are met").

193. *See* Philip M. Podsakoff & Larry J. Williams, *The Relationship Between Job Performance and Job Satisfaction*, *in* GENERALIZING FROM LABORATORY, *supra* note 188, at 207, 244 ("[T]he pattern of relationships obtained in the laboratory is generally consistent with those obtained in field research.").

194. *See* Daniel R. Ilgen, *Laboratory Research: A Question of When, Not If*, *in* GENERALIZING FROM LABORATORY, *supra* note 188, at 257, 257. Ilgen was summarizing the results of numerous studies comparing laboratory results with field results contained in GENERALIZING FROM

- Finally, Markman and Medin noted in 2002 that "[i]n general, results from laboratory studies have held up surprisingly well when tested, for example, on the floor of a casino or at the race track."[195]

Mitchell's main example of psychology's lack of realism relates to probabilities and frequencies.[196] As explained elsewhere in this article,[197] over the last twenty years studies have demonstrated that humans are often poor statistical reasoners.[198] Subsequent studies indicate that human performance on some tests can be improved if the questions are presented in a frequency rather than a probability format.[199] This is all well and good. Unfortunately, while Mitchell attempts to characterize the questions in the K-T studies as involving "unnatural and unfamiliar formats,"[200] in the real world people often have to deal with problems presented as probabilities.[201]

2. Lack of Feedback and Learning Opportunities

Mitchell points out that social scientists using between-subject tests that are based on one-shot decision situations and thus fail to test for feedback, learning, and market interaction effects may not get an accurate picture when describing decision-making behavior in repetitive markets.[202] This seems a fairly obvious point, but any implication that subjects are being tricked into giving wrong answers is inaccurate.[203]

LABORATORY. Altbough not every single study found strong generalizability from laboratory experiments, the strong thrust of most of the studies was consistent with Ilgen's conclusion. *See also* Kuhberger, *supra* note 136, at 45 ("[F]raming research has stepped outside the lab to a considerable degree . . . [demonstrating that] experts are also influenced by framing, but maybe to a lesser degree than students.").

195. Arthur B. Markman & Douglas L. Medin, *Decision Making, in* 2 STEVEN'S HANDBOOK OF EXPERIMENTAL PSYCHOLOGY 413, 451 (D.L. Medin & H. Pashler eds., 3d ed. 2002).

196. Mitchell, *Pessimism, supra* note 12, at 1988-1992.

197. *See infra* notes 368-378 and accompanying text.

198. *See* Evans, *supra* note 165, at 24-25 (noting that although "man" was formerly regarded as a good intuitive statistical reasoner, over the past two decades "evidence has accumulated that the way in which subjective probabilities are formed is apparently subject to a wide variety of biases").

199. *See, e.g.*, G. Gigerenzer, *The Bounded Rationality of Probabilistic Mental Models, in* RATIONALITY, *supra* note 165, at 284, 293-94.

200. Mitchell, *Pessimism, supra* note 12, at 1992.

201. *See infra* note 381 and accompanying text.

202. *See* Mitchell, *Pessimism, supra* note 12, at 1977-79.

203. Regarding studies involving statistical reasoning, leading experts recently wrote:

> The accusation that psychologists have been devising parlor tricks, which people are susceptible to in the laboratory context but either do not encounter or could solve in real world contexts, seems less plausible in view of the research reported here. First, for each problem we have reported, some of the subjects showed by their answers (and

The background controversy, which Mitchell explicates, is interesting. Whereas psychologists in the heuristics and biases school tend to study one-shot decisions, experimental economists tend to create markets and allow subjects to play repeatedly to determine whether eventually their decision making will tend toward the rational and create efficient markets.[204] The debate, therefore, involves comparing apples and oranges, decision making in the one-shot setting versus decision making in a repeat game setting.

To the extent that many decisions made by individuals are not repeated frequently (e.g., decisions to buy a house, to buy a car, to invest an inheritance, to take a new job, to vote as a juror, and so on), the behavioralists' approach is obviously the more appropriate one.[205] Moreover, experience seldom teaches investors to cure their overconfidence.[206] Often, even many trials and large amounts of feedback do not make Chicago Man out of K-T Man; many psychological research results "do live through shockingly high levels of feedback and multiple trials."[207]

> often by the rationales for their answers, subsequently elicited) an appreciation of the statistical principles that in previous work other subjects failed to appreciate. A more reasonable explanation of the success of some of our subjects is that they are more skilled at statistical reasoning than the other subjects rather than that they saw through the experimenters' tricks. Second, the factors that make statistical reasoning more or less likely (for example, recognition of heterogeneity and of the role played by chance) do not sound like factors that make people more or less dupable by experimenters but rather like factors that make the appropriateness of statistical reasoning more or less obvious. Third, statistical training markedly influences answers to the sort of problems we studied. This suggests that it is not problem- or context-produced illusions that make people unable to solve statistical problems, but simply lack of statistical knowledge.

Richard E. Nisbett et al., *The Use of Statistical Heuristics in Everyday Inductive Reasoning, in* PSYCHOLOGY OF INTUITIVE JUDGMENT, *supra* note 8, at 510, 530-31.

204. Whereas psychologists may be accused, and are by Mitchell, of manipulating their experiments to produce quirks in human decision making, Mitchell, *Pessimism, supra* note 12, at 1971, economists are often accused of similar manipulations with opposite goals. *See* Douglas L. Medin & Max H. Bazerman, *Broadening Behavioral Decision Research: Multiple Levels of Cognitive Processing,* 6 PSYCHONOMIC BULL. & REV. 533, 536 (1999) (suggesting that "much of experimental economics consists of contrived experiments created in order to show convergence" of actual decision making and optimal decision making).

205. *See* Hillel J. Einhorn & Robin M. Hogarth, *Confidence in Judgment: Persistence of the Illusion of Validity,* 85 PSYCHOL. REV. 395, 413-14 (1978) (noting that many situations do not provide feedback in a way that facilitates learning).

206. *See* Donald C. Langevoort, *Selling Hope, Selling Risk: Some Lessons for Law from Behavioral Economics About Stockbrokers and Sophisticated Customers,* 84 CAL. L. REV. 627, 639 (1996) (citing Sheryl B. Ball et al., *An Evaluation of Learning in the Bilateral Winner's Curse,* 48 ORGANIZATIONAL BEHAV. & HUM. DECISION PROCESSES 1, 17-18 (1991), and Berndt Brehmer, *In One Word: Not from Experience, in* JUDGMENT AND DECISION-MAKING: AN INTERDISCIPLINARY READER 705 (Hal R. Arkes & Kenneth R. Hammond eds., 1986) (making the point that feedback is seldom as clear and unambiguous as needed to cure overconfidence)).

207. Medin & Bazerman, *supra* note 204, at 536.

Even the repeat games of experimental economist and Nobel Prize winner Vernon Smith demonstrated that the stock market, the market most likely of all markets to be efficient, often is not.[208] Stock markets are often subject to irrational bubbles[209] and are rife with inefficiencies well documented by behavioral finance research.[210] Even repeat professional players in the financial markets—players who have time and incentive to perform well—are subject to many of the heuristics and biases that have been identified in lay persons by one-shot laboratory experiments.[211]

---

208. *See* Justin Fox, *Is the Market Rational?* FORTUNE, Dec. 9, 2002, at 116 (noting that Vernon Smith's "economic experiements have also shot holes in efficient-markets dogma"); *see also* George A. Akerlof & Janet L. Yellen, *Can Small Deviations from Rationality Make Significant Differences to Economic Equilibria?*, 75 AM. ECON. REV. 708, 708 (1985) (giving examples of situations where small deviations from rationality can have major impact); Kenneth L. Fisher & Meir Statman, *Cognitive Biases in Market Forecasts*, J. PORTFOLIO MGMT., Fall 2000, at 72, 72 (arguing that five cognitive biases, including overconfidence, confirmation, representativeness, anchoring, and hindsight, underlie the illusion of validity that causes investors to make poor decisions); Thomas Russell & Richard Tbaler, *The Relevance of Quasi Rationality in Competitive Markets*, 75 AM. ECON. REV. 1071, 1071 (1985) (showing tbat "the knee-jerk reaction of some economists that competition will render irrationality irrelevant is apt only in very special cases, probably rarely observed in the real world").

209. *See generally* MATHIAS BINSWANGER, SPECULATIVE BUBBLES AND ECONOMIC GROWTH: NEW DIMENSIONS IN THE COEVOLUTION OF REAL AND FINANCIAL MARKETS (1999) (accurately projecting that the 1990s boom was a bubble); EDWARD CHANCELLOR, DEVIL TAKE THE HINDMOST: A HISTORY OF FINANCIAL SPECULATION (1999) (examining "Tulipmania" and other early bubbles); PETER M. GARBER, FAMOUS FIRST BUBBLES: THE FUNDAMENTALS OF EARLY MANIAS (2000) (same); ROBERT J. SHILLER, IRRATIONAL EXUBERANCE (2000) (predicting, accurately as it turned out, that the dot com stock market run-up of the mid- to late-1990s was a classic bubble).

210. Many books extensively document behavioral anomalies in the operation of supposedly efficient stock markets. *See, e.g.*, ADVANCES IN BEHAVIORAL FINANCE (Richard H. Thaler ed., 1993); GARY BELSKY & THOMAS GILOVICH, WHY SMART PEOPLE MAKE BIG MONEY MISTAKES AND HOW TO CORRECT THEM (1999); HERSH SHEFRIN, BEYOND GREED AND FEAR: UNDERSTANDING BEHAVIORAL FINANCE AND THE PSYCHOLOGY OF INVESTING (2000); ANDRE SHLEIFER, INEFFICIENT MARKETS: AN INTRODUCTION TO BEHAVIORAL FINANCE (2000); RICHARD H. THALER, QUASI RATIONAL ECONOMICS (1991); RICHARD H. THALER, THE WINNER'S CURSE: PARADOXES AND ANOMALIES OF ECONOMIC LIFE (1992); LARS TVEDE, THE PSYCHOLOGY OF FINANCE (1999); *see also* Kent Daniel et al., *Investor Psychology and Security Market Under- and Overreactions*, 53 J. FIN. 1839, 1865 (1998) (suggesting that investor overconfidence helps account for stock price behavior); Werner De Bondt & Richard H. Thaler, *Does the Stock Market Overreact?*, 40 J. FIN. 793 (1985) (showing that investors overreact to new information and underweigh older information when making investment decisions); Terrence Odean, *Are Investors Reluctant to Realize Their Losses?*, 53 J. FIN. 1775 (1998) (explaining how investors' susceptibility to loss aversion affects the stock market); SENDHIL MULLAINATHAN & RICHARD H. THALER, BEHAVIORAL ECONOMICS, (Nat'l Bureau of Econ. Research, Working Paper No. W7948, Oct. 2000) (explaining that the limits of arbitrage prevent even sophisticated financial markets from eliminating the effects of psychological decision-making limitations via market forces, learning and evolution), bttp://papers.ssrn.com/sol3/papers.cfm?abstract_id=245733.

211. Langevoort has explained that because even for professional investors in the stock markets feedback is seldom prompt and unambiguous, various biases in decision making can be persistent. *See* Langevoort, *Taming the Animal Spirits*, *supra* note 10, at 135; Langevoort, *supra* note 206, at 636-41. *See generally* Werner F. M. De Bondt & Richard M. Thaler, *Do Security*

3. Conversational Cues and Demand Characteristics

In his litany of potential problems with experimental research, Mitchell's next complaint addresses two related concepts. First are conversational cues—the notion that subject responses viewed as erroneous by experimenters are "actually reasonable responses to conversational cues contained in the experimental stimuli."[212] Second are demand characteristics; the notion that subjects, knowing that they are in an experiment, make surmises regarding the experimenter's purpose and shape their responses accordingly.[213]

Conversational cues and demand characteristics are part of a family of features of experimental research that can potentially confound results because "alert, aware participants are actively seeking cues in the research setting to inform them of what they are expected to do or what they should do in order to present themselves in a favorable light."[214]

Analysts Overreact?, AM. ECON. REV., May 1990, at 52, 52 (finding that securities analysts display the availability heuristic in making forecasts); De Bondt & Thaler, *supra* note 210, at 793 (noting that there is substantial evidence that stock professionals, like lay people, do not process new information consistent with Bayes's rule but instead overreact to new information); Bradford De Long et al., *The Survival of Noise Traders in Financial Markets*, 64 J. BUS. 1, 5-6 (1991) (noting overconfidence of several types of professionals including CIA analysts, experienced psychologists, and physicians and explaining how irrational "noise traders" can survive and even prosper in supposedly efficient financial markets); David Hirshleifer et al., *Security Analysts and Trading Patterns When Some Investors Receive Information Before Others*, 49 J. FIN. 1665, 1686 (1994) (noting evidence of overconfidence by professional stock investors); Robert A. Prentice, *The SEC and MDP: Implications of the Self-Serving Bias for Independent Auditing*, 61 OHIO ST. L.J. 1597, 1622-25 (2000) (surveying evidence that stock professionals are subject to the self-serving bias); Russ Wermers, *Mutual Fund Herding and the Impact on Stock Prices*, 54 J. FIN. 581, 618 (1999) (providing evidence that mutual fund managers are subject to irrational herding effects); GILLES HILARY & LIOR MENZLY, DO PAST SUCCESS LEAD ANALYSTS TO BECOME OVERCONFIDENT? 2 (Univ. of Chi., Apr. 2001) (providing evidence of overconfidence by professional stock analysts), http://papers.ssrn.com/sol3/papers.cfm?abstract_id=261476. *But see* Michael B. Mikhail et al., *Do Security Analysts Improve Their Performance with Experience?*, 35 J. ACCT. RES. 131, 155 (Supp. 1997) (finding evidence that sell-side security analysts following a specific firm do tend to improve performance with experience).

212. Mitchell, *Pessimism*, *supra* note 12, at 1981.

213. *See* CARLSMITH ET AL., *supra* note 69, at 280-81 (describing demand characteristics); Donald P. Judges, *Two Cheers for the Department of Justice's Eyewitness Evidence: A Guide for Law Enforcement*, 53 ARK. L. REV. 231, 253 (2000) (noting that demand characteristics occur when "the subject's responses are distorted by his or her understanding of the investigator's goals").

214. Brewer, *supra* note 73, at 8; *see also* Ara Norenzayan & Norbert Schwartz, *Telling What They Want to Know: Participants Tailor Causal Attributions to Researchers' Interests*, 29 EUR. J. SOC. PSYCHOL. 1011, 1019 (1999) (finding that subjects responding to a questionnaire tended to give different answers depending on whether the researcher was identified as a social scientist or as a personality psychologist); Wayne E. Ormond & Lorne M. Sulsky, *Influence of Raters' Self-Consciousness and Appraisal Purpose on Leniency and Accuracy of Performance Ratings: A Critique*, 82 PSYCHOL. REP. 32, 33 (1998) (criticizing another study on grounds that its results

a. Conversational Cues

One of Mitchell's main sources on conversational cues is Professor Schwarz, who points out that according to the tacit rules of conversation in everyday life, people expect to be given relevant information.[215] So, if an experimenter gives information to a subject, the subject will consider that information, because the subject assumes that there was a reason for the experimenter to include it.[216] This means that if subjects in an experiment are told that the usual assumption that information is relevant does not hold, they will be less likely to fall prey to various biases and inefficient heuristics studied in the literature.[217]

Fortunately, conversational cues are simply irrelevant to experiments regarding many of the heuristics and biases in the K-T tradition.[218] However, they are particularly relevant to the dilution effect (the tendency for the addition of irrelevant information to cause decision makers to form less extreme judgments than those based only on relevant information),[219] and have been studied in that regard. Some studies show that subjects are unable to disregard irrelevant information when making decisions, even after having several opportunities to do so.[220] Experiments indicate that decision makers are worse off with more information than with less, because they cannot ignore the non-diagnostic information.[221] However, Tetlock and colleagues, like Schwarz, suspected that the subjects persisted in considering the irrelevant information in part because they assumed,

---

could be explained primarily by demand characteristics introduced by the experimental procedure).

215. Mitchell, *Pessimism, supra* note 12, at 1981-82 (quoting Norbert Schwarz, *Judgment in a Social Context: Biases, Shortcomings, and the Logic of Conversation,* 26 ADVANCED EXPERIMENTAL SOC. PSYCHOL. 123, 154-56 (1994)).

216. *See* Norbert Schwarz, *Social Judgment and Attitudes: Warmer, More Social, and Less Conscious,* 30 EUR. J. SOC. PSYCHOL. 149, 152 (2000); Philip E. Tetlock, *Intuitive Politicians, Theologians, and Prosecutors: Exploring the Empirical Implications of Deviant Functionalist Metaphors, in* PSYCHOLOGY OF INTUITIVE JUDGMENT, *supra* note 8, at 582, 594.

217. *See* Schwarz, *supra* note 216, at 152.

218. Conversational cues would have no particular impact, for example, upon studies of the availability heuristic, anchoring and adjustment, attribution theory, group judgments, overconfidence, overoptimism, and many other effects.

219. *See generally* Richard E. Nisbett et al., *The Dilution Effect: Nondiagnostic Information Weakens the Implications of Diagnostic Information,* 13 COGNITIVE PSYCHOL. 248 (1981). Thus, jurors provided with some relevant evidence indicating defendant's guilt are more likely to find the defendant guilty than jurors provided with that same evidence *and* with additional *irrelevant* information.

220. *See* N. John Castellan, Jr., *Multiple-Cue Probability Learning with Irrelevant Cues,* 9 ORGANIZATIONAL BEHAV. & HUM. PERFORMANCE 16, 26 (1973) (finding that subjects were unable to ignore irrelevant information even after a large number of trials).

221. *Id.* at 26.

consistent with normal conversational practice, that the experimenters would not have provided the information if it had no bearing upon the task. In experiments, they found evidence to support their hypothesis.[222]

This finding scarcely means that the results of the original studies of the dilution effect are not valuable, however. For example, auditors have been shown to be subject to the dilution effect.[223] Unfortunately for auditors, a client that floods its outside auditor with irrelevant information is not going to advertise to the auditor the irrelevance of the information. Similarly, sellers of products and promoters of securities can bombard consumers and investors with irrelevant data and thereby reduce the accuracy of their decision making. If they intend to manipulate or defraud, they are unlikely to red flag for special attention the irrelevant information that they include in their messages. The outside auditors and consumers, consistent with conversational norms, will likely believe, erroneously, that the information has been provided because it bears upon the decisions they must make.[224]

b. Demand Characteristics

Researchers are well aware of the impact of demand characteristics (also called demand effects).[225] These occur when subjects in experiments act to try to please the experimenter.[226]

---

222. Philip E. Tetlock et al., *Revising the Value Pluralism Model: Incorporating Social Content and Context Postulates, in* THE PSYCHOLOGY OF VALUES 25 (Clive Seligman et al. eds., 1996), *cited in* Tetlock *supra* note 216, at 594.

223. *See, e.g.*, Steven M. Glover, *The Influence of Time Pressure and Accountability on Auditors' Processing of Nondiagnostic Information,* 35 J. ACCT. RES. 213, 223 (1997) (finding that accountants are subject to the dilution effect); Karl Hackenbrack, *Implications of Seemingly Irrelevant Evidence in Audit Judgment,* 30 J. ACCT. RES. 126 (1992) (same); Vicky B. Hoffman & James M. Patton, *Accountability, the Dilution Effect, and Conservatism in Auditors' Fraud Judgments,* 35 J. ACCT. RES. 227, 233 (1997) (same).

224. *See generally* Alex Chernev, *The Effect of Common Features on Brand Choice: Moderating Role of Attribute Importance,* 23 J. CONSUMER RES. 304 (1997) (exploring how adding information affects consumer decision making); Tom Meyvis & Chris Janiszewski, *Consumers' Beliefs About Product Benefits: The Effect of Obviously Irrelevant Product Information,* 28 J. CONSUMER RES. 618 (2002) (same).

225. *See* Daniel Kahneman & Amos Tversky, *On the Study of Statistical Intuitions,* 11 COGNITION 123, 124, 132, 135 (1982) (finding that the results of earlier studies which had concluded that exposure to the color pink reduced muscle strength primarily reflected demand characteristics); *see also* Norbert Schwarz, *Self-Reports: How the Questions Shape the Answers,* 54 AM. PSYCHOLOGIST 93, 96 (1999) (reporting that whether survey letterhead said "Institute for Personality Research" or "Institute for Social Research" affected responses to survey on reasons for mass murder); Jeffrey M. Smith et al., *The Influence of Color and Demand Characteristics on Muscle Strength and Affective Ratings of the Environment,* 113 J. GEN. PSYCHOL. 289, 297 (1986).

226. *See* F.J. MCGUIGAN, EXPERIMENTAL PSYCHOLOGY: A METHODOLOGICAL APPROACH 287-91 (1960) (explaining demand characteristics).

Researchers whose experimental designs create the opportunity for significant demand characteristics often find their results under attack.[227]

Because psychologists are aware of the effects of demand characteristics, they are constantly refining experimental procedures to minimize or eliminate them.[228] There are ways of planning and designing research operations so that the number of potentially confounding factors associated with the independent variable can be reduced.[229] One of these methods is to utilize the between-subjects research design that Mitchell criticizes.[230] Others include (a) using a coherent and believable cover story;[231] (b) keeping experimenters, confederates, and others who come into contact with the subjects as unaware as possible of each subject's condition;[232] (c) using different rooms and different experimenters when subjects are asked to do two separate but related tasks;[233] (d) using research instruments based on

227. *See* Norman J. Finkel, *Commonsense Justice and Jury Instructions: Instructive and Reciprocating Connections,* 6 PSYCHOL. PUB. POL'Y & L. 591, 613-14 (2000) (attacking a related study for not adequately guarding against demand effects); Ormond & Sulsky, *supra* note 214, at 32.

228. *See, e.g.,* Olivier Corneille et al., *Judgeability Concerns: The Interplay of Information, Applicability, and Accountability in the Overattribution Bias,* 76 J. PERSONALITY & SOC. PSYCHOL. 377, 381 (1999) (designing study of the overattribution bias to eliminate effects of conversational cues and demand characteristics); Duncan Cramer & Natalie Buckland, *Effect of Rational and Irrational Statements and Demand Characteristics on Task Anxiety,* 129 J. PSYCHOL. 269, 274 (1995) (controlling for demand characteristics); John M. Govern & Lisa A. Marsch, *Inducing Positive Mood Without Demand Characteristics,* 81 PSYCHOL. REP. 1027, 1032 (1997) (testing a new technique for altering subjects' mood without demand effects in order to study the impact of that altered mood); David Wiseman & Irwin P. Levin, *A New Laboratory Method for Altering Positive Affect,* 76 PSYCHOL. REP. 1103, 1106 (1995) (developing a new method for altering positive affect without demand effects in order to study the impact of that altered affect).

229. Brewer, *supra* note 73, at 8.

230. *See* Smith, *supra* note 74, at 23 ("[P]articipants in a within-participants design see more than one condition and are thus in a better position to guess at the experimental hypotheses."); Finkel, *supra* note 73, at 923 (noting that within-subjects designs are more likely to inadvertently create demand characteristics than between-subjects designs).

231. *See* CARLSMITH ET AL., *supra* note 69, at 283; Ronald S. Friedman & Jens Forster, *The Effects of Approach and Avoidance Motor Actions on the Elements of Creative Insight,* 79 J. PERSONALITY & SOC. PSYCHOL. 477, 481 (2000) (using a cover story to eliminate self-perception effects); Smith, *supra* note 74, at 35.

232. *See* Smith, *supra* note 74, at 35. For example, in a study of media influence on public perceptions of air bag safety, Feigenson kept experimenters in the dark regarding the subjects' condition and the purpose of the study in order to eliminate demand characteristics. *See* Neil R. Feigenson, *Air Bag Safcty: Media Coverage, Popular Conceptions, and Public Policy,* 7 PSYCHOL. PUB. POL'Y & L. 444, 476-77 (2001).

233. *See* John A. Bargh & Tanya L. Chartrand, *The Mind in the Middle: A Practical Guide to Priming and Automaticity Research,* in HANDBOOK OF RESEARCH METHODS, *supra* note 73, at 253, 267.

the real world;[234] and (e) supplementing laboratory results with *unsolicited* responses gathered from various sources such as anecdotes, public testimony, police reports, and the like.[235]

As just one specific example, in studying the phenomenon of anchoring and adjustment, experimenters have often used as anchors numbers that were *obviously* uninformative (such as the spin of a roulette wheel) in order to insure that subjects did not incorporate the number into their decision process on the assumption that it must be relevant simply because the experimenter mentioned it.[236]

The fact, noted earlier, that well designed laboratory experiments tend to produce the same results as field studies[237] indicates that researchers are doing a reasonably good job of minimizing distortions caused by conversational cues, demand characteristics, and related phenomena. Mitchell is certainly correct in arguing that psychologists must remain highly sensitive to these problems. Sechrest and Bootzin properly urge psychologists to do more of their research in real life settings.[238] But studies such as those by Kunreuther and his colleagues, which found irrationality in real-life decisions that mirrored laboratory results with regard to insurance policies, should give pause to those who attack the laboratory.[239] Studying how to minimize demand characteristics is important not just in the laboratory, but also in the real world; real world jurors' responses to voir dire questions are affected by their

---

234. *See* Richard L. Wiener & Dennis P. Stolle, *Trial Consulting: Jurors' and Attorneys' Perceptions of Murder*, 34 CAL. W. L. REV. 225, 242 (1997) (comparing the predictions of public defenders with the decisions of jurors who had filled out jury questionnaires to study how effectively attorneys instinctively gauge whether jurors would or would not be favorable to their case). To minimize demand characteristics, Wiener and Stolle used a questionnaire based substantially on a real juror questionnaire commonly used by public defender offices. *Id.*

235. *See* Neil M. Malamuth & James V.P. Check, *The Effects of Mass Media Exposure on Acceptance of Violence Against Women: A Field Experiment*, 15 J. RES. PERSONALITY 436, 437 (1981) (using such unbiased sources to study effects of media exposure on male acceptance of rape myths); *see also* CARLSMITH ET AL., *supra* note 69, at 283-92 (discussing other means of reducing demand characteristics).

236. Gretchen B. Chapman & Eric J. Johnson, *Incorporating the Irrelevant: Anchors in Judgments of Belief and Value*, *in* PSYCHOLOGY OF INTUITIVE JUDGMENT, *supra* note 8, at 120, 122 (citing J. EDWARD RUSSON & PAUL J.H. SHOEMAKER, DECISION TRAPS 90 (1989)).

237. *See supra* notes 169-195 and accompanying text.

238. Lee B. Sechrest & Richard R. Bootzin, *Psychology and Inferences About Public Policy*, 2 PSYCHOL. PUB. POL'Y & L. 377, 380 (1996).

239. *See* HOWARD KUNREUTHER ET AL., DISASTER INSURANCE PROTECTION: PUBLIC POLICY LESSONS 1-4, 237 (1978) (finding, consistent with laboratory experiments, that low probability events are systematically discounted by individuals); *see also* Johnson et al., *supra* note 87, at 50 (finding strong evidence supporting framing effects in real life insurance decisions made by consumers under two different statutory schemes).

perceptions of judges' and attorneys' expectations,[240] as are witness responses to questions from police and lawyers.[241]

4. Use of Ecologically Suspect Formats

Mitchell's next argument is that decision-making errors are often exaggerated by the ecologically suspect format of psychological experiments, specifically the "decontextualized, abstract, or unnatural nature of the research setting and research task."[242] Aside from an example of the Wason four-card selection task,[243] which tests conditional reasoning performance,[244] Mitchell points only to the probability/frequency debate discussed earlier.[245] Because the evidence is mixed as to whether frequency formats actually improve performance over probability formats, and because people frequently confront probability formats in real life,[246] the impact of Mitchell's argument is limited. More importantly, again consider the substantial evidence cited above indicating that the same phenomena discovered in the sterile environment of the laboratory have been replicated repeatedly in real world decision making.[247]

---

240. *See* Linda L. Marshall & Althea Smith, *The Effects of Demand Characteristics, Evaluation Anxiety, and Expectancy on Juror Honesty During Voir Dire*, 120 J. PSYCHOL. 205, 213 (1986) (finding that juror responses were affected most by evaluation anxiety, less by previous jury duty, and less still by expectancy effects and demand characteristics).

241. *See* Gary P. Wells & Eric P. Seelau, *Eyewitness Identification: Psychological Research and Legal Policy on Lineups*, 1 PSYCHOL. PUB. POL'Y & L. 765, 767 (1995) (noting that "[f]actors that can create interpretation difficulties for an experiment can create similar problems for a lineup" and giving demand characteristics as an example).

242. Mitchell, *Pessimism, supra* note 12, at 1985.

243. The Wason test is explained *infra* note 358. It is discussed in some detail there.

244. Mitchell points out that only 10% of some types of subjects can solve abstract forms of the problem. *See* Stephen J. Hoch & Judith E. Tschirgi, *Logical Knowledge and Cue Redundancy in Deductive Reasoning*, 13 MEMORY & COGNITION 453, 456 (1985) (finding a 48% success rate with subjects holding a master's degree), *cited in* Mitchell, *Pessimism, supra* note 12, at 1986. That percentage rises substantially when the test is given in a more familiar setting. *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) (noting that only one percent of subjects give the correct response when the problem is stated abstractly, that 7-9% succeed when given some versions of the problem with a factual setting, but that the success rate rose to 74% when the problem was phrased with facts that allowed the subject to recall past experience with the content of the problem, the relationship expressed, and a counter-example to the rule governing the relationship); *see also* Mitchell, *Pessimism, supra* note 12, at 1987 (giving an example of a more familiar context for the question).

245. *See supra* notes 197-201 and accompanying text.

246. *See infra* note 381.

247. *See supra* notes 169-195 and accompanying text.

5. Deprivation of Decision Tools

Finally, Mitchell argues that researchers in the K-T tradition exaggerate non-normative response findings by depriving subjects of decision aids such as instructional texts, calculators, or computers.[248] Mitchell points out that a behavioral decision theorist has placed a Bayesian calculator online in order to assist decision makers in updating odds in conformance with Bayes' theorem.[249]

Of course, decision aids are irrelevant to most of the heuristics and biases identified in the K-T tradition. They could be relevant to probabilistic reasoning tasks, but it seems unlikely that online Bayesian calculators are going to be in widespread use in general society anytime soon, and as Mitchell himself admits "[t]he literature on decision aids indicates that people may be reluctant to use mechanical decision tools in place of their own judgment."[250] In other words, the fact that we can all carry calculators around with us in the real world if we so choose is not going to cure us of overoptimism, overconfidence, loss aversion, the endowment effect, or most other heuristics and biases identified in the K-T literature.

C. *Does Behavioral Decision Theory Have Features That Limit the Importance of the Research for the Legal System?*

In the final section of his attack on psychology research, Mitchell makes a multi-faceted argument that it is perilous to extrapolate the findings of behavioral decision research to the real legal world.[251] The arguments are interesting and worthy of individual analysis, even though much of the heuristics and biases literature has already been confirmed in real world decision making, as noted above.[252]

---

248. Mitchell, *Pessimism, supra* note 12, at 1993.

249. Michael H. Birnbaum, Bayesian Calculator, *at* http://psych.Fullerton.edu/mbirnbaum/bayes/BayesCalc.htm (last visited Nov. 9, 2003).

Bayes' Theorem is a rational method for updating probabilities. *See generally* ROBYN M. DAWES, RATIONAL CHOICE IN AN UNCERTAIN WORLD 323-26 (1988).

250. Mitchell, *Pessimism, supra* note 12, at 1995 n.175 ("[A] common theme emerging from recent decision aid research is that decision makers are reluctant to relinquish their judgments in favor of decision aids." (quoting 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))); 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) (finding that decision aids will be avoided unless they minimize the decision maker's overall level of effort expenditure).

251. Mitchell, *Pessimism, supra* note 12, at 1995-2017.

252. *See supra* notes 169-195 and accompanying text.

1. Real World Success Versus Normative Coherence

The first facet of Mitchell's argument is that laboratory results do not necessarily translate to the real world.[253] He begins with an example that is rather easily turned against him. In one of his few attacks on a specific behavioralist theory, Mitchell notes that Russell Korobkin has argued that because of cognitive imperfections, "the party preparing the first contract draft may 'gain a powerful advantage in negotiations.'"[254] In response, Mitchell hazards a totally unsupported supposition that "[t]he original terms supposedly accepted with the assistance of cognitive imperfections, could well lead to a fair and efficient transaction."[255]

Unfortunately for Mitchell's argument, the form contracts that consumers and investors sign are, because of the merchants' and promoters' self-serving bias, relentlessly one-sided, and therefore often unfair and inefficient.[256] Professor Farnsworth, the reporter for the Restatement (Second) of Contracts, noted that in his own experience in legal practice "no one in any of the corporations or in the law firm ever suggested that the forms should be drafted other than as one-

253. Mitchell, *Pessimism, supra* note 12, at 1995-2002.

254. *Id.* at 1996 (quoting Russell Korobkin, *Inertia and Preference in Contract Negotiation: The Psychological Power of Default Rules and Form Terms,* 51 VAND. L. REV. 1583, 1627 (1998)).

255. *Id.*

256. As Slawson has noted:

> Forms standardized to achieve economies of mass production and mass merchandising will also, under the present system, almost certainly be unfair, because if they were not, their issuers would probably lose money. An unfair form will not deter sales because the seller can easily arrange his sales so that few if any buyers will read his forms, whatever their terms, and he risks nothing because the law will treat his forms as contracts anyway. The user of an unfair form does not even stand to lose any significant number of future sales because the contingencies against which his forms provide him protection are normally of a kind which only infrequently occur (although when they do, the buyer may lose a great deal). When such a contingency arises the buyer will not usually be in a position to compare the form he bought with others he might have bought instead. Most buyers probably believe (correctly) that the forms they could have bought from a competing seller would have been just as bad anyway. An unfair form thus normally constitutes a costless benefit which a seller refuses at his peril. If he fails to take advantage of it, his competitors will. Competitive pressures have worked so long and so thoroughly to make standard forms unfair that we no longer even notice the unfairness. Standard credit agreements commonly allow the lender to call the entire unpaid balance, plus costs of collection, should even a single payment be a moment late, or, not uncommonly, should the lender just wake up some morning feeling "insecure," but it is rare that either provision occasions even a judicial comment. A standard agreement recently signed by a colleague of [Slawson's] contained provisions disclaiming all representations and warranties of year, model, mileage, price or design-change prior to delivery (!), but he signed it without thought.

W. David Slawson, *Standard Form Contracts and Democratic Control of Lawmaking Power,* 84 HARV. L. REV. 529, 531-32 (1971).

sidedly in the interests of the corporate client as possible."[257] Behavioral considerations, rather than efficient bargaining, explain why consumers and investors continue to sign these one-sided contracts.[258]

Mitchell then attacks a study done by Professors Guthrie and Rachlinski, along with Federal Magistrate Andrew Wistrich.[259] They surveyed 200 magistrate judges attending a federal judicial conference and found that the judges' responses indicated a susceptibility to various heuristics and biases (anchoring, framing, hindsight bias, representativeness heuristic, and the egocentric bias)[260] that was consistent with laboratory results involving students and others.

Mitchell's attack, and it is one fairly lodged against any laboratory experiment in this format, is that the judges' answers, although they appeared irrational when judged against the normative response, may have been rational if the judges' main goal was simply to complete the questionnaire as quickly as possible so that they could get out to the golf course.[261] The problem with Mitchell's argument is that the judges' responses were not random; they were systematically biased in a way consistent with other studies done involving other subjects. Hastie and Viscusi's study involving judges found the same effect.[262] Because laboratory results have repeatedly been supported by findings in real world settings,[263] it is unlikely that there is any particular problem in most psychological surveys of subjects having as their main goal finishing the survey quickly.[264]

---

257. Alan Farnsworth, *On Trying to Keep One's Promises: The Duty of Best Efforts in Contract Law*, 46 U. PITT. L. REV. 1, 44 (1984); *see also* Robert A. Hillman & Jeffrey J. Rachlinski, *Standard-Form Contracting in the Electronic Age*, 77 N.Y.U. L. REV. 429, 444 (2002) ("Businesses often delegate the job of drafting [contract] terms to lawyers, who believe that they can best serve their client by composing an arsenal of one-sided terms without regard to the business environment, or for that matter, anything else.").

258. *See* Robert A. Prentice, *Contract-Based Defenses in Securitics Fraud Litigation: A Behavioral Analysis*, 2003 U. ILL. L. REV. 337, 341 (2003); *see also infra* notes 491-503 and accompanying text.

259. Guthrie et al., *supra* note 10, *cited in* Mitchell, *Pessimism*, *supra* note 12, at 1996-97.

260. *Id.* at 778 ("Judges, it seems, are human. Like the rest of us, their judgment is affected by cognitive illusions that can produce systematic errors in judgment.").

261. *See* Mitchell, *Pessimism*, *supra* note 12, at 1999.

262. Hastie & Viscusi, *supra* note 66, at 917 (finding "massive" hindsight bias by jurors and lesser but still substantial hindsight bias by judges).

263. *See supra* notes 169-195 and accompanying text.

264. Even college students, who are more likely to be irresponsible in participating in such an experiment than a federal judge, have generally been found to be "a fairly good proxy for 'real people.'" Guthrie, *supra* note 1, at 1156 ("[S]everal studies have found that experts display roughly the same biases as college students or the same biases at somewhat reduced levels." (citations omitted) (quoting PLOUS, *supra* note 135, at 258)).

2. Individual Decision Making Versus Group Decision Making

Next, Mitchell points out that the individual unit of analysis in most behavioral studies is the individual, whereas many legal decisions are made by groups, such as panels of judges and juries.[265] This is true, and many psychologists wish to study these dynamics and understand when and under what circumstances group deliberation will change decision making. Unlike psychologists, economists generally are not interested in studying differences between individual and group decision making. One of the biggest flaws of economic reasoning has been its "extraordinary bias towards individualism. Most economists tend to assume not just that we make decisions in isolation, but that we nearly always should do so."[266] Moreover, the psychology studies done to date "support the conclusion that group deliberation is unlikely to remedy the effects of a powerful individual judgment bias, like [for example] the hindsight" bias.[267] Many studies find little or no difference between individual and group deliberations.[268] The anchoring bias, for example, has been demonstrated in group decision making as well as individual decision making.[269] The over-optimism and self-efficacy biases have been

265. Mitchell, *Pessimism, supra* note 12, at 2002-05.

266. Michael Prowse, *The Psychology Behind Prizes for a Dismal Science,* FIN. TIMES, Oct. 19/20, 2002, at II (comparing psychological and economic approaches); *see also* Stephen M. Bainbridge, *Why a Board? Group Decision Making in Corporate Governance,* 55 VAND. L. REV. 1, 2 (2002) ("Economic analysis tends to focus on the decisions of individuals. This emphasis likely stems from the underlying model of rational choice, which posits an autonomous individual who makes rational choices that maximize his satisfactions.").

267. Reid Hastie & W. Kip Viscusi, *Juries, Hindsight, and Punitive Damages Awards: Reply to Richard Lempert,* 51 DEPAUL L. REV. 987, 992 (2002). *See generally* Terry Connolly & Edward W. Bukszar, *Hindsight Bias: Self-Flattery or Cognitive Error?,* 3 J. BEHAV. DECISION MAKING 205, 205 (1990) (noting that 122 studies of the hindsight bias show that it is "robust to variations in method, population, and task"); Hastie & Viscusi, *supra* note 66, at 917 (reporting results of a study finding "massive" hindsight bias by juries); Norbert L. Kerr et al., *Bias in Judgment: Comparing Individuals and Groups,* 103 PSYCHOL. REV. 687 (1996) (reviewing the literature and finding no evidence that juries are less biased than individual jurors); Dagmar Stahlberg et al., *We Knew It All Along: Hindsight Bias in Groups,* 63 ORGANIZATIONAL BEHAV. & HUM. DECISION PROCESSES 46, 56 (1995) (finding that "groups are just as prone to hindsight bias as individuals when making hypothetical predictions" but that groups are better at recalling previous judgments).

268. *See, e.g.*, William C. Thompson et al., *Jurors' Sensitivity to Variations in Statistical Evidence* 15-20 (unpublished manuscript), *cited in* Rowe, *supra* note 84, at 547 n.10 (discussing a base rate experiment that found that group decision process did not lessen subjects' insensitivity to strength of statistical evidence).

269. Robert W. Rutledge, *The Effects of Group Decisions and Group-Shifts on Use of the Anchoring and Adjustment Heuristic,* 21 SOC. BEHAV. & PERSONALITY 215, 224 (1993) (finding that groups are susceptible to anchoring effects in a manner similar to individuals).

shown to be even greater in groups than in individuals,[270] as has the overconfidence bias.[271]

Sunstein recently summarized the evidence on the effectiveness of groups at improving decision making:

> Are groups able to avoid the judgment errors made by individuals? The evidence is mixed. In general, groups tend to polarize: they tend [to end up] in a more extreme position in line with their predeliberation tendencies. At the same time, groups have been found to make better decisions than individuals with respect to certain statistical problems. There is some evidence that groups are slightly better at avoiding the problems created by use of the availability heuristic. On the other hand, some evidence suggests that the use of the representativeness heuristic is actually amplified. It seems clear that group processes do not eliminate the use of heuristics, and it remains to be found whether and when they reduce or increase the resulting errors.[272]

Similarly, in 1996 Kerr and his colleagues reviewed all the studies they could find regarding differences in individual and group bias, and concluded that there is little difference between group and individual bias.[273] The differences that did exist did "not show a simple, consistent pattern of relative bias."[274] Kerr and colleagues noted the problem that this poses for Chicago Man advocates who suggest that collective decision making should cancel out judgmental errors:

> Though this may be correct for aggregate public opinion, it is premised on a statistical analogy—the law of large numbers—that is clearly incompatible with actual interactive group decision making under some likely social decision schemes. . . . More important, this argument does not apply to judgment biases . . . which are systematic rather than random. At best, our analyses offer an existence proof that collective rationality can sometimes be superior to individual rationality, but they also suggest that over a large and plausible region of relevant parameter space, group decision making actually exacerbates the biases observed in individual decisions.[275]

3. Framing Research Confounds and Confuses

In his next point, Mitchell returns to familiar ground, taking yet another swipe at the phenomenon of framing effects.[276] Perhaps

---

270. *See, e.g.*, Chip Heath & Forest J. Jourden, *Illusion, Disillusion and the Buffering Effect of Groups*, 69 ORGANIZATIONAL BEHAV. & HUM. DECISION PROCESSES 103, 114 (1997) ("[G]roups do not increase positive evaluations so much as they preserve them.").

271. *See* Chip Heath & Richard Gonzales, *Interaction with Others Increases Decision Confidence But Not Decision Quality: Evidence Against Information Collection Views of Interactive Decision Making*, 61 ORGANIZATIONAL BEHAV. & HUM. DECISION PROCESSES 305, 322 (1995) (finding that when subjects interacted with others, their confidence stayed stable or increased, but their decision making did not improve).

272. CASS R. SUNSTEIN, HAZARDOUS HEURISTICS 14 (Univ. of Chi., Olin Law & Econ. Working Paper No. 165 (2d Series), Oct. 2002), http://ssrn.com/abstract_id=344620.

273. Kerr et al., *supra* note 267, at 713.

274. *Id.*

275. *Id.* at 713-14.

276. Mitchell, *Pessimism*, *supra* note 12, at 2002-11. Framing effects occur, for example, when people's responses, perceptions, or preferences change simply because of a simple

framing effects are Mitchell's favorite whipping boy because of all of the heuristics and biases that have been studied in the K-T tradition, framing effects have proved to be the most complex. Psychologists have demonstrated that they are indeed an entire family of effects.[277] Mitchell accurately notes that

> whereas this research is typically summarized as revealing that people are risk averse on positively framed problems and risk seeking on negatively framed problems, 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 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.[278]

Mitchell then proceeds to lay out the complications in fair and accurate detail, concluding that "there is no universal framing effect that can be easily translated into legal doctrine."[279] His point is a strong one and deserves but two responses. First, his own source clearly demonstrates that although it has a complicated nature, the framing effect is "a reliable phenomenon,"[280] and economists' and legal commentators' theories will suffer if they ignore it. Second, legal decision scholars seem to be aware of the complications with framing theory and none, to my knowledge (and Mitchell points out none)[281] has offered a policy prescription based on an admittedly faulty assumption that framing effects are simple, consistent, and universal.[282]

---

reframing of the question, outcome or situation. *See* Christopher C. Fennell & Lee Anne Fennell, *Fear and Greed in Tax Policy: A Qualitative Research Agenda,* 13 WASH. U. J.L. & POL'Y 75, 87 (2003) (noting that "[r]esearch indicates that the surrounding context or 'frame' can be extremely important in processing or 'coding' an event."); Richard Zeckhauser & W. Kip Viscusi, *Risk Management Strategies: The Risk Management Dilemma,* 545 ANNALS 144, 149 n.7 (1996) (noting that "'[t]he framing of risk problems has a considerable effect on how risks are viewed and what preferences are expressed.").

277. *See, e.g.,* 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) (identifying three categories of framing—risky choice framing, attribute framing, and goal framing—and finding that "[w]ithin each framing type, results show substantial consistency").

278. Mitchell, *Pessimism, supra* note 12, at 2005-06.

279. *Id.* at 2011.

280. Kuhberger, *supra* note 136, at 23.

281. Every footnote in this section to Mitchell's paper refers to a psychology journal. None references a law journal containing a policy prescription.

282. Behavioralists virtually always refer to framing effects, accurately, as "systematic" rather than as universal. *See, e.g.,* Timur Kuran & Cass R. Sunstein, *Availability Cascades and Risk Regulation,* 51 STAN. L. REV. 683, 705 (1999); Howard Latin, *"Good" Warnings, Bad Products, and Cognitive Limitations,* 41 UCLA L. REV. 1193, 1237 (1994).

4. Neglect of Systematic Information-Processing Modes

In the final section of his methodology article, Mitchell chides legal decision theorists for not having embraced with due vigor the latest theories regarding dual processing of information. Mitchell appropriately calls attention to this exciting research.[283] Indeed, Kahneman's latest book addresses this research in great detail.[284] The key notion, expressed in slightly different ways by competing schools of thought, is that the human mind has two processing systems that can operate simultaneously. What is often termed "System 1" is an intuitive system that works automatically, rapidly, and effortlessly. We often use this system without even realizing it. System 2 is a reflective system that is controlled, effortful, deductive, and self-aware.[285] In the anchoring and adjustment phenomenon, it is System 1 that automatically anchors on a number given, and System 2 that consciously adjusts away from that anchor (but usually not far enough).[286] System 1 automatically believes everything it is told; System 2 adjusts for known facts that might qualify that belief.[287]

While Mitchell cannot be faulted for insisting that legal decision theorists keep up to date with the latest research in psychology, some have already incorporated this work into their writings.[288] Moreover, he once again offers not a single policy prescription by any legal decision theorist that he believes is undermined by an insufficient emphasis on dual processing modes of decision making. Ultimately, Mitchell suggests that we should "(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

283. Mitchell, *Pessimism, supra* note 12, at 2011-17.

284. *See* PSYCHOLOGY OF INTUITIVE JUDGMENT, *supra* note 8.

285. Kahneman & Frederick, *supra* note 72, at 51. *See generally* KEITH E. STANOVICH, WHO IS RATIONAL? STUDIES OF INDIVIDUAL DIFFERENCES IN REASONING 144-47 (1999) (introducing a dual-processing model); 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) (giving background on these theories); 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 (2000) (sketching a summary of dual process models).

286. Kahneman & Frederick, *supra* note 72, at 51.

287. *Id.* at 57.

288. *See, e.g.*, Paul B. Marrow, *The Unconscionability of a Liquidated Damage Clause: A Practical Application of Behavioral Decision Theory,* 22 PACE L. REV. 27, 54-57 (2001) (describing a prominent dual processing theory); Paul Slovic, *Rational Actors and Rational Fools: The Influence of Affect on Judgment and Decision-Making,* 6 ROGER WILLIAMS U. L. REV. 163, 167, 182 (2000) (emphasizing the interplay of affect and dual-process theories of information processing); Cass R. Sunstein, *Probability Neglect: Emotions, Worst Cases, and Law,* 112 YALE L.J. 61, 84-85 (2002) (describing "system 1" versus "system 2" thinking).

legal contexts, and for whom and (2) determine whether active interventions within the legal setting can trigger the rationale mode of thought if the benefits of such interventions outweigh their costs."[289] It is psychologists in the K-T Man tradition, not Chicago Man economists, who hold the promise of developing these theories and making these determinations.

D. Final Insights

Empirical psychological research is complex, and Mitchell rightly calls attention to its limitations. Mitchell's two main points in the first article appear to be that (a) psychology is a questionable science whose methods systematically overstate limitations on human judgment, and (b) legal decision theorists have inexpertly applied its precepts to legal doctrine. He overstates both points.

Regarding the methods and manners of the psychology discipline, Mitchell cites Krueger for the proposition that "[i]nvestigators demonstrate bias by detecting it. They rarely attempt to detect rational judgment."[290] But this is not true. As Mitchell demonstrates with his articles, in any academic field a good gateway to publication is to take a contrarian position, to attempt to discredit an established theory. Psychology professors could make their careers by discrediting the main themes of K-T research. Economists,[291] game theorists,[292] philosophers,[293] and others have done their level best to discredit most of the heuristics and biases attributed to K-T Man.

---

289. Mitchell, *Pessimism, supra* note 12, at 2017.

290. *Id.* at 1955 (quoting Joachim Krueger, *The Bet on Bias: A Foregone Conclusion?,* 9 PSYCOLOQUY 1, ¶ 4 (1998), *at* http://www.cogsci.soton.ac.uk/cgi/psyc/newpsy?9.46).

291. For example, after Tversky and Kahneman and Lichtenstein and Slovic demonstrated the phenomenon of preference reversals, which are inconsistent with the assumption that Chicago Man's preferences are intransitive, Pommerehne and colleagues and Reilly attempted to debunk the theory. *Compare* Sarah Lichtenstein & Paul Slovic, *Reversals of Preference Between Bids and Choices in Gambling Decisions,* 89 J. EXPERIMENTAL PSYCHOL. 46 (1971), *and* Tversky & Kahneman, *supra* note 135, *with* Werner Pommerehne et al., *Economic Theory of Choice and the Preference Reversal Phenomenon: A Reexamination,* 72 AM. ECON. REV. 569 (1982), *and* Robert Reilly, *Preference Reversal: Further Evidence and Some Suggested Modifications in Experimental Design,* 72 AM. ECON. REV. 576 (1982). They failed. *See generally* Paul Slovic & Sarah Lichtenstein, *Preference Reversals: A Broader Perspective,* 73 AM. ECON. REV. 596, 597 (1983).

292. *See* David Grether & Charles Plott, *Economic Theory of Choice and the Preference Reversal Phenomenon,* 69 AM. ECON. REV. 623, 632 (1979) (discovering that the introduction of incentives actually strengthened preference reversals, contrary to the authors' original intention to demonstrate that incentives would make preference reversals disappear).

293. L. Jonathan Cohen, *Can Human Irrationality Be Experimentally Demonstrated?,* 4 BEHAV. & BRAIN SCI. 317 (1981).

Furthermore, prominent and incredibly productive psychologist Gerd Gigerenzer and his colleagues in Germany spent a decade attempting to discredit Kahneman and Tversky's work. They argued that most of the "defects" identified in human reasoning were actually "fast and frugal" heuristics that were incredibly efficient.[294] After a decade of critiques that, to Gigerenzer's credit, forced researchers in the Kahneman and Tversky camp to improve their methodologies, tweak their theories, and make some concessions, Gigerenzer is now publishing books featuring his own list of reasoning limitations (such as "illusory certainty") and emphasizing man's weakness in dealing with probabilities.[295]

Mitchell's own sources emphasize that psychology is a respected science with research that produces results that are often as reliable as those of medical science and that uses many of the same methods as the hard sciences.[296] Research results in psychology are roughly as consistent as those in physics,[297] and as reliable as many prominent findings in medical science.[298] Stanovich, one of Mitchell's

---

294. *See generally* BOUNDED RATIONALITY: THE ADAPTIVE TOOLBOX (Gerd Gigerenzer & Reinhard Selten eds., 2001); GERD GIGERENZER, ADAPTIVE THINKING: RATIONALITY IN THE REAL WORLD (2000); GERD GIGERENZER ET AL., SIMPLE HEURISTICS THAT MAKE US SMART (1999) [hereinafter GIGERENZER ET AL., SIMPLE HEURISTICS].

295. *See* GERD GIGERENZER, CALCULATED RISKS: HOW TO KNOW WHEN NUMBERS DECEIVE YOU 14, 37 (2002); *see also* Eldar Shafir, *Intuitions About Rationality and Cognition, in* RATIONALITY, *supra* note 165, at 260, 279 ("Arguments about natural selection and adaptation notwithstanding, some of our ways of making decisions may be truly 'maladaptive.'").

It is fair to say that while Kahneman and Tverksy studied humans' cognitive errors at the expense of their cognitive successes, Gigerenzer and colleagues did just the opposite, with considerable success. *See* GIGERENZER ET AL., SIMPLE HEURISTICS, *supra* note 294, at 217-18.

296. *See* STANOVICH, *supra* note 64, at 114 (arguing that psychology is a respected science); *see also* Larry Hedges, *How Hard Is Hard Science, How Soft Is Soft Science?*, 42 AM. PSYCHOL. 443. 450 (1987) ("There is substantial support for the contention that rigorous reviews of some kinds of social science research reveal very consistent results.").

297. *See* Hedges, *supra* note 296, at 451 (concluding after an examination of groups of studies in both physical sciences and social sciences that "research results in the physical sciences are not markedly more consistent than those in the social sciences . . . [and] [t]he notion that experiments in the social sciences produce relatively inconsistent (empirically noncumulative) results is not supported by these data either"); Henrion & Fischhoff, *supra* note 179, at 666-67 ("The underestimation of uncertainty of physical constants and compilations of recommended values seems to be pervasive."); Ronald C. Serlin & Daniel K. Lapsley, *Rational Appraisal of Psychological Research and the Good-Enough Principle, in* HANDBOOK FOR DATA ANALYSIS, *supra* note 98, at 199, 225 (showing that research in psychology and physics share many strengths and limitations, and concluding that "hypothesis-testing in psychology, when fortified by the good-enough principle, is not rationally disadvantaged when compared against hypothesis testing in physics").

298. *See* Robert Rosenthal, *How Are We Doing in Soft Psychology?*, 45 AM. PSYCHOLOGIST 775, 775-76 (1990) (noting several examples in which medical science has made recommendations for action based on statistical results no stronger than those typically found in social science research).

most frequent sources, notes that "findings in cognitive psychology have met the basic test of replicability. Many of the fundamental laws of information processing have been observed in dozens of laboratories all over the world."[299]

As Tetlock, one of Mitchell's favorite sources, notes, research programs on judgment and choice "have been phenomenally successful, triggering an avalanche of discoveries of when judgment and choice deviate from conventional standards of scientific or economic rationality."[300] More to the point, Tetlock urges:

> It would be curmudgeonly—and, even worse, wrong—to deny that both research programs pass classic philosophy-of-science tests of knowledge advancement, an all-too-rare achievement in the behavioral and social sciences. A small set of explanatory constructs—judgmental heuristics, the framing of outcomes, the psychophysics of gain and loss functions—organizes a vast array of findings and stimulates falsifiable hypotheses that hold up in an impressive array of settings.[301]

In other words, to some extent we are now arguing about details. Mitchell essentially admits this. He does not argue that the Chicago Man model in any way approximates how people actually act.[302] He does not reject the psychological analysis of law.[303] It is the details he quibbles over, and the scientific work of nailing down those details will be ongoing for a long, long time. In the meantime, one can argue plausibly that current methods *understate*, rather than overstate (as Mitchell claims), limitations on human judgment and decision making.[304]

Regarding the physical sciences, keep in mind that the scientists doing work in those areas are subject to the heuristics and biases that affect everyone else, including overconfidence, anchoring and adjustment, the confirmation bias, and others. Henrion and Fischhoff recently noted the role that the psychology of buman decision making, particularly overconfidence, has played into disconcertingly large errors made over the years in measurements of physical quantities, such as the velocity of light, Planck's constant, or the rest mass of the electron. *See* Henrion & Fischhoff, *supra* note 179, at 666.

299. STANOVICH, *supra* note 64, at 114. Stanovich also notes that even in social psychology, where some of the basic criticisms of psychology research seem most apt, "evidence has indicated that the laboratory-derived relationships and theories do in fact predict behavior in a variety of other situations involving different types of individuals." *Id.* at 117.

300. Tetlock, *supra* note 216, at 582.

301. *Id.*

302. Mitchell, *Pessimism, supra* note 12, at 1936 ("[M]y 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.").

303. *Id.* at 1937 ("[M]y criticisms also should not be understood as a rejection of the psychological analysis of law.").

304. ERIC VAN DEN STEEN, SKILL OR LUCK? BIASES OF RATIONAL AGENTS 4, 17 (MIT Sloan Sch. of Mgmt., Working Paper No. 4255-02, June 2002) ("[E]xperimental results often *underestimate* the true impact of behavioral biases. . . . [E]xperiments that test for the self-serving bias typically restrict the actions the subjects can take. In everyday life, people have much more freedom. This implies that such structured experiments will tend to under-estimate

Regarding the import of psychological evidence to legal analysis, Mitchell is unable to credibly undermine a single specific policy proposal by a legal decision theorist.[305] This is not because the heuristics and biases literature is without faults; it is because legal decision theorists are generally aware of its faults and take them into account in making policy prescriptions. Indeed, in his more candid moments, Mitchell admits that "legal decision theorists generally express caution about their endeavor and note the preliminary nature of much of their work. . . ."[306]

---

the practical relevance of these biases." (emphasis added)), http://papers.ssrn.com/sol3/papers.cfm?abstract_id=319972.

Tversky and Kahneman note that

> the conjunction error is only a symptom of a more general phenomenon: People tend to overestimate the probabilities of representative (or available) events and/or underestimate the probabilities of less-representative events. The violation of the conjunction rule demonstrates this tendency even when the "true" probabilities are unknown or unknowable. *The basic phenomenon may be considerably more common than the extreme symptom by which it was illustrated.*

Tversky & Kahneman, *supra* note 93, at 45 (emphasis added); *see also* Guthrie, *supra* note 1, at 1158 ("[P]eople may be more likely to rely on framing and other cognitive shortcuts when they are confronted with complicated rather than straight-forward decisions . . . suggest[ing] that framing and other phenomena of this sort might have a greater impact on real-world decision making than on simplified laboratory decision making.").

305. The closest Mitchell comes to a direct attack on a particular policy prescription in his *Pessimism* article is when he challenges the evidence regarding the hindsight bias. Mitchell, *Pessimism*, *supra* note 12, at 1933-35 n.45. While he is correct in concluding that there exists as yet no conclusive evidence that jurors deliberating as a group are affected by the hindsight bias, we should recall that (a) the evidence that individual decision makers are infected by the hindsight bias is overwhelming, Dan L. Burk & Mark A. Lemley, *Is Patent Law Technology-Specific?*, 17 BERKELEY TECH. L.J. 1155, 1198 n.185 (noting that the hindsight bias is "well-documented"); (b) most biases infect individual and group decision makers similarly, *see supra* notes 267-275 and accompanying text; (c) there is no evidence that the hindsight bias does not infect group decision makers and some evidence that it does, *see supra* note 267 and accompanying text; and (d) one of Mitchell's favorite sources in the area, *see* Mitchell, *Pessimism*, *supra* note 12, at 1931-32 nn.41-43, admits that "[g]iven the seeming ubiquity of [the hindsight] bias in human decision making, it is likely [that juries] do [exhibit the hindsight bias]." Lempert, *supra* note 66, at 881.

Although I agree with Lempert that we should be hesitant to modify any long-standing policy based on a single study, the law has long recognized the existence of the hindsight bias and tried to manage it with doctrines such as the business judgment rule in corporate law, *see* Arkes & Schipani, *supra* note 10, at 587, and the rule against allowing admission of evidence regarding subsequent repairs in products liability cases. Even Judge Posner admits that part of the reason for this rule of evidence might be the hindsight bias, although naturally he attempts to present an economic rationalization. *See* Posner, *Evidence*, *supra* note 34, at 1545 ("The subsequent-repairs rule may also be justified by concerns with hindsight bias, but these concerns seem exaggerated and in any event could be dealt with by other measures.").

306. Mitchell, *Pessimism*, *supra* note 12, at 1933. In another passage, Mitchell argues that legal decision theorists speak in blanket terms and fail to recognize studies that qualify broad conclusions about human reasoning. *Id.* at 1944. He makes particular reference to an article on base rates, Koehler, *supra* note 97, at 1, that legal decision theorists such as Guthrie, Rachlinski and I have cited in our work, rather than ignoring it. *See* Guthrie et al., *supra* note 10, at 806

This certainly does not mean that errors have not crept into the policy precepts of legal decision theorists, or that they will not do so. But it is the proponents of K-T Man, rather than those of Chicago Man, who have a fighting chance of usefully formulating legal doctrine based on how people actually make decisions. Nowhere in Mitchell's articles does he indicate that he would disagree with Pouncy's recent assessment that "[t]he work of Daniel Kahneman and Amos Tversky convincingly demonstrated that the rational choice model of human motivation was at best grossly incomplete, and at worst, simply wrong."[307]

IV. HOW MUCH INDIVIDUAL VARIATION EXISTS IN HUMAN REASONING AND WHAT ARE ITS IMPLICATIONS?

Mitchell's second article, his *Equal Incompetence* article,[308] erects and tears down a giant straw man. He claims that while economists err by assuming that man is always rational, legal decision theorists err by assuming that man is always irrational. Contrary to this claim of assumed "equal incompetence" on behalf of all people in all situations, legal decision theorists recognize individual and situational variations. Some of Mitchell's own cited sources make this explicit by stating that the psychological evidence shows that people "frequently"[309] or "often"[310] think in ways that depart from accepted norms of rationality.[311]

---

n.135; Prentice, *supra* note 36, at 159 n.147; Rachlinski, *Heuristics and Biases, supra* note 10, at 85 n.117.

307. Pouncy, *supra* note 6, at 302; *see also* Dailey, *supra* note 5, at 1603 ("The effort to revise the economic model of human decisionmaking with findings from cognitive psychology is an especially important development in law, where much of the behavior under study has already, almost by definition, failed the traditional test of rationality . . . ."); Pouncy, *supra* note 6, at 308 ("The fact that rationality, as a decisional heuristic, cannot meaningfully explain much of human behavior has been recognized by at least two generations of anthropologists and a generation of psychologists, but has only recently begun to be acknowledged in doctrinal analysis and in law and economics jurisprudence.").

308. Mitchell, *Incompetence, supra* note 12.

309. *Id.* at 2 n.2 ("Psychologists who study human judgment and choice have learned that people *frequently* fall prey to cognitive illusions that produce systematic errors in judgment." (emphasis added) (citing Guthrie et al., *supra* note 10, at 777)).

310. *Id.* ("Actual judgments show systematic departures from models of unbiased forecasts, and actual decisions *often* violate the axioms of expected utility theory." (emphasis added) (citing Christine Jolls et al., *A Behavioral Approach to Law and Economics,* 50 STAN. L. REV. 1471, 1477 (1998))).

311. *See, e.g.*, James A. Fanto, *Quasi-rationality in Action: A Study of Psychological Factors in Merger Decision-Making,* 62 OHIO ST. L.J. 1333, 1344 (2001) (stating that while applying behavioral literature in an analysis of takeover activity, "people act quasi-rationally, not rationally nor irrationally").

Mitchell cites to one of my articles in which I responded to a line of decisions stemming from law and economics scholars/judges who assumed that auditors would not do anything illegal because to do so would be irrational in light of the damage it would do to their reputations if uncovered.[312] In addition to arguing that it could, in fact, be arguably rational for auditors to audit recklessly,[313] I made the general point that people often think heuristically and are often subject to a catalogue of biases that can lead to conclusions that are normatively irrational.[314] In that article, I was very careful to point out where there was evidence that trained auditors were less susceptible to various biases than were the general run of people,[315] and where there was evidence that they were equally or even more susceptible.[316]

In another article that dealt with protecting investors from fraud, I evaluated a proposal that essentially would allow sophisticated investors to opt out of any governmental protection from securities fraud.[317] Again, I examined a number of heuristics and biases that potentially lead investors away from the rational path when they make investment decisions.[318] I was very sensitive to the argument that experienced investors might not be as susceptible to reasoning errors as inexperienced investors, so I repeatedly cited

312. Prentice, *supra* note 36, at 133.

313. *Id.* at 199-217.

314. *Id.* at 139-81.

315. *Id.* at 146 (citing studies showing that auditors may not be as subject to the confirmation bias as other people, although noting contra studies); *id.* at 151-52 (citing studies showing that experienced auditors may commit fewer errors of recall than accounting students, but also noting contrary studies); *id.* at 161 (citing studies indicating that auditors are generally better than others at distinguishing between more objective and less objective sources of information); *id.* at 167 (citing studies tending to show that auditors' natural conservatism tends to minimize some of the adverse effects of their behavioral biases).

316. *Id.* at 145 (citing studies showing that auditors use rule of thumb heuristics like everyone else and that they often act inconsistently with Bayesian notions of probability); *id.* at 148-49 (citing studies tentatively indicating that trained auditors are subject to the hindsight bias just like lay people); *id.* at 154-55 (citing several studies indicating that auditors are generally subject to the overconfidence bias, but noting a few contrary findings); *id.* at 156-57 (citing studies showing that auditors' judgment can be misled by framing effects); *id.* at 157 (citing studies showing that "[a]uditors . . . are not much letter than lay people at calculating probabilities" and tend to act inconsistently with Bayes' theorum); *id.* at 165 (citing studies showing that even expert auditors are subject to the anchoring and adjustment bias); *id.* at 167 (citing studies showing that auditors, like others, are influenced by the order in which they process information); *id.* at 169 (citing several studies showing that accountants are affected by the self-serving bias).

317. Robert A. Prentice, *Whither Securities Regulation? Some Behavioral Observations Regarding Proposals for Its Future,* 51 DUKE L.J. 1397 (2002).

318. *Id.* at 1454-89.

studies showing that experienced investors are often (not always) equally susceptible to such biases.[319]

After recounting numerous behavioral weaknesses in human cognition, I made the point that Mitchell insists legal decision theorists ignore:

> Just as a thorough recounting of an ex-spouse's faults ultimately makes him sound worse than he really was, the totality of the foregoing description of behavioral research as it applies to investors makes them sound more dunderheaded than they usually are. *Certainly not all heuristics and biases apply at all times and in all settings to all investors.*[320]

When Thomas Ulen notes that the "central distinguishing contention" of the legal decision theory movement is that people are "imperfectly rational,"[321] Mitchell seems to read it as a claim that people are "perfectly irrational," and this simply is not the case.[322] As Tversky and Kahneman themselves recently wrote, the heuristics and biases literature "neither assumes nor entails that people are perceptually or cognitively inept."[323] Because legal decision scholars are necessarily sensitive to the fact that the complex human reasoning and judgment processes are anything but universal and uniform, Mitchell is unable in either of his articles to meaningfully challenge a single policy prescription by any legal decision scholar.

Nonetheless, Mitchell's discussion advances the debate about human decision making and he is no doubt correct when he points out that

---

319. *Id.* at 1455 (noting that research shows that even professional securities analysts do not choose to acquire most information that is available to them and are thus creatures of bounded rationality); *id.* at 1460-61 (noting studies showing that educated people, professionals in general, and stock analysts in specific tend to be victimized by the overconfidence bias); *id.* at 1464-65 (referring to evidence showing that most people have difficulty telling when they are being deceived and noting that scams work on all classes of people); *id.* at 1469 (citing evidence that even sophisticated investors tend to credit oral communications more than written communications and therefore pay insufficient attention to legal disclaimers contained in the fine print); *id.* at 1470 (citing studies indicating that professional investors are subject to the availability bias); *id.* at 1471 (citing studies showing that professional investors use the representativeness heuristic); *id.* at 1476 (citing examples indicating that the phenomenon of social proof affects sophisticated investors as well as lay investors); *id.* at 1482 (citing study showing that financial models assembled by sophisticated investors often improperly ignore low probability events); *id.* at 1483 (citing studies showing that securities analysts are subject to the anchoring and adjustment phenomenon).

320. *Id.* at 1489 (emphasis added).

321. Thomas S. Ulen, *Evolution, Human Behavior, and Law: A Response to Owen Jones's Dunwody Lecture*, 53 FLA. L. REV. 931, 933 (2001).

322. Most behavioral decision theorists would agree with Piattelli-Palmarini's statement that "[i]rrational, and even exaggeratedly so, we frequently are, but not *necessarily*." PIATTELLI-PALMARINI, *supra* note 65, at 40. *See generally* Mitchell, *Incompetence*, *supra* note 12.

323. Tversky & Kahneman, *supra* note 93, at 47.

> [D]ifferences in education, training, cognitive capacity, thinking dispositions, sex, and cultural background *across* individuals appear to be reliably associated with different levels of cognitive performance. Furthermore, emotional differences, developmental differences, and different modes of mental processing appear to be associated with different levels of cognitive performance *within* individuals. Therefore, depending on the characteristics of the individual and the system of thought activated in a particular decision making situation, the behavior of different groups of individuals and the behavior of the same individual may very considerably, from perfect rationality to seeming irrationality.[324]

Still, people are not always rational,[325] and Mitchell does not claim that they are. He is correct in pointing out that, conversely, people are not always irrational. The question is whether departures from rationality are sufficiently systematic to be useful in making legal policy.[326] When behavioral research was first imported into the economics literature, many economists argued that the various departures from rationality that were increasingly being documented were sufficiently random that they would cancel each other out and therefore not require any refinement of the rational man assumption.[327] However, evidence that many of these heuristics and biases are systematic rather than random is overwhelming.[328]

---

324. Mitchell, *Incompetence, supra* note 12, at 87 (emphasis added).

325. For the few scientists who have argued the "Panglossian notion that people's judgments are hardly ever biased," Gilovich and Griffin point out that they themselves "use a variety of methodological safeguards such as double-blind experimental designs to make sure their own observations are not contaminated by bias," and ask, "Are the observations of scientists so much more prone to bias than the individuals they study?" Gilovich & Griffin, *supra* note 158, at 9.

326. An analogy may be drawn to the efficient market debate. It has become obvious over the years that investors are not always rational. Rather, they are subject to all the heuristics and biases that Tversky, Kahneman, and others have discovered. *See* SHILLER, supra note 209, at 135-68 (explaining how investors fall subject to various cognitive fallacies). Nonetheless, efficient market advocates have argued that the roughly equal prevalence of under- and over-reactions demonstrated by investors in the empirical research implies that market prices are informationally efficient. *See, e.g.*, Eugene F. Fama, *Market Efficiency, Long-Term Returns and Behavioural Finance,* 49 J. FIN. ECON. 283 (1998). However, Bloomfield and others recently noted that "[t]his claim is correct only if we are unable to predict the situations in which we will observe under- and overreactions. Our experiments show that we can make such predictions in the laboratory, simply by knowing the reliability of investors' information." Robert Bloomfield et al., *Underreactions, Overreactions and Moderated Confidence,* 3 J. FIN. MARKETS 113, 130 (2000).

327. *See, e.g.,* RICHARD A. POSNER, ECONOMIC ANALYSIS OF LAW 19 (5th ed. 1998) ("Economics is concerned with explaining and predicting tendencies and aggregates rather than the behavior of each individual person; and in a reasonably large sample, random deviations from normal rational behavior will cancel out.").

328. One of Mitchell's primary lines of argument is predicated on the work of Stanovich and West that tends to show that certain types of problems with "irrationality" in the heuristics and biases literature are not universal in that people with higher cognitive ability (smarter) or with better cognitive styles will not be as subject to the studied biases. *See* Mitchell, *Incompetence, supra* note 12, at 142-47. It is well to remember that in a recent book summarizing that line of studies and a huge amount of behavioral decision literature, Stanovich concluded that errors tended to be systematic rather than random. STANOVICH, *supra* note 285, at 252; *see also id.* at 48 ("As an explanation of the repeated failure of subjects in the heuristics and biases literature to display normatively appropriate behavior, a strong version of the performance error view does

Mitchell himself does not seem to believe that the heuristics and biases that limit man's reasoning ability are random. Indeed, in introducing the above-quoted paragraph, he states that "a growing body of empirical research demonstrat[es] that individuals vary widely, *and predictably*, in their propensities to act rationally."[329] If these variations are predictable, then there is reason to believe that behavioral and cognitive research can lead to fruitful policy prescriptions.

A. Individual Differences in Rational Behavior

1. Differences Across Individuals

a. Education

To examine Mitchell's specific arguments, consider first his point that education and training can improve people's normative reasoning performance. Based on that contention, Mitchell states, "when the legal decision theorists tell their audience that the educated and uneducated equally fall prey to cognitive illusions, they are wrong."[330] Mitchell cites no legal decision theorist who has told his or her audience that educated and uneducated people *always* fall prey to the same illusions, and I cannot imagine that one has done so. Nor does Mitchell cite any legal decision scholar who has based a policy prescription upon such an assumption.

Mitchell could have cited several theorists who have claimed that educated and uneducated people *usually* fall prey to the same illusions at roughly the same rates, and Mitchell would have great difficulty discrediting such statements. For example, there are studies indicating:

- That knowledge can be a cause, rather than a cure, for bias, as experienced subjects "are prone to 'illusory correlations'—the perception of patterns in random data that conform with *a priori* theories."[331]

---

not seem tenable. The deviations from normative responding have repeatedly been shown to be systematic."); Blumenthal, *supra* note 1, at 42 (noting, in regard to the economic argument that variations from rationality are random and will cancel each other out, that "empirical research shows this to be unlikely at best, false at worst").

329. Mitchell, *Incompetencc, supra* note 12, at 86 (emphasis added).

330. *Id.* at 87.

331. Evans, *supra* note 165, at 26.

- That statistical experts are sometimes as prone to committing the conjunction fallacy as lay people.[332]
- That increased knowledge often correlates with *increased* overconfidence.[333]
- That trained physicians are highly subject to framing effects.[334]
- That "many physicians still have difficulties drawing diagnostic inferences from statistics."[335]
- That it is very "difficult for even highly educated people to make inferences on the basis of probabilities."[336]
- That "[e]xperienced researchers are also prone to the same biases [as laymen] when they think intuitively."[337]
- That "expert auditors do not behave differently from novice auditors."[338]
- That experts often are no better than laypeople at making predictions.[339]
- That judges, although less susceptible than jurors, are subject to "a strong hindsight bias effect."[340]

---

332. PIATTELLI-PALMARINI, *supra* note 65, at 66 ("What is really surprising is that there is no great difference in the average responses from the 'uninformed' subject (that is, one who has no real notion of the laws of probability) and those of statistical experts. There is in fact a slight difference between the two groups: those who know something about statistics make more errors than the uninformed and also more than the experts. Even the experts, in fact, err more than the uninformed."); *see also* Tversky & Kahneman, *supra* note 93, at 26 (noting that the conjunction fallacy is committed "not only by statistically naïve undergraduates, but even by highly sophisticated respondents").

333. PIATTELLI-PALMARINI, *supra* note 65, at 119 (noting that in some studies of overconfidence, "the discrepancy between correctness of response and overconfidence *increases* as the respondent is *more knowledgeable*").

334. Barbara J. McNeil et al., *On the Elicitation of Preferences for Alternative Therapies,* 306 NEW ENG. J. MED. 1259, 1262 (1982) (finding that a significant majority of clinical doctors were subject to the framing effect). For example, if told that there was a mortality rate of 7% within 5 years of an operation, doctors hesitated to recommend it. But if told there was a 93% survival rate within 5 years of the operation, they were more inclined to recommend it. *See id; see also* Kuhberger, *supra* note 136, at 42 (noting that experts are subject to framing effects, although perhaps not to as significant a degree as students).

335. GIGERENZER, *supra* note 295, at 90.

336. *Id.* at 37.

337. Tverksy & Kahneman, *supra* note 116, at 18.

338. Jean Bédard, *Expertise in Auditing: Myth or Reality?,* 14 ACCT. ORGS. & SOC'Y 113, 121 (1989).

339. ROBIN M. HOGARTH, EDUCATING INTUITION 157 (2001) ("[E]xperts are not necessarily more accurate than novices when making certain types of predictions.").

340. Hastie & Viscusi, *supra* note 267, at 988; *see also* Hastie & Viscusi, *supra* note 66, at 901; W. Kip Viscusi, *How Do Judges Think About Risk?,* 1 AM. L. & ECON. REV. 26, 36-58 (1999) (reporting evidence of a further study demonstrating cognitive errors by experienced jurists).

Similar to psychology experiments, studies in experimental economics tend to find little difference between students and market professionals in laboratory experiments. *See* Sheryl B.

- That experts are more biased than people with intermediate levels of experience in predicting how long it will take novices to learn the basics of their skill.[341]
- That overconfidence effects have been documented in physicians, clinical psychologists, lawyers, engineers, security analysts, and other experts.[342]
- That expert auto mechanics typically consider only a small subset of the possible faults that can occur in a car and hence underestimate the probability of a breakdown.[343]
- That professional financial analysts are consistently too optimistic in estimating earnings,[344] consistently overreact to new information,[345] and tend not to learn from experience.[346]
- That professional blackjack players tend to be subject to the omission bias.[347]
- That basketball coaches, players, and experienced play-by-play commentators believe in the "hot hand" even though the phenomenon has no statistical validity.[348]

Ball & Paula-Ann Cech, *Subject Pool Choice and Treatment Effects in Economic Laboratory Research,* 6 RES. EXPERIMENTAL ECON. 239, 257 (1996) (finding after survey of many studies that only one produced much evidence of differences between students and market professionals).

341. Pamela J. Hinds, *The Curse of Expertise: The Effects of Expertise and Debiasing Methods on Prediction of Novice Performance,* 5 J. EXPERIMENTAL PSYCHOL.: APPLIED 205, 217 (1999).

342. Griffin & Tversky, *supra* note 179, at 230; *see also* GEOFFREY FRIESEN & PAUL A. WELLER, QUANTIFYING COGNITIVE BIASES IN ANALYST EARNINGS FORECASTS 30 (Univ. of Iowa, Oct. 2002) (finding that both overconfidence and cognitive dissonance affected financial analyst decisions and that the impact of overconfidence was "substantial"), http://papers.ssrn.com/sol3/papers.cfm?abstract_id=364700.

343. Baruch Fischhoff et al., *Fault Trees: Sensitivity of Estimated Failure Probabilities to Problem Representation,* 4 J. EXPERIMENTAL PSYCHOL.: HUM. PERCEPTION & PERFORMANCE 330, 342-43 (1978).

344. Thomas G. Calderon, *Predictive Properties of Analysts' Forecasts of Corporate Earnings,* 29 MID-ATLANTIC J. BUS. 41, 56 (1993) (noting that his study's "finding that analysts' forecasts reflect an upward bias is consistent with several prior studies in the area").

345. *See* De Bondt & Thaler, *supra* note 211, at 57 ("The same pattern of overreaction found in the predictions of naïve undergraduates is replicated in the predictions of stock market professionals.").

346. John Jacob et al., *Expertise in Forecasting Performance of Security Analysts,* 28 J. ACCT. & ECON. 51, 80 (1999) (noting also that their findings are consistent with Kanheman and Tversky's heuristics and biases literature).

347. Gideon Keren & Willem A. Wagenaar, *On the Psychology of Playing Blackjack: Normative and Descriptive Considerations with Implications for Decision Theory,* 114 J. EXPERIMENTAL PSYCHOL.: GEN. 133, 142 (1985) (suggesting regret as a potential explanation for why blackjack players "stand" with 16 or less even though the odds are better if they take another card because studies show people regret bad results that stem from action more than bad results that stem from inaction).

- That even with extensive experience, people have difficulty overcoming the winner's curse.[349]

Clearly, even experts with worlds of education often find themselves subject to the same biases as the average person, Mitchell's protests notwithstanding. That said, experts do sometimes make decisions differently than lay people,[350] and legal decision theorists are well aware of this fact.

i. Deductive Reasoning

Of all the heuristics and biases that have been studied, Mitchell concentrates on only a few, beginning with deductive reasoning.[351] Mitchell cites my statement regarding the confirmation bias[352] that people "tend to preferentially solicit evidence that confirms their hypothesis. Disconfirming evidence gets the short end of the perceptual stick."[353] Mitchell implies that my conclusion is wrong,[354] but there is substantial evidence, cited in my article, indicating that even scientists, trained auditors, and statisticians are subject to the confirmation bias. Scientists reviewing reports find those reports that agree with the scientists' preexisting opinions on the subject matter to be more persuasive than reports taking positions with which the scientists disagree.[355] Experienced auditors are more sensitive to

---

348. Gilovich et al., *supra* note 95, at 616; Koehler & Conley, *supra* note 95.

349. *See* Yoella Bereby-Meyer & Brit Grosskopf, *Overcoming the Winner's Course: An Adaptive Learning Experience* 24-27 (AOM Conflict Mgmt. Div. 2002 Mtgs. No. 13496, Aug. 2002) (finding that it is very difficult to debias the winner's curse which is, of course, the tendency to overpay in an auction scenario), http://papers.ssrn.com/sol3/papers.cfm?abstract_id =324201.

350. Markman & Medin, *supra* note 195, at 451-53 (citing studies).

351. Mitchell, *Incompetence, supra* note 12, at 87-90.

352. Generally speaking, the confirmation bias is the tendency to seek out and process information that confirms rather than disconfirms our pre-existing opinions. D. Micbael Risinger et al., *The Daubert/Kumho Implications of Observer Effects in Forensic Science: Hidden Problems of Expectation and Suggestion,* 90 CAL. L. REV. 1, 7 (2002). For example, if exposed to two contradictory studies regarding the death penalty, death penalty opponents will tend to find the one that undermines use of the death penalty to be the more convincing and probative, whereas death penalty supporters will tend to find the one that supports use of the death penalty to be more convincing and probative. Charles G. Lord et al., *Biased Assimilation and Attitude Polarization: The Effects of Prior Theories on Subsequently Considered Evidence,* 37 J. PERSONALITY & SOC. PSYCHOL. 2098, 2108 (1979).

353. Prentice, *supra* note 36, at 145-46, *quoted in* Mitchell, *Incompetence, supra* note 12, at 88 n.50.

354. Mitchell, *Incompetence, supra* note 12, at 88 (suggesting that such a conclusion might arise if a behavioralist read only a few studies)

355. *See* Jonathan J. Koehler, *The Influence of Prior Beliefs on Scientific Judgments of Evidence Quality,* 56 ORGANIZATIONAL BEHAV. & HUM. DECISION PROCESSES 28, 47 (1993) (finding that scientists judge research reports that agree with their previously held views to be of

information that confirms their initial hypotheses than to information that disconfirms those hypotheses.[356] Bazerman surveyed the studies in the area and concluded that "the tendency to exclude disconfirming information in the search process is not eliminated by the formal scientific training that is expected of statisticians."[357]

Mitchell focuses primarily on just one area of the confirmation bias—experiments involving the Wason card test.[358] Most subjects err by turning over cards that potentially confirm the rule, rather than cards that would definitively disconfirm it. Mitchell cites a study finding that 48% of subjects with a Master's degree or higher educational attainment correctly answered the problem.[359] Four observations are in order. First, more than half of the highly educated subjects in this study erred. Second, other studies have found that highly educated people did much worse than this on the Wason card test.[360] Third, the Wason card studies are but a small part of the evidence supporting the confirmation bias and Mitchell does not address all of the other studies that show such a bias in many settings. If Mitchell has any evidence indicating that scientists, judges, and other highly educated persons do not tend toward gathering information that supports (rather than rejects) their

---

higher quality than reports that disagree, and citing several other studies finding similar results).

356. *See* E. Michael Bamber et al., *An Examination of the Descriptive Validity of the Belief-Adjustment Model and Alternative Attitudes to Evidence in Auditing*, 22 ACCT. ORGS. & SOC'Y 249, 263 (1997) ("[I]n the evaluation of evidence auditors are confirmation prone in that they are more sensitive to evidence that confirms the initial hypothesis. This tendency holds over experience levels (staff and advanced senior), and over contexts that do and do not involve irregularities.").

357. *See* BAZERMAN, *supra* note 58, at 35-36 (citing several studies finding a confirmation bias among lay people and even statisticians).

358. Stephen Hoch and Judith Tschirgi describe this abstract logic test in the following manner:

> 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 antecedents (p and –p, respectively), and the cards 18 and 5 represent the consequents (q and –q). The solution is to turn over A and 5 (p and –q), because the rule is violated only by cards pairing a vowel with an odd number [i.e., A,5 (p,-q) and 5,A (-q, p)]. Most studies have found that only about 10% of the subjects can solve abstract forms of the problem.

Hoch & Tschirgi, *supra* note 244, at 453 (citing P.C. Wason, *Reasoning*, *in* NEW HORIZONS IN PSYCHOLOGY (B. Foss ed., 1966)).

359. *Id.*, *cited in* Mitchell, *Incompetence*, *supra* note 12, at 89 n.53.

360. *See, e.g.*, Leda Cosmides & John Tooby, *Beyond Intuition and Instinct Blindness: Toward an Evolutionarily Rigorous Cognitive Science*, 50 COGNITION 41, 62-63 (1994) (concluding that "even formal training in logical reasoning does little to boost performance" on the Wason test); Einhorn & Hogarth, *supra* note 205, at 399-400 (finding that a majority of statisticians "failed to make the appropriate response" in the Wason test).

opinions and processing that information in a confirming rather than disconfirming way, he does not cite it.[361] Fourth, the fact that people with Masters' degrees and Doctorates err in their approach to the Wason card problem "only" slightly more than half the time (as opposed to 90% for some other groups)[362] does not mean that legal policy and analysis cannot benefit substantially from information about the confirmation bias. Many legal rules are aimed at protecting more vulnerable groups of society. The unconscionability doctrine, for example, will not be invoked to aid educated, sophisticated consumers.[363] That does not mean that we do not need such a doctrine or that behavioral decision experiments cannot inform us as to its proper boundaries and applications.[364]

ii. Inductive Reasoning

Next, Mitchell addresses inductive reasoning, indicting legal decision scholars for their claim that people are not good "intuitive statisticians."[365] He suggests that some people are good statistical reasoners and that others can benefit from training.[366] There is, however, very strong evidence that while Chicago Man is skilled at probabilistic reasoning, most flesh-and-blood men and women resemble K-T Man.[367] Because of the availability heuristic, people tend to believe that available and salient risks are much more significant than statistics would indicate.[368] When strong emotions

361. The Wason test is one setting in which we can observe whether decision makers tend to use confirming or disconfirming strategies of hypothesis testing. But it is only one setting, and it is irrelevant to another important aspect of the confirmation bias—the tendency to interpret new evidence as confirming rather than disconfirming previously held opinions. *See, e.g.*, John M. Darley & Paget H. Gross, *A Hypothesis-Confirming Bias in Labeling Effects*, 44 J. PERSONALITY & SOC. PSYCHOL. 20, 28 (1983) (finding that when subjects witness a child taking an academic test, "[t]hose who believed the child came from a high socioeconomic class reported that her performance indicated a high ability level, whereas those who believed the child came from a low socioeconomic class reported that the identical performance indicated a substantially lower level of ability"); Koehler, *supra* note 355, at 28 (scientists reviewing reports); Lord et al., *supra* note 352, at 2098 (subjects studying arguments relating to the death penalty).

362. *See supra* note 358.

363. *See, e.g.*, Riesett v. W.B. Doner & Co., 293 F.3d 164, 173 (4th Cir. 2002) (holding that the unconscionability doctrine is irrelevant for contract between sophisticated parties); Stenke v. Masland Dev. Co., 394 N.W.2d 418, 424 (Mich. Ct. App. 1986) (same).

364. *See generally* Marrow, *supra* note 288 (using behavioral concepts to analyze unconscionability issues).

365. Mitchell, *Incompetence*, *supra* note 12, at 90-91.

366. *Id.*

367. GIGERENZER, *supra* note 295, at 37 (noting that it is "difficult for even highly educated people to make inferences on the basis of probabilities").

368. *See* Richard H. Pildes & Cass R. Sunstein, *Reinventing the Regulatory State*, 62 U. CHI. L. REV. 1, 60-62 (1995) (explaining how the availability heuristic can induce lay people to

(including hope and fear) are involved in decisions, people often attempt little assessment of probability at all.[369] Chicago Man would attend to base rates in most decision making, but the evidence as to whether most real people do so is very equivocal.[370] Rational decision makers would incorporate new information consistent with Bayes' Theorem, but again the evidence is very strong that most decision makers do not do so.[371] Instead, they tend to overreact to small risks,[372] although at the very far extremes of the probability scale they tend to ignore probabilities completely.[373] They systematically overestimate the likelihood of conjunctive events and underestimate the likelihood of disjunctive events.[374] They often believe, irrationally,

---

overestimate certain types of risks and urging policymakers to educate the public). *See generally* BAZERMAN, *supra* note 58, at 7 (noting that when Peter Lynch ran Fidelity's Mutual Fund he bought securities of bland firms because the more well-known a firm was, the more its stock was likely to be overvalued); Donald V. Moser, *The Effects of Output Interference, Availability, and Accounting Information on Investors' Predictive Judgments,* 64 ACCT. REV. 433, 444 (1989) (finding availability bias in investors' predictive earnings judgments); Amos Tversky & Daniel Kahneman, *Availability: A Heuristic for Judging Frequency and Probability, in* JUDGMENT UNDER UNCERTAINTY, *supra* note 8, at 164, 178 (concluding that "production of a compelling scenario is likely to constrain future thinking").

369. *See* Cass R. Sunstein, *Beyond the Precautionary Principle,* 151 U. PA. L. REV. 1003, 1044 (2003) ("[W]hen strong emotions are involved . . . large-scale variations in probabilities will matter little—even when those variations unquestionably should matter."). Sunstein relies on the studies of Rottenstreich & Hsee. *See* Yuval Rottenstreich & Christopher Hsee, *Money, Kisses, and Electric Shocks: On the Affective Psychology of Risk,* 12 PSYCHOL. SCI. 185 (2001), *cited in* Sunstein, *supra,* at 1044.

370. Koehler makes a strong argument that the "base rate fallacy" is overstated. Koehler, *supra* note 97, at 2 ("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."); Jonathan J. Koehler, *The Base Rate Fallacy Myth,* 49 PSYCOLOQUY 1, ¶ 1.2 (1993), http://psycprints.ecs.soton.ac.uk/archive/00000343 (similar). However, many psychologists disagree. *See, e.g.,* BAZERMAN, *supra* note 58, at 19-20; ROBYN M. DAWES, EVERYDAY IRRATIONALITY: HOW PSEUDO-SCIENTISTS, LUNATICS, AND THE REST OF US SYSTEMATICALLY FAIL TO THINK RATIONALLY 91 (2001); REID HASTIE & ROBYN M. DAWES, RATIONAL CHOICE IN AN UNCERTAIN WORLD 111-13 (2001); PLOUS, *supra* note 135, at 115-16.

371. *See generally* HASTIE & DAWES, *supra* note 370, at 111-13 (summarizing numerous studies, some done by Kahneman & Tversky); Ward Edwards, *Conservatism in Human Information Processing, in* JUDGMENT UNDER UNCERTAINTY, *supra* note 8, at 359, 359 ("It turns out that opinion change is very orderly, and usually proportional to numbers calculated from Bayes's theorem—but it is insufficient in amount."); Joseph F. Funaro, *An Empirical Analysis of Five Descriptive Models for Cascaded Inference,* 14 ORGANIZATIONAL BEHAV. & HUM. PERFORMANCE 186, 186 (1975) (noting that most studies find that "intuitive opinion revisions are conservative in comparison to the optimal revisions specified by Bayes' theorem").

372. *See* Stephen D. Sugarman, *A Restatement of Torts,* 44 STAN. L. REV. 1163, 1173 (1992) (noting that empirical evidence indicates that people tend to "overreact to small risks of which they are somewhat aware").

373. PIATTELLI-PALMARINI, *supra* note 65, at 130.

374. HASTIE & DAWES, *supra* note 370, at 137 (citing studies). Conjunctive events are the combination of two simple events, say throwing a 1 and a 6 with two dice. *Id.* at 169. Disjunctive events are more inclusive, such as throwing a 1 or a 6, *or both. Id.*

in the gambler's fallacy[375] and the "hot hand."[376] They misunderstand the concept of regression toward the mean.[377] They suffer from denominator blindness, in that they are likely to view four accidents out of 100 products sold as quite similar to four accidents out of 1,000 products sold.[378]

Critics of the heuristics and biases literature hang much of their criticism on studies indicating that when problems framed as probabilities are reframed as frequencies, people often do much better.[379] However, as Mitchell would admit, there is also substantial evidence that reframing questions as frequencies often does not improve problem solving.[380] Furthermore, reformatting questions from probabilities to frequencies does not make the limitation on human decision making go away, nor does it help much when it is clear that as people go through life they will often be faced with

---

375. The gambler's fallacy occurs when someone sees a flipped coin land tails up four times in a row. PLOUS, *supra* note 135, at 113. Many people believe that on the next flip it is more likely than not that the coin will land heads up. *Id.*

376. The "hot hand" error is the belief that a basketball player who has hit a few shots in a row now has the "hot hand" and is likely to continue to shoot at an average above his usual rate. Even professional coaches and players believe in the phenomenon, although it has been thoroughly debunked. *See* Gilovich et al., *supra* note 95, at 601-13; Koehler & Conley, *supra* note 95.

377. HOGARTH, *supra* note 339, at 124; Daniel Kahneman & Amos Tversky, *Conflict Resolution: A Cognitive Perspective, in* BARRIERS TO CONFLICT RESOLUTION 44, 46 (Kenneth J. Arrow et al. eds., 1995). For example, many people believe fervently in the "*Sports Illustrted* jinx," the notion that a player or team that appears on the cover of this sports magazine is likely to suffer bad luck. But the players or team are often chosen for the cover because of some extraordinary performance, so when they revert to their normal performance it seems like they have been jinxed. *See* JOHN ALLEN PAULOS, A MATHEMETICIAN PLAYS THE STOCK MARKET 106 (2003) (explaining the so-called jinx)

378. W. Kip Viscusi & Richard J. Zeckhauser, *The Denominator Blindness Effect: Accident Frequencies and the Misjudgment of Recklessness* 22-23 (Harvard Law & Econ. Discussion Paper No. 387, Oct. 2002), http://papers.ssrn.com/sol3/papers.cfm?abstract_id=357160.

379. The most interesting view is based in evolutionary psychology, suggesting that humans are not well-suited to using probability theory, which was invented only a few hundred years ago. However, they are arguably better at using frequencies, because this they have been doing while making decisions for tens of thousands of years. *See* Gary L. Brase et al., *Individuation, Counting, and Statistical Inference: The Role of Frequency and Whole-Object Representations in Judgment Under Uncertainty,* 127 J. EXPERIMENTAL PSYCHOL.: GEN. 3, 4 (1998).

380. *See* Jonathan St. B.T. Evans et al., *Frequency Versus Probability Formats in Statistical Word Problems,* 77 COGNITION 197, 212 (2000) (finding that use of frequencies improves probabilistic reasoning only if framed in a very specific way that makes the problem easier to understand, and concluding that "our findings cast serious doubts upon the widely cited claim that frequency formats facilitate correct statistical reasoning in quantitative word problems"); 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) (finding that use of frequency formats for ratings by psychologists and psychiatrists for likelihood of future violence and assessed risk of violence does not necessarily improve decision making by minimizing format effects of response scales).

problems framed as probabilities.[381] The evidence regarding the difficulties people face in assessing probabilities does point us toward a useful policy prescription arising from behavioral research—that required disclosures should be framed as frequencies rather than probabilities wherever possible or, better yet, presented in both formats.

Mitchell really does not dispute that people are generally inadequate statistical reasoners. Rather, his main point is that, contrary to some studies, there is definitely evidence that a person trained in statistics has a greater likelihood of using statistical reasoning and a better quality of reasoning.[382] Nothing could better illustrate the value of legal decision theory. Whereas the rational man model simply assumes that everyone effectively reasons inductively, behavioral evidence indicates that people do not, but that they can

---

381. As Stanovich notes:

> I am living in a technological society where I must: decide which HMO to join based on just such statistics, figure out whether to invest in a Roth IRA, decide what type of mortgage to purchase, figure out what type of deductible to get on my auto insurance, decide whether to trade in a car or sell it myself, decide whether to lease or to buy, think about how to apportion my TIAA/CREF retirement funds, and decide whether I would save money by joining a book club. And I must make all of these decisions based on information represented in a manner for which my brain is not adapted (in none of these cases have I coded individual frequency information from my own person [sic] experience). In order to reason normatively in all of these domains (in order to maximize my own personal utility), I am going to have to deal with probabilistic information represented in nonfrequentist terms—in representations that the cognitive ecologists have shown are different from my well-adapted algorithms for dealing with frequency information.
>
> . . . The problem is that in a symbol-oriented postindustrial society, we are presented with paper-and-pencil problems all the time, and much of what we know about the world comes not from the perception of actual events but from abstract information preprocessed, prepackaged, and condensed into symbolic codes such as probabilities, percentages, tables, and graphs (the voluminous statistical information routinely presented in *USA Today* comes to mind).

STANOVICH, *supra* note 285, at 206-07.

Koehler and colleagues make the same point, suggesting that we must

> accept that subjective probabilities are not only natural, but inescapable. A historical review of the use of concepts related to "chance" in more than 500 years of English literature noted that "[w]ith one exception, all quotations found . . . are subjective probabilities. They all are expressions of degrees of belief, at least in a poetic sense, that an event will happen." Consider the following statement, taken from a recent financial column: "Three months ago, I might have put the risk of an Asia-wide credit crunch . . . at less than 10%, now I'd say it is approaching 30% to 40%." How can this statement about a unique, not to say unprecedented, event, be reworded in frequency terms?

Koehler et al., *supra* note 186, at 715 (citation omitted).

382. Mitchell, *Incompetence, supra* note 12, at 37 (citing Richard E. Nisbett et al., *The Use of Statistical Heuristics in Everyday Inductive Reasoning,* 90 PSYCHOL. REV. 339, 358 (1983)). Mitchell is careful not to overstate his claim in this area, for Nisbett, Mitchell's primary source, admits that "Kahneman and Tversky have shown repeatedly that statistical expertise provides no such guarantee against errors [in inductive reasoning]." Nisbett et al., *supra*, at 359 (citation omitted).

improve with training. This suggests that judges should be given at least minimal statistical training to improve their statistical reasoning. It might further suggest that in certain cases, jurors should be given such training as well, because Mitchell cites sources indicating that substantial improvements can be gained with minimal training when that training is based on psychological principles rather than on the traditional statistical approach.[383]

Mitchell's references to behavioral studies showing that training in economics can improve cost-benefit reasoning[384] and help subjects avoid sunk-cost effects[385] also point to policy prescriptions, such as requiring courses in economics in college and perhaps high school. Mitchell is hoist on his own petard, for the very behavioral studies that Mitchell himself references provide important insights into legal policy making that rational man economic analysis does not. If one can set aside Mitchell's inaccurate claim that legal decision theorists claim or believe that people are universally, uniformly, and apparently irreparably irrational, one can gain extremely useful insights from his excellent discussion of the nuances of the behavioral literature.

b. Cognitive Capacity

Mitchell next mines the controversial (but, in my view, substantially accurate) research of Keith Stanovich and Richard West,

---

383. *See* Mitchell, *Incompetence, supra* note 12, at 92 n. 65 (citing Peter Sedlmeier & Gerd Gigerenzer, *Teaching Bayesian Reasoning in Less Than Two Hours,* 130 J. EXPERIMENTAL PSYCHOL.: GEN. 380, 396 (2001)). Sedlmeier and Gigerenzer argued that a psychological approach to teaching Bayesian reasoning can greatly improve upon the traditional statistical approach to teaching the subject. Sedlmeier & Gigerenzer, *supra,* at 396.

384. No behavioral decision theorist that I know of would claim that people with professional training in economics could not reason more consistently with economic principles than people with no such training. Apparently, people with such training do, as expected, reason more consistently. *See* Richard Larrick et al., *Teaching the Use of Cost-Benefit Reasoning in Everyday Life,* 1 PSYCHOL. SCI. 362, 365-69 (1990).

385. In making decisions, people tend (irrationally, according to economists) to value sunk costs. *See* Hal R. Arkes & Catherine Blumer, *The Psychology of Sunk Cost,* 35 ORGANIZATIONAL BEHAV. & HUM. DECISION PROCESSES 124, 124 (1985). The sunk cost phenomenon appears to affect how NBA teams handle their draft picks. *See* Barry M. Staw & Ha Hoang, *Sunk Costs in the NBA: Why Draft Order Affects Playing Time and Survival in Professional Basketball,* 40 ADMIN. SCI. Q. 474, 487 (1995) ("Regressions showed that the higher a player was taken in the college draft, the more time he was given on the court, even after controlling for such other logical predictors of playing time, such as performance, injury, and trade status."). And it has been suggested as an influence upon United States decision making during the Viet Nam war. *See* MICHAEL HOWARD, THE CAUSES OF WAR 232 (1984) (quoting Secretary of Defense Robert McNamara as writing "[w]e could not simply walk away from an enterprise involving two administrations, five allied countries, and thirty-one thousand dead as if we were switching off a television channel").

who have provided evidence that persons with higher cognitive capacity are significantly more likely to give normative responses in most of the heuristics and biases tests commonly used by behavioral psychologists.[386] Generally speaking, Stanovich and West find that subjects with greater cognitive capacity do better on statistical reasoning tests, the Wason selection task, tests for framing effects, tests for the conjunction fallacy, and others.[387] They tend not to do better on tests involving the false consensus effect, the overconfidence effect, and the influence of noncausal base rates.[388]

The results Stanovich and West find are very interesting and certainly support the uncontroversial point that not all people are equally subject to all heuristics and biases. These findings do not undermine traditional notions of K-T Man, however. Note first that although subjects with greater cognitive capacity often do better on these tests than subjects who are not as "smart," these "smarter" subjects nonetheless often err. They are subject to many of the effects and biases and use many of the heuristics identified by Kahneman and Tversky, although they are not as affected as people with less cognitive capacity. Stanovich concludes that cognitive capacity differences can account for systematic discrepancies between actual performance and normative models to only a "moderate extent."[389] After performing some of the most important experiments in the area and reporting on them in a recent book, Stanovich remains strongly in the K-T camp, as the ultimate sentence in his book notes that "[i]t does seem that some human behavior is systematically irrational."[390]

Second, the findings of Stanovich and West strongly bolster the mainstream behavioral psychologists against the attack of some who have claimed that the results Kahneman, Tversky, and their collaborators have labeled as their normative benchmark are not actually normative. These critics argue that it is the subjects who are getting it right and the psychologists who are getting it wrong.[391]

---

386. *See* Mitchell, *Incompetence, supra* note 12, at 94-95 (citing numerous works by Stanovich and West).

387. STANOVICH, *supra* note 285, at 39-40.

388. *Id.* Stanovich and West found similar differences between groups of people with different thinking dispositions. *Id.* at 153-89.

389. *Id.* at 210.

390. *Id.* at 252. More recently, Stanovich and West concluded that their findings "support the notion that the normative/descriptive discrepancies that remain after computational limitations have been accounted for reflect a systematically suboptimal intentional-level psychology." Keith E. Stanovich & Richard F. West, *Individual Differences in Reasoning: Implications for the Rationality Debate?, in* PSYCHOLOGY OF INTUITIVE JUDGMENT, *supra* note 8, at 421, 440.

391. Mitchell raises this question himself. *See* Mitchell, *Pessimism, supra* note 12, at 1941-43.

However, if smarter people side with the psychologists, the critics are probably in error.[392]

Third, this information that individuals vary in their capacity to make decisions rationally is very useful information. "[C]onventional economic theories as exposited do not discriminate among types of individuals,"[393] but psychologists and legal decision theorists do. The law already often takes into account the fact that children or the elderly or the mentally infirm may be uniquely vulnerable.[394] This additional information may allow us to sharpen the protective focus of the law. More obviously, and as noted earlier, not all consumers need be equally subject to framing effects or other reasoning limitations for legal decision makers to take into account how marketers can use those effects to market tobacco and other products.[395] All consumers need not be equally vulnerable to sharp practices for the unconscionability doctrine to be enforced to protect those who are. Not all citizens need be subject to the status quo bias for a legislature to make a big change in consumer choices by altering the status quo. Not all jurors need be equally vulnerable to the hindsight bias for it to be sensible to give an instruction to attempt to mitigate its effects.

As with the previous section, Mitchell should be congratulated for advancing the debate about the role of behavioral psychology research in the legal realm. His articles, contrary to his warnings, make this body of research appear more, not less, valuable, as it will

---

392. STANOVICH, *supra* note 285, at 66 ("Yet despite the normative controversies surrounding all of these tasks, each one displayed the same type of association with cognitive ability—more intelligent individuals were more likely to give the response traditionally considered normative.").

More recently, Stanovich and West noted that "we should resist the conclusion that individuals with more computational power are systematically computing the *non*normative response. Such an outcome would be an absolute first in a psychometric field that is 100 years and thousands of studies old." Stanovich & West, *supra* note 390, at 427 (citations omitted).

393. David M. Grether, *Bayes Rule as a Descriptive Model: The Representativeness Heuristic*, 95 Q.J. ECON. 537, 555 (1980).

394. Obviously legislatures pass special rules to protect children from exploitative advertising and pornography. *See* Child Online Protection Act (COPA), 47 U.S.C. § 231 (2000); Donald W. Garner & Richard J. Whitney, *Protecting Children from Joe Camel and His Friends: A New First Amendment and Federal Preemption Analysis of Tobacco Billboard Regulation*, 46 EMORY L.J. 479, 487-89 (1997) (describing a raft of state and local laws aimed at protecting children from tobacco advertising). Courts use doctrines such as unconscionability and undue influence to protect the elderly in contract cases. *See, e.g.*, Bennett v. Bailey, 597 S.W.2d 532, 535 (Tex. App. 1980) (refusing to enforce an unconscionable contract against an elderly widow). Rules requiring contractual capacity protect the mentally infirm.

395. Jon D. Hanson & Douglas A. Kysar, *Taking Behavioralism Seriously: The Problem of Market Manipulation*, 74 N.Y.U. L. REV. 630, 724 (1999) (arguing that manufacturers will be driven by market forces to exploit consumer weaknesses).

help identify the specific cases where people are particularly subject to biases.

2. Differences Within Individuals

Mitchell's next point is that "[i]ndividuals vary over time in their propensity to engage in rational behavior."[396] He cites substantial psychological literature to this effect; legal decision theorists are thus also aware of this fact and take it into account in their policy prescriptions.

As Mitchell recognizes, "the most-studied factor with regard to intra-individual changes in rational behavior is affective state."[397] The Chicago Man model has no explanation for the effect that emotion has on decision making,[398] but psychologists are well aware of it,[399] and legal scholars interested in behavioral theory have often taken it into account. Sunstein, for example, recently cited the work of Rottenstreich and Hsee, who found in a series of very interesting studies that when strong emotions are involved, people often attempt little assessment of probability at all.[400] The strong fears they have of a particularly salient risk may lead them to spend large sums of money to try to avoid it, even though the probability of it occurring is low.[401] The strong hopes they have of winning the lottery may cause them to spend substantial sums of money on lottery tickets, even though the chances of winning the lottery are vanishingly thin. Sunstein explained the implications of this affect heuristic for environmental regulation, which include the fact that we tend to overregulate the risks that people respond to emotionally and to

---

396. Mitchell, *Incompetence, supra* note 12, at 98.

397. *Id.* at 99.

398. *But see* Richard H. McAdams, *Accounting for Norms,* 1997 WIS. L. REV. 625, 629-30 (noting that some economists are starting to modify the traditional rational man model, infusing it with such complexities as emotion, altruism, habit, and risk misperception). However, Blumenthal notes that when scholars begin incorporating the impact of emotions on decision making into their models they tend to "fundamentally undercut their own arguments and proposals." Blumenthal, *supra* note 1, at 25.

399. *See* Slovic et al., *supra* note 94, at 398 ("[T]he importance of affect is being recognized increasingly by decision researchers."); *see also* ANTONIO R. DAMASIO, DESCARTES' ERROR: EMOTION, REASON, AND THE HUMAN BRAIN 3-19 (1994) (discussing a landmark study of how physical brain damage can leave logical reasoning intact but nonetheless impair effective decision making by injuring centers of the brain controlling emotion).

400. *See* Sunstein, *supra* note 369, at 1010, (citing Rottenstreich & Hsee, *supra* note 369, at 185).

401. *See* Slovic et al., *supra* note 94, at 409 (noting that this affect heuristic helps explain "why societal concerns about hazards such as nuclear power and exposure to extremely small amounts of toxic chemicals fail to recede in response to information about the very small probabilities of the feared consequences from such hazards").

underregulate greater risks that do not prompt such a strong affective response.[402]

Studies indicate that when cigarette advertising (by use of the word "natural" or by featuring scenes of nature) increases the positive affect associated with smoking, it also reduces potential smokers' perception of the risk of cigarette smoking.[403] Hanson and Kysar have explained how marketers, especially tobacco marketers, have manipulated the affect heuristic (and others) in a way that may justify more legal liability and regulation.[404]

Jonathan Koehler and I recently noted how the strong affective response that arises in situations of counterfactual ("what if. . .") thinking may influence judgments.[405] Exploring extensive literature in counterfactual thinking[406] and anticipated regret theory,[407] we noted that studies have shown that mock jurors will tend to impose stiffer punishments against muggers who attack a person on her way home and to return higher damage awards against careless drivers who hit a person on her way home if the victim is taking an *unusual* route home.[408] In an unusual setting, it is easy for jurors to imagine "if only she had been taking her usual route home, this wouldn't have happened." This counterfactual thinking increases the jurors' emotional response to the facts of the case and, in turn, affects their judgments. This is evidence of a *normality bias*. Other studies show

---

402. Sunstein, *supra* note 369, at 1046 ("If probabilities are neglected, especially when emotions are engaged, then the principle will operate through excessive public concern with certain low-probability hazards."). One of Kahneman and Tversky's most persistent critics, Gerd Gigerenzer, agrees with Sunstein that people generally misperceive risks. GIGERENZER, *supra* note 295, at 238 ("John Q. Public does not always fear the situations that are actually most likely to hurt or kill him and other people.").

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

404. Jon D. Hanson & Douglas A. Kysar, *Taking Behavioralism Seriously: Some Evidence of Market Manipulation*, 112 HARV. L. REV. 1420, 1467-1565 (1999).

405. *See* Robert A. Prentice & Jonathan J. Koehler, *A Normality Bias in Legal Decision Making,* 88 CORNELL L. REV. 583, 616-21 (2003).

406. *See supra* note 99 and accompanying text.

407. According to Loomes & Sugden:

> The essential notion underlying regret theory is that people tend to compare their actual situations with the ones they would have been in, had they made different choices in the past. If they realize that a different choice would have led to a better outcome, people may experience the painful sensation of regret; if the alternative would have led to a worse outcome, they may experience a pleasurable sensation we call "rejoicing." When faced with new choice situations, people remember their previous experiences and form expectations about the rejoicing and regret that the present alternatives might entail. They then take these expectations into account when making their decisions.

Graham Loomes & Robert Sugden, *A Rationale for Preference Reversal,* 73 AM. ECON. REV. 428, 428 (1983).

408. *See* Prentice & Koehler, *supra* note 405, at 616-21.

that when intent, injury, and all other factors are held constant, jurors will punish more severely defendants whose acts were active rather than passive (for example, the physician who unplugged the life support machine as opposed to the physician who failed to plug it in when she had the chance to).[409] This is evidence of an *omission bias*. Koehler and I experimentally tested the effect of these two biases in an attempt to determine which of the biases (which reinforce one another in most settings) would trump the other in a factual situation where they conflicted.[410] We noted the importance of affect to decision making and most definitely did not assume that all individuals are equally irrational in all situations, for we sought to determine in which settings which bias would predominate.

As Mitchell himself notes, other legal decision theorists such as Chris Guthrie[411] and Russell Korobkin[412] have also studied the implications of regret theory for legal decision making. Mitchell concedes that "[b]y bringing regret aversion into the decision calculus, we complicate the model of litigation behavior but also may gain descriptive accuracy, for the avoidance and achievement of certain affective states appears to be an important part of many decisions, particularly settlement decisions."[413] Descriptive accuracy is indispensable to effective policy making. Because psychologists have a greater interest in the impact of affect on human decision making than do economists, their descriptions of how people make decisions are likely to be more accurate than economic models.

Mitchell concludes the section by arguing that experimental simulations should be supplemented "with observational research, archival and case studies, and interviews and surveys of actual jurors."[414] I agree. Indeed, the behavioral literature is rich with varied approaches to studying similar problems, and a strength of the new legal decision theory research is that all of these various types of methodologies can be tapped to improve descriptive and explanatory accuracy.[415] Psychologists in general recognize the benefit of tapping

---

409. *See* Prentice & Koehler, *supra* note 405, at 587.

410. *Id.*

411. Chris Guthrie, *Better Settle Than Sorry: The Regret Aversion Theory of Litigation Behavior*, 1999 U. ILL. L. REV. 43.

412. Korobkin, *supra* note 254, at 1583.

413. Mitchell, *Incompetence*, *supra* note 12, at 104.

414. *Id.* at 104-05.

415. *See, e.g.*, HANDBOOK OF RESEARCH METHODS, *supra* note 73. This large tome begins with a significant section on research design and issues of inference validity. It then explores numerous types of research methodology, including behavioral observation and coding, small group research, event-sampling, survey research, content analysis, narrative analysis, and others. Finally, it discusses various data analysis strategies. This book and numerous others set forth the wide variety of research designs used by psychologists to develop the literature being

into many different forms of supporting evidence.[416] Contrary to Mitchell's implication, so do legal decision theorists.

Consider my article regarding the self-serving bias.[417] My essential point was that we should be very wary of allowing large accounting firms to offer all manner of nonaudit services, including legal services, to audit clients because the self-serving bias would create a strong danger that the firms would attempt to please their audit clients in order to preserve the stream of nonaudit revenue.[418] In making that point I used several forms of evidence to bolster my argument in just the manner that Mitchell suggests.

- After introducing the concept of the self-serving bias and noting the debate regarding its origins (cognitive vs. motivational), I reported a wide range of studies using various forms of empiricism that found self-serving behavior by attorneys, physicians, investment bankers, securities analysts, expert witnesses, scientists, and judges.[419] For example, many studies have documented that physicians order more tests and longer treatments when they refer patients to laboratory facilities that they own than when they refer patients to facilities owned by others.[420]

---

tapped by Guthrie, Korobkin, Langevoort, Rachlinski, Sunstein, and others. *See, e.g.*, RESEARCH METHODS IN PSYCHOLOGY (Glynis M. Breakwell et al. eds., 2d ed. 2000); ROBERT ROSENTHAL ET AL., CONTRASTS AND EFFECT SIZES IN BEHAVIORAL RESEARCH: A CORRELATIONAL APPROACH (2000); ROBERT ROSENTHAL & RALPH L. ROSNOW, ESSENTIALS OF BEHAVIORAL RESEARCH: METHODS AND DATA ANALYSIS (1984); ROBERT ROSENTHAL, JUDGMENT STUDIES: DESIGN, ANALYSIS, AND META-ANALYSIS (1987); HOI K. SUEN & DONALD ARY, ANALYZING QUANTITATIVE BEHAVIORAL OBSERVATION DATA (1989).

416. STANOVICH, *supra* note 64, at 119 ("[P]sychologists should always be concerned that their experimental conclusions not rely too heavily on any one method or particular subject population.").

417. Prentice, *supra* note 211, at 1597.

418. *Id.* at 1668-70.

419. *Id.* at 1620-29.

420. *See* Thomas L. Carson, *Conflicts of Interest*, 13 J. BUS. ETHICS 387, 394 (1994) ("[W]hen physicians are paid according to how much work they do for their patients, many physicians succumb to the temptation to provide their patients with unnecessary, even dangerous treatments."); Thomas S. Crane, *The Problem of Physician Self-Referral Under the Medicare and Medicaid Antikickback Statute*, 268 JAMA 85, 86 (1992) (citing studies by the Office of the Inspector General and the Florida Health Care Cost Containment Board showing how physicians respond to financial incentives in their treatment practices); David Hemenway et al., *Physicians' Responses to Financial Incentives: Evidence from a For-Profit Ambulatory Care Center*, 322 NEW ENG. J. MED. 1059, 1062 (1990) (finding that when compensation of physicians changed from a flat fee to a bonus system where they could earn extra revenue by ordering laboratory tests for patients on machines owned by employers, physicians increased their orders for lab tests by 23%); Bruce J. Hillman et al., *Frequency and Costs of Diagnostic Imaging in Office Practice—Comparison of Self-Referring and Radiologists-Referring Physicians*, 323 NEW ENG. J. MED. 1604, 1606 (1990) (finding that self-referring physicians employed diagnostic imaging at least

- Then I reported the results of several laboratory studies of accountants doing nonaudit work, such as tax and consulting, that revealed a self-serving bias. For example, some of the studies noted that tax professionals' judgments as to the appropriateness of a tax treatment varied greatly, depending on whether they learned that a particular position was favored by their client.[421]
- Turning to auditors specifically, I reported the results of a number of audit laboratory studies that reflected a self-serving bias.[422] Most of the studies involved practicing auditors and almost all found a self-serving bias. For example, one study presented auditors with an ambiguous situation and "found that the auditors acted in a manner consistent with the self-serving bias in that they used aggressive interpretations of accounting standards to allow their clients to take aggressive positions when the auditors' engagement risk (risk of fines, censure, litigation, loss of reputation, etc.) was moderate, but used conservative interpretations of the same standards in order to require conservative positions by their clients when the auditors' engagement risk was high."[423]
- Noting specifically that laboratory results cannot be assumed to always reproduce the real world, I examined several archival and case studies of actual auditor behavior,[424] including Wright and Wright's examination

---

four times as frequently than their colleagues who referred patients to others); Bruce J. Hillman et al., *Physicians' Utilization and Charges for Outpatient Diagnostic Imaging in a Medicare Population,* 268 JAMA 2050, 2050 (1992) ("[S]elf-referral resulted in 1.7 to 7.7 times more frequent performance of imaging examinations than radiologist-referral."); Jean M. Mitchell & Elton Scott, *Physician Ownership of Physical Therapy Services,* 268 JAMA 2055, 2055 (1992) (finding that visits per patient were 39% to 45% higher in facilities where referring physicians were joint venturers and that more revenues were generated from patients with well-paying insurance).

421. *See* C. Brian Cloyd & Brian C. Spilker, *The Influence of Client Preferences on Tax Professionals' Search for Judicial Precedents, Subsequent Judgments, and Recommendations,* 74 ACCT. REV. 299, 301 (1999) (finding that one-half of subject tax professionals recommended the client-preferred position even though a panel of experts concluded that there was only a 14% chance it would be sustained if challenged).

422. Prentice, *supra* note 211, at 1640-44.

423. *Id.* at 1642 (citing Karl Hackenbrack & Mark W. Nelson, *Auditors' Incentives and Their Application of Financial Accounting Standards,* 71 ACCT. REV. 43 (1996)).

424. *Id.* at 1644-49.

of 186 sample audits,[425] Krishnan and Krishnan's study using data from 1,837 public companies,[426] and Citron and Taffler's examination of ten years' worth of going concern qualifications in the United Kingdom.[427] These studies and most of several others that I cited, performed in a number of settings and in different countries, found a strong self-serving bias.[428] For example, despite the former Big Five firms' claims to Congress that litigation risk forced them to resign accounts, Scholz found that they did so *only* if the client was in relatively poor financial condition; if the client was bringing in a large stream of revenue and was not in danger of failing, litigation risk was not correlated with resignation.[429]

- I then examined several studies, some arising from laboratory experiments and some not, that studied the psychology of *individual* auditors in an attempt to explain why it might be in their self-interest to coddle an audit client even though it was not in the audit *firm's* rational economic interest to do so.[430]

In short, I used a variety of types of empirical studies conducted by a variety of scholars in a variety of settings, including both the laboratory and the real world, to undermine the simplistic notion promulgated by law-and-economics judges that courts act appropriately when they simply assume that auditors placed in an inherently conflicted situation will, because of the reputational

---

425. Arnold Wright & Sally Wright, *An Examination of Factors Affecting the Decision to Waive Audit Adjustments,* 12 J. ACCT. AUDITING & FIN. 15, 33 (1997) (finding a "strong positive association . . . between the likelihood [that audit firms waived an adjustment] and client size, a surrogate for audit fees").

426. Jagan Krishnan & Jayanthi Krishnan, *The Role of Economic Trade-Offs in the Audit Opinion Decision: An Empirical Analysis,* 11 J. ACCT. AUDITING & FIN. 565, 583 (1996) (finding that an audit firm is more likely to issue a qualified opinion "the higher the litigation risk . . . [and] the lower the client's decile position in the auditor's portfolio").

427. David B. Citron & Richard J. Taffler, *The Audit Report Under Going Concern Uncertainties: An Empirical Analysis,* 22 ACCT. & BUS. RES. 337, 344 (1992) (finding that auditors in the UK tended to issue going concern qualifications prior to a client's bankruptcy only if the company was *both* in a very weak financial condition and in imminent danger of failing).

428. I located and reported one study which found little evidence of the self-serving bias. *See* Timothy J. Louwers, *The Relation Between Going-Concern Opinions and the Auditor's Loss Function,* 36 J. ACCT. RES. 143 (1998).

429. Prentice, *supra* note 211, at 1647 (citing Susan Scholz, *Auditor Litigation Risk and Auditor Resignations: Some Empirical Evidence* (1997) (unpublished paper on file with author)).

430. *Id.* at 1650-53.

constraint, always act in the manner economists assume is rational.[431] Those of us who argued against allowing audit firms to offer legal services won the debate, not because of the persuasiveness of my article, but because Arthur Andersen acted in the manner predicted by the studies cited in my article and thereby brought about the Enron debacle,[432] which functionally ended any substantial encroachment by accounting firms into the field of legal representation in this country for the foreseeable future.[433]

Thus, Mitchell has overstated the differences in rationality across individuals and between individuals. More importantly, uniformity in this regard is not the basis for behavioralist policy prescriptions, nor need it be. Psychologists are well aware of these individual differences. Importantly, legal decision theorists have, even before Mitchell suggested it, used many different forms of evidence beyond experimental simulations to reach their conclusions.

B. Situational Differences

Mitchell's next major claim is that legal decision theorists believe in "universal characteristics of human cognition"[434] and therefore do not recognize that "*all situations do not elicit the same type of thinking.*"[435] This claim sets up yet another straw man, for much of the behavioral literature is composed of experiments designed to determine under what circumstances various biases appear and

---

431. This point of view had been promulgated by law-and-economics judges and adopted in many courts. *See, e.g.*, Melder v. Morris, 27 F.3d 1097, 1103 (5th Cir. 1994); DiLeo v. Ernst & Young, 901 F.2d 624, 629 (7th Cir. 1990); Retsky Family L.P. v. Price Waterhouse LLP, No. 97-C7694, 1998 U.S. Dist. LEXIS 17459, at *26 n.2 (N.D. Ill. Oct. 21, 1998); Stamatio v. Hurco Cos., 885 F. Supp. 1180, 1185 (S.D. Ind. 1995); *In re* Software Toolworks, Inc. Sec. Litig., 789 F. Supp. 1489, 1499 (N.D. Cal. 1992), *aff'd in part, rev'd in part*, 50 F.3d 615 (9th Cir. 1994).

432. In order to preserve a stream of mostly nonaudit revenue that it hoped would soon grow to $100 million annually, Arthur Andersen looked the other way when it learned of Enron's financial shenanigans. *See* Chris Ayres, *Enron Memo "Reveals $2bn Smoking Gun,"* TIMES (London), Jan. 18, 2002, at 1, 2002 WL 4174798. This is consistent with additional academic commentary published after my articles. *See, e.g.*, Don A. Moore et al., *Auditor Independence, Conflict of Interest, and the Unconscious Intrusion of Bias* 32 (Harvard NOM Working Paper No. 02-40, 2002) (finding that people's private opinions are easily biased in ways that are consistent with the interests of partisans with whom they are affiliated, and these people tend not to realize the strength of the bias or to be able to correct for it), http://papers2.ssrn.com/paper.taf?ABSTRACT_ID=324261.

433. Congress responded to the Enron debacle by passing the Sarbanes-Oxley Act of 2002. Pub. L. No. 107-204, 116 Stat. 745 (2002). Section 201 of Sarbanes-Oxley amends Section 10A of the Securities Exchange Act of 1934, 15 U.S.C. 78j-1 (2000), to add subsections (g) and (h), which make it unlawful for an accounting firm providing audit services to provide various non-audit services including "legal services."

434. Mitchell, *Incompetence, supra* note 12, at 105.

435. *Id.* at 105-06 (emphasis added).

under what circumstances they do not. Mitchell cannot, I assert, find a single legal decision theorist who believes that people think the same way in all settings.[436]

Mitchell's larger goal here is to point out some of the potential weaknesses of experimental research in the behavioral field and he raises issues worth exploring. He argues that for various reasons people will not necessarily make the same decisions in the real world as they make in the laboratory because of the difference in contextual factors. Therefore, he asserts, experiments done by behavioral psychologists are of limited value.[437] Arguably, Mitchell not only ignores the fact that legal decision theorists (unlike rational man economists) recognize that context matters and consider it when making their policy prescriptions, but also overstates his case, especially in light of all the real world confirmation that already exists for K-T Man.[438]

1. Accountability

Mitchell's first specific point is that accountability can affect decision makers. His implication, and it is a fair one, is that psychological experiments are often conducted in laboratory conditions where subjects have nothing at stake. If subjects are held accountable for their decisions in some fashion, that is, if they have to explain their decision, or are punished for bad decisions or rewarded for good decisions, they might perform differently. Noting that "any theory of legal decisionmaking that fails to incorporate accountability effects is materially incomplete,"[439] Mitchell argues that

> [L]ess careful scholars present accountability effects as either uniformly negative or positive, apparently as needed to bolster their particular arguments . . . . For instance, Professor Prentice, in a recent article arguing that self-serving judgmental biases pose serious problems in the work of auditors, portrays accountability conditions as producing uniformly negative effects on cognitive performance (that is, as exacerbating bias).[440]

I've been called worse than a "less careful scholar," but because it bolsters my point that Mitchell is the one who tends toward careless

---

436. Behavioralists can be accused of setting up straw men as well in the form of Chicago Man, which is so easy to debunk. However, Chicago Man (or some similar version of *homo economicus*) is the "standard model in economics." McFadden, *supra* note 6, at 75; *see also id.* at 74. ("The rational consumer model is so deeply entwined in economic analysis, and in broad terms so plausible, that it is hard for many economists to imagine that failures of rationality could infect major economic decisions or survive market forces.").

437. *See generally* Mitchell, *Incompetence, supra* note 12, at 105-09.

438. *See supra* notes 169-195 and accompanying text.

439. Mitchell, *Incompetence, supra* note 12, at 113.

440. *Id.* at 112-13.

characterizations, let me set out the full passage in my article to which Mitchell objected:

> Laboratory studies [showing that auditors act in a self-serving manner] are numerous, but they only go so far. These results may certainly be questioned because subjects are usually given anonymity [and are therefore shielded from accountability]. Lord has found evidence indicating that subjects given anonymity are less likely to issue a qualified opinion than subjects who were held accountable for their decisions. On the other hand, some studies show that imposing accountability produces few changes in result; furthermore, these laboratory studies are consistent with the self-serving behavior accounting firms exhibit in the real world [in studies explored in the next section].[441]

Even a quick reading of my passage clearly demonstrates that I did not claim that accountability effects are uniformly negative (or positive). Rather, I cited Lord's study,[442] which produced some evidence indicating that accountability *can minimize the effects of the self-serving bias*, and then noted that other studies have had less success in using accountability to debias decision making. In other words, I did just what Mitchell suggests that I should have done but claims that I did not—I presented evidence showing that the issue is unresolved.[443]

Mitchell criticizes me further for quoting two articles from the mid-1980s by Tetlock to support the point that accountability does not always work to debias decision making.[444] His implication is that these two articles are dated. This is a fair complaint, but much more recently, in 1999, Tetlock and a coauthor concluded that "[t]wo decades of research now reveal that (a) only highly specialized subtypes of accountability lead to increased cognitive effort; [and] (b) more cognitive effort is not inherently beneficial; sometimes it makes matters even worse. . . ."[445] Thus, new research strengthens, rather than weakens, the case I made.[446]

---

441. Prentice, *supra* note 211, at 1644-45 (citations omitted).

442. Alan T. Lord, *Pressure: A Methodological Consideration for Behavioral Research in Auditing*, AUDITING: J. PRAC. & THEORY, Fall 1992, at 89, 103, *cited in* Prentice, *supra* note 211, at 1643-44 & n.232.

443. Mitchell admits that several leading behavioral scholars (Seidenfeld, Langevoort, Rasmussen) have in their writings carefully noted the effects of accountability on decision making and tempered their conclusions accordingly. Mitchell, *Incompetence*, *supra* note 12, at 112 n.137.

444. *Id.* at 113 n.138 (referring to Philip E. Tetlock, *Accountability: The Neglected Social Context of Judgment and Choice*, 7 RESEARCH IN ORGANIZATIONAL BEHAVIOR 297 (1985) and Philip E. Tetlock & Richard Boettger, *Accountability: A Social Magnifier of the Dilution Effect*, 57 J. PERSONALITY & SOC. PSYCHOL. 388 (1989)).

445. Jennifer S. Lerner & Philip E. Tetlock, *Accounting for the Effects of Accountability*, 125 PSYCHOL. BULL. 255, 270 (1999).

446. Some studies show that requiring people to give reasons for their decisions may actually cause them to make *worse* decisions rather than better by inducing the decision makers to focus on factors that in the long run are not the important ones to them. *See, e.g.*, Timothy D. Wilson

Despite his mischaracterizations, Mitchell again does a real service by pointing out that under some circumstances accountability can affect decision making, sometimes improving it and sometimes exacerbating problems. Psychologists and legal decision theorists continue to work to learn how and under what circumstances decision making can be debiased.[447] This is important work, for if

---

& Jonathan W. Schooler, *Thinking Too Much: Introspection Can Reduce the Quality of Preferences and Decisions*, 60 J. PERSONALITY & SOC. PSYCHOL. 181, 185 (1991) (finding that analyzing reasons can lead to preferences and decisions that correspond less with expert conclusions); *see also* Oswald Huber & Gabriele Seiser, *Accounting and Convincing: The Effect of Two Types of Justification on the Decision Process*, 14 J. BEHAV. DECISION MAKING 69, 70 (2001) ("Several studies have found that the continuous commitment to a failing course of action (sunk cost effect) is *stronger* when decision makers know they have to justify their choice." (emphasis added)); Itamar Simonson & Peter Nye, *The Effect of Accountability on Susceptibility to Decision Errors*, 51 ORGANIZATIONAL BEHAV. & HUM. DECISION PROCESSES 416, 440 (1992) (reporting studies finding that accountability did not reduce irrational reliance on sunk costs, did not reduce preference reversals, but did reduce escalation of commitment).

Still, it must be emphasized that there are some situations in which some heuristics seem to be moderated by accountability. As Tetlock has also noted, "[a]ccountability manipulations are much more effective in attenuating certain cognitive biases—primacy, overattribution, overconfidence—when participants learn of being accountable prior to (as opposed to after) exposure to the evidence on which they are basing their judgments." Tetlock, *supra* note 216, at 589. On the other hand, Tetlock notes, accountability has had no effect regarding "insensitivity to base-rate information, giving more predictive weight to causal as opposed to merely statistical relationships between variables, preference reversals as a function of choice-versus-matching elicitation procedures, insensitivity to sample size, and the conjunction fallacy." *Id.* at 592. Finally, we cannot forget that accountability sometimes exacerbates bias. Tetlock reasons that in such cases, "efforts by subjects to put on the cognitive equivalent of their 'Sunday's best'—to become more thoughtful and to have a 'neat justification package' readily available—lead to judgments that depart ever more markedly from those that expert observers deem rational." *Id.*

See generally Robert de Hoog & Godfried van der Wittenboer, *Decision Justification, Information Structure and the Choice of Decision Rules*, *in* NEW DIRECTIONS IN RESEARCH ON DECISION MAKING 191, 202 (Berndt Brehmer et al. eds., 1986) (finding experimentally that "the necessity to justify one's decision vis a vis others has no effect on the kind of decision rules selected for arriving at a decision"); Huber & Seiser, *supra*, at 80 (finding that although accountability in the form of requiring subjects to justify their decision led to an increase in tbe amount of information used and to a more elaborate choice process, the decision heuristics did not change).

447. *See generally* Chapman & Johnson, *supra* note 236, at 125 (reviewing results of studies indicating that making subjects aware of the effects of an anchor did not decrease the anchor's affect on their judgments); Nisbett et al., *supra* note 203, at 526 (providing evidence for the argument that "statistical reasoning about everyday events should be highly trainable"); Norbert Schwarz & Leigh Ann Vaughn, *The Availability Heuristic Revisited: Ease of Recall and Content of Recall as Distinct Sources of Information*, *in* PSYCHOLOGY OF INTUITIVE JUDGMENT, *supra* note 8, at 103, 112-14 (reporting studies finding that forcing subjects in studies of the hindsight bias to argue against the inevitability of the reported outcome sometimes slightly attenuates the bias, but other times exacerbates it); Tversky & Kahneman, *supra* note 93, at 27 (finding that attempting to induce a reflective attitude by respondents did not reduce their susceptibility to the representativeness heuristic); Neil D. Weinstein & William M. Klein, *Resistance of Personal Risk Perceptions to Debiasing Interventions*, *in* PSYCHOLOGY OF INTUITIVE JUDGMENT, *supra* note 8, at 313, 322-23 (reporting results of experiments showing tbat various techniques for debiasing the overoptimism bias in regard to familiar health problems tended to be ineffective or to actually exacerbate the bias); Timothy D. Wilson et al., *Mental Contamination and the Debiasing*

improvements could be made, say to jury decision making in light of the hindsight bias, it would be all to the good. That is what legal decision theorists are attempting to accomplish.

2. Hypothetical Consequences

Mitchell's next point, which is related to his previous point, is the notion that laboratory experiments do not model real decision-making processes because they involve only hypothetical consequences. Here, at least, Mitchell admits that legal scholars applying behavioral principles have recognized this concern with behavioral research and have often designed their studies and qualified their conclusions so as to take it into account.[448] Once he recognizes that point, there is not much punch left in his argument because studies generally show that there is not much difference between laboratory decision making and real world decision making.[449]

Mitchell relies heavily on an excellent article by Hertwig and Ortmann that challenges psychologists to conform their research more closely to the conventions of experimental economics.[450] While there is no doubt much that psychologists can learn about research from experimental economists (and vice versa),[451] let us focus on Mitchell's specific point—that economists usually provide financial incentives to experimental subjects, while psychologists typically do not.

Problem, *in* PSYCHOLOGY OF INTUITIVE JUDGMENT, *supra* note 8, at 185, 192 (noting that efforts to debias judgments often lead to unnecessary correction or overcorrection, creating judgments that are no more accurate than the original debiased judgment).

448. Mitchell, *Incompetence*, *supra* note 12, at 114 n.141 (noting Garvin, Guthrie, and Korobkin).

449. *See supra* notes 161-201 and accompanying text. There certainly are studies showing differences in some contexts when incentives are introduced. *See, e.g.*, Vernon L. Smith & James M. Walker, *Monetary Rewards and Decision Cost in Experimental Economics*, 31 ECON. INQUIRY 245, 259 (1993) [hereinafter Smith & Walker, *Monetary Rewards*] (surveying thirty-one experimental studies showing the effects of monetary rewards and opportunity cost, finding some irrationality even in the presence of rewards but that several studies show improvement in the presence of rewards and in virtually all cases rewards reduced the variance of the data around the predicted outcome); Vernon L. Smith & James M. Walker, *Rewards, Experience and Decision Costs in First Price Auctions*, 31 ECON. INQUIRY 237, 244 (1993) [hereinafter Smith & Walker, *Experience*] (finding that increases in payoffs improved but did not eliminate nonoptimal bidding).

450. Ralph Hertwig & Andreas Ortmann, *Experimental Practices in Economics: A Methodological Challenge for Psychologists?*, 24 BEHAV. & BRAIN SCI. 383 (2001).

451. *See* Smith & Walker, *Monetary Rewards*, *supra* note 449, at 246 ("Our fundamental view [as experimental economists] is that the experimentalist has as much to learn from experimental subjects about subjective rationality, as human decision makers have to learn from the models that we call 'rational.'").

Why is "[t]he idea that preferences are only revealed by real incentives . . . deeply embedded in economists' worldview"?[452] A primary reason why economists usually use financial incentives in their experiments relates to the conventions of their field—"experimental economists who do not use [financial incentives] at all can count on not getting their results published."[453] Another reason is simply that "economic theory lends itself to straightforward translations into experiments employing financial incentives."[454]

Psychology, on the other hand, often studies the many decisions and behaviors that occur in the real world to which monetary incentives have little relevance.[455] To study these decisions using financial incentives is often "inappropriate, and not valid."[456] Health decisions, for example, are often made within parameters completely different than decisions involving money, so use of financial incentives in studies of such decisions would be counterproductive.[457] "Payoffs are useless when there is no right answer. This is typically true in studies of judgment, as opposed to decision making."[458] Indeed, psychologists often study the effects of various reward systems and values that stimulate behavior and to study them solely through a financial lens would not be viable.[459]

---

452. John K. Horowitz & K.E. McConnell, *Values Elicited From Open-Ended Real Experiments*, 41 J. ECON. BEHAV. & ORG. 221, 222 (2000).

453. Hertwig & Ortmann, *supra* note 450, at 390.

454. *Id.*

455. *See* Hasker P. Davis & Robert L. Durham, *Economic and Psychological Experimental Methodology: Separating the Wheat from the Chaff*, 24 BEHAV. & BRAIN SCI. 405, 406 (2001) ("[I]n the 'real world' people engage in all sorts of behaviors that are not monetarily relevant (e.g., memory tasks, problem solving, social interactions, child rearing, jury decision making, etc.).").

456. *Id.*

457. Tim Rakow has noted:

> [E]conomic theories of utility maximization are frequently applied to health care decisions. However, people readily recognize scenarios involving life expectancy and money as distinct classes of decision, and how they categorize decisions is seen to be related to their preferences. The parameters of decisions involving health (such as temporal discount rates) can be quite different from those involving money. Furthermore, contrary to predictions that might be made on the basis of experiments with financial incentives, people can be reluctant to trade or gamble life expectancy for improved quality of life. Thus, there is the possibility that an understanding of some classes of decisions are best served by experiments involving non-financial incentives.

Tim Rakow, *Theorize It Both Ways?*, 24 BEHAV. & BRAIN SCI. 425, 426 (2001) (citations omitted).

458. Jonathan Baron, *Purposes and Methods*, 24 BEHAV. & BRAIN SCI. 403, 403 (2001).

459. Davis & Durham, *supra* note 455, at 406 ("[S]ocial psychologists and personality theorists have already investigated the differential effects of different reward systems and values (intrinsic vs. extrinsic) on various behaviors or traits in order to determine their differential effects. They do, in fact, exist under some circumstances and not in others. To suggest that the study of the myriad of human activities is best exemplified by a single reward system is not a viable approach." (citations omitted)).

As Alvin Roth has noted:

Not only do many decisions not involve financial issues, but also all decisions are in a sense hypothetical. As Kuhberger notes:

> Why do psychologists helieve that the study of reactions in imagined situations is a legitimate means of studying real decision hehavior? To he sure, in other areas of psychology (for instance, in psychophysics), such methods would be considered extremely questionable if not absurd. The reason is that decision making—rather than, for example, perception—is hypothetical at its very core. When making a decision, we anticipate hypothetical states of the world, we consider events that could or could not obtain, we consider feelings we do not have yet. At the time of decision, none of these outcomes, events, or feelings, is real, but all are hypothetical. That is, in essence, decision making consists of the manipulation of hypothetical mental contents. Thus, decision researchers have some justification in assuming that people's real decisions can profitably be investigated by asking them to make hypothetical decisions.[460]

> [R]elying exclusively on purely monetary incentives would miss some important phenomena. Kahneman and Thaler, for example, showed that subjects who were given a coffee mug valued it more highly than subjects who had not been given one. To the extent that people treat possessions differently from money, this would have been a hard effect to observe if the only payoffs available to subjects had been monetary.

Alvin E. Roth, *Form and Function in Experimental Design,* 24 BEHAV. & BRAIN SCI. 427, 427 (2001).

460. Anton Kuhberger, *Why Use Real and Hypothetical Payoffs?,* 24 BEHAV. & BRAIN SCI. 419, 420 (2001); *see also* David B. Wiseman & Irwin P. Levin, *Comparing Risky Decision Making Under Conditions of Real and Hypothetical Consequences,* 66 ORGANIZATIONAL BEHAV. & HUM. DECISION PROCESSES 241, 243, 248-49 (1996) (noting that "past research has largely supported the position that the decision a person makes under hypothetical circumstances is a reasonably valid predictor of the decision that person would make in the same context with real consequences" and finding in their experiments that "subjects did not differ in their indicated preference . . . as a function of whether the consequence of their choice would or would not be incurred").

Financial incentives create their own problems in research design,[461] but the most important point is that, as a general rule, "real world" behavior is very similar to laboratory behavior. Numerous studies have found nonrational behavior even in the presence of financial incentives.[462] Camerer and Hogarth reviewed seventy-four studies and found that "[t]he most common result is that incentives did not

---

461. Goodie notes:

> [T]he question is not as simple as 'financial incentives versus no incentives.' In some settings, a financial incentive might be no incentive, and other incentives might be real incentives. More generally, choices motivated by points, tokens, the approval of the experimenter or the (learned or unlearned) rewards of correct answers per se are actual choices, not hypothetical ones. Such choices may or may not be as strongly motivated as those that earn money, but they are not less real.

Adam S. Goodie, *Are Scripts or Deception Necessary When Repeated Trials Are Used? On the Social Context of Psychological Experiments*, 24 BEHAV. & BRAIN SCI. 412, 412 (2001).

Harrison and Rutstrom note additionally:

> Consider, for example, the popular use of "lab dollars." These are a lab currency used in the experiment itself, and then converted to some local currency at the end of the experiment. Invariably, these lab dollars have lots of zeroes after them, so that instead of bidding $30, one might observe a subject bidding 30,000 "lab pesos." The purported reason for using this device is to give the subjects greater incentive to report monetary responses at a finer level of detail than if a field currency were used. The problem is that this will occur only if the subject suffers from some illusion with respect to the exchange rate between lab currency and field currency. Because such illusion is bound to vary across subjects, one has lost control over incentives. At the very least, the incentives will be much lower than intended, reducing saliency and increasing noise in the data. In the worst case, payoff dominance problems may cause results to be biased.

Glenn W. Harrison & E. Elisabet Rutström, *Doing It Both Ways—Experimental Practice and Heuristic Context*, 24 BEHAV. & BRAIN SCI. 413, 414 (2001).

462. *See, e.g.*, Linda Babcock et al., *Biased Judgments of Fairness in Bargaining*, 85 AM. ECON. REV. 1337 (1995) (finding a self-serving bias even when subjects were given a cash bonus for coming closest to view of objective third party); Jane Beattie & Graham Loomes, *The Impact of Incentives upon Risky Choice Experiments*, 14 J. RISK & UNCERTAINTY 155, 165 (1997) (concluding that the issue merited further research but finding that "the overwhelming weight of the evidence in this article supports Camerer's view that the absence (or presence) of financial incentives is not a crucial factor in encouraging (or discouraging) violations of standard axioms in pairwise choice problems"); Adam S. Goodie & Edmund Fantino, *An Experientially-Derived Base-Rate Error in Humans*, 6 PSYCHOL. SCI. 101, 105 (1995) (finding that adding financial incentives led to no significant improvement in subjects' base-rate neglect); Grether, *supra* note 393, at 555 (finding little evidence of reduction in the impact of the representativeness heuristic in the presence of incentives); Grether & Plott, *supra* note 292, at 632 (finding that incentives created stronger, not weaker, preference reversals); Horowitz & McConnell, *supra* note 452, at 235-36 (finding that even in the presence of financial incentives subjects ignored opportunity costs in conflict with economic logic); Amos Tversky & Daniel Kahneman, *Advances in Prospect Theory: Cumulative Representation of Uncertainty*, 5 J. RISK & UNCERTAINTY 297, 315 (1992) (finding, in study of choices between risky prospects, not "much difference between subjects who were paid a flat fee and subjects whose payoffs were contingent on their decisions"); Daniel J. Zizzo et al., *A Violation of the Monotonicity Axiom: Experimental Evidence on the Conjunction Fallacy*, 41 J. ECON. BEHAV. & ORG. 263, 273 (2000) (finding that monetary incentives and dynamic feedback did not reduce subjects' susceptibility to the conjunction fallacy).

affect mean performance."[463] In some judgment and decision tasks, incentives often hurt performance.[464] In some areas where concentration and attention are important, financial incentives often do improve decision making by increasing the subject's motivation and/or attention,[465] but Stone and Ziebart note that "financial incentives are no panacea for eliminating decision biases. Instead, incentives appear to increase the extent of attention given to a task, but also to increase potentially distracting emotions."[466] Kuhberger and colleagues list several examples of important areas where no significant differences were found between hypothetical and real decisions, and conclude "the general consensus among psychologists seems to be that hypothetical choices give a reasonable, qualitatively correct picture of real choices."[467] They then pursued their own study and found that similar preference reversals regarding framing were obtained using both hypothetical and real decisions for both small and large payoffs.[468] These results are supported by the numerous studies cited above where laboratory results were confirmed in the field.[469]

Interestingly, several studies have been done of decision making on television game shows where large financial incentives are involved; the results tend to mirror those found in laboratory

---

463. Colin F. Camerer & Robin M. Hogarth, *The Effects of Financial Incentives in Experiments: A Review and Capital-Labor-Production Framework,* 19 J. RISK & UNCERTAINTY 7, 22 (1999). In many of these studies, the incentives did reduce variation, although the mean was unaffected. *Id.* at 23.

464. *Id.* at 21.

465. Kahneman and Tversky recognized this long ago when they wrote, "[i]ncentives do not operate by magic. They work by focusing attention and by prolonging deliberation. Consequently they are more likely to prevent errors that arise from insufficient attention and effort than errors that arise from misperception or faulty intuition." Amos Tversky & Daniel Kahneman, *Rational Choice and the Framing of Decisions, in* CHOICES, VALUES, AND FRAMES, *supra* note 8, at 209, 222.

466. Dan N. Stone & David A. Ziebart, *A Model of Financial Incentive Effects in Decision Making,* 61 ORGANIZATIONAL BEHAV. & HUM. DECISION PROCESSES 250, 259 (1995). The authors conclude that "after controlling for these mediating variables financial incentives had no effect on choice accuracy." *Id.* at 258. In other words, in their experiment the authors found that financial incentives increased the amount of information processed, improved the sequence, and improved the variability. On the other hand, the incentive also increased the negative effects and the negative consequences of the latter canceled out the positive benefits of the former.

467. Kuhberger et al., *supra* note 17, at 1164; *see also* David A. Asch et al., *Omission Bias and Petussis Vaccination,* 14 MED. DECISION MAKING 118, 121 (1994) (discussing a similar study in the real world regarding parents' decisions to vaccinate their children with DPT and finding a similar bias); Ilana Ritov & Jonathan Baron, *Reluctance to Vaccinate: Omission Bias and Ambiguity,* 3 J. BEHAV. DECISION MAKING 263, 275 (1990) (finding a strong omission bias in decision to vaccinate).

468. Kuhberger et al., *supra* note 17, at 1167-70.

469. *See supra* notes 169-195 and accompanying text.

experiments with small or no financial incentives.[470] Similarly, psychologists and others have performed studies in third world nations where they could afford to offer monetary incentives that were meaningful to the subjects being studied; again, the subjects' decision making tended to be subject to the same heuristics and biases as the decision making of subjects in United States laboratory experiments.[471] Camerer and Hogarth have done the most extensive study, which Mitchell quotes extensively and fairly.[472] They tell us:

> Critics and referees very commonly assert that if the stakes were just high enough the rationality rejection would disappear. While several studies have tried to make rationality violations disappear—in utility theory paradoxes, ultimatum bargaining, and voting experiments—none have succeeded in clearly overturning anomalies. . . . [S]ince *all* established anomalies have survived these kinds of hostile attacks, uninformed critics should quit talking as if simply raising the stakes would make the effects disappear. So far, that hasn't proved true; and nothing in any sensible understanding of human psychology suggests it would.[473]

---

470. *See e.g.*, Roel M.W.J. Beetsma & Peter C. Schotman, *Measuring Risk Attitudes in a Natural Experiment: Data from the Television Game Show Lingo*, 111 ECON. J. 821, 846-47 (2001) (finding risk aversion and overconfidence in Dutch television game show); Randall W. Bennett & Kent A. Hickman, *Rationality and the "Price Is Right,"* 21 J. ECON. BEHAV. & ORG. 99, 105 (1993) ("Evidence from 'The Price Is Right' game show documents persistent decision-making errors despite the penalties associated with the sub-optimal behavior and the ample opportunity for contestants to formulate strategy before-hand."); Jonathan Berk et al., *The Price Is Right, But Are the Bids? An Investigation of Rational Decision Theory*, 86 AM. ECON. REV. 954, 965-66 (1996) (reporting results "consistent with the hypothesis that behavior on *The Price Is Right* is explained by bounded rationality" rather than by rational decision theory); Robert Gertner, *Game Shows and Economic Behavior: Risk-Taking on "Card Sharks,"* 108 Q.J. ECON. 507, 507 (1993) (finding that contestants on "Card Sharks" game show "do not act in a manner consistent with expected-utility maximization"); Andrew Metrick, *A Natural Experiment in 'Jeopardy!'*, 85 AM. ECON. REV. 240, 252 (1995) (reporting that a study of performances on the "Jeopardy" game show demonstrated that "[s]uboptimal choice can persist despite the three mitigating factors of high stakes, an identifiable market mechanism, and an opportunity for players to learn"); MARCO HAAN ET AL., THE WEAKEST LINK: A FIELD EXPERIMENT IN RATIONAL DECISION MAKING 31 (Univ. of Groningen, Mimeo Working Paper, Feb. 2002) (finding that contestants on "The Weakest Link" act in a boundedly rational manner and exhibit fairness concerns inconsistent witb the rational man model), http://papers.ssrn.com/sol3/papers.cfm?abstract_id=303163.

471. Steven J. Kachelmeier & Mohamed Shehata, *Examining Risk Preferences Under High Monetary Incentives: Experimental Evidence from the People's Republic of China*, 82 AM. ECON. REV. 1120, 1137 (1992) (finding a systematic difference between low- and high-prize lotteries in some respects, but reporting that "even at the highest prize level used in this study, we still find a pattern of strongly risk-seeking preferences for low-probability gain prospects, consistent with the predictions of Tversky and Kahneman"). Binswanger, in less sophisticated experiments, did find that Indian peasants became more risk-averse in high prize conditions. *See* Hans P. Binswanger, *Attitudes Toward Risk: Experimental Measurement in Rural India*, 62 AM. J. AGRIC. ECON. 395 (1980); Hans P. Binswanger, *Attitudes Toward Risk: Theoretical Implications of an Experiment in Rural India*, 91 ECON. J. 867, 888 (1981).

472. Mitchell, *Incompetence*, *supra* note 12, at 117-19.

473. Camerer & Hogarth, *supra* note 463, at 33-34. Summarizing both the Hertwig and Ortmann study and the Camerer and Hogarth study, Hilton recently noted:

> Experimental incentives do sometimes have an effect of improving rationality, but not always (only in 23 out of the 43 studies created by combining the Hertwig-Ortmann

If Mitchell is simply reminding us that for some mental tasks, incentives can reduce departures from the norm, that psychologists would do well to study phenomena with financial incentives as well as those without, and that we should pay attention to which types of behavior are affected by financial incentives and which are not, then he is doing a fine service.[474] If he purports to undermine in any substantial way the validity of behavioral research, including the Kahneman-Tversky line, he fails.[475] As Sunstein recently noted, "the heuristics operate even when the stakes are large."[476]

> and Camerer-Hogarth reviews). . . . In seven others there was no effect of incentives, and in a remarkable 13 cases they had negative effects. Indeed, given that Wall Street bond traders dealing day after day in millions of dollars show irrationalities predicted by prospect theory it would be surprising if small experimental learning and incentives eliminated irrationality.

Denis J. Hilton, *Is the Challenge for Psychologists to Return to Behaviorism?*, 24 BEHAV. & BRAIN SCI. 415, 415 (2001) (citing Z. Shapira, *Aspiration Levels and Risk Taking: A Theoretical Model an Empirical Study on the Behavior of Government Bond Traders* (2000) (unpublished manuscript, New York University).

474. One of the most interesting aspects of the impact of financial incentives was recently explored by Sunstein in connection with Professor Solomon Asch's famous experiment on conformity. SUNSTEIN, *supra* note 181, at 13-15. Asch found that large percentages of subjects could be induced to say that one line was longer than an obviously shorter line if a number of the experimenter's confederates first stated the same erroneous conclusion. *See* SOLOMON E. ASCH, SOCIAL PSYCHOLOGY (1952); Solomon E. Asch, *Studies of Independence and Conformity: A Minority of One Against a Unanimous Majority,* 70 PSYCHOL. MONOGRAPHS NO. 416 (1956). Studies of incentives show that the rate of conformity is decreased when monetary incentives are added *if* the task is easy, but *increased if the task is hard. See* Robert Baron et al., *The Forgotten Variable in Conformity Research: Impact of Task Importance on Social Influence,* 71 J. PERSONALITY & SOC. PSYCHOL. 915, 924 (1996) (reporting studies indicating that "heightening incentives for accuracy actually heightened participants' susceptibility to an inaccurate group consensus"); SUNSTEIN, *supra* note 181, at 13-15 (discussing these and related studies).

475. *See also* Maya Bar-Hillel and Efrat Neter, *How Alike Is It? Versus How Likely Is It?: A Disjunction Fallacy in Probability Judgments, in* PSYCHOLOGY OF INTUITIVE JUDGMENT, *supra* note 8, at 82, 96 (finding disjunction fallacy committed by 74% of those studied in a hypothetical betting condition and by 72%, a small and insignificant difference, of those studied in a real betting condition where it was in the subjects' best interests to find the right answer rather than the answer they believed the experimenters desired); Roger Buehler et al., *Inside the Planning Fallacy: The Causes and Consequences of Optimistic Time Predictions, in* PSYCHOLOGY OF INTUITIVE JUDGMENT, *supra* note 8, at 250, 261-62 (reviewing substantial evidence indicating that incentives actually exacerbate the "planning fallacy," the tendency to be overoptimistic in prediction of completion times for future events); Chapman & Johnson, *supra* note 236, at 125 ("[T]he evidence about the influence of incentives on anchoring is mostly negative."); David M. Grether, *Testing Bayes Rule and the Representativeness Heuristic: Some Experimental Evidence,* 17 J. ECON. BEHAV. & ORG. 31, 54, 56 (1992) (finding that financial incentives had little effect on base rate neglect, concluding that "modeling and understanding behavior under uncertainty is a more demanding and difficult task than many economists have thought"); Schwarz & Vaughn, *supra* note 447, at 117 (reviewing the results of several studies showing that "incentives rarely attenuated reliance on the availability heuristic"); Slovic & Lichtenstein, *supra* note 291, at 596 (noting that incentives do little to diminish preference reversals); Tversky & Kahneman, *supra* note 93, at 33 (reporting study finding that a conjunction error was made by 65% of subjects playing an experimental game with real payoffs and by 62% of subjects playing the same game with only hypothetical payoffs); J. Frank Yates et al., *Probability Judgment Across Cultures, in*

Addressing accountability, incentives, and other forms of debiasing, Tetlock, one of Mitchell's favorite sources, warned skeptics such as Michell in 2002 that

> [E]fforts at debiasing have thus far had mixed success. Kahneman and Tversky anticipated these results when they wrote 15 years ago: "Incentives do not operate by magic. They work by focusing attention and by prolonging deliberation. Consequently they are more likely to prevent errors that arise from insufficient attention and effort than errors that arise from misperceptions or faulty intuition." In this view, motivating people to think harder will often backfire, amplifying biases rather than attenuating them. Effective learning takes place only under difficult-to-satisfy conditions: It requires accurate and timely feedback about the relations between antecedent conditions and appropriate responses. Skeptics (and we have some sympathy with this camp) have an uphill battle.[477]

C. *Implications for Legal Theory*

In the final section of his *Equal Incompetence* article, Mitchell assesses the implications that his arguments have for application of legal decision theory to important legal issues.[478]

1. The Realism Versus Parsimony Trade-Off

Mitchell quotes Jolls, Sunstein, and Thaler who proclaim the goal of many behavioralists: "We believe that a behavioral approach imposes discipline on economic theorizing because assumptions cannot be imported at will. In a behavioral approach, assumptions about behavior should accord with empirically validated descriptions of actual behavior."[479]

In other words, legal decision theorists seek to make policy prescriptions based on psychology's best understanding of how people actually make decisions, rather than based upon a stylized model that few people believe represents how people actually make decisions. Mitchell argues that this cannot be done because the variations in individual reasoning that he highlights undermine the "equal incompetence" view that he imputes to legal decision theorists.[480] Not to beat a dead horse, but legal decision theorists do not believe in

---

PSYCHOLOGY OF INTUITIVE JUDGMENT, *supra* note 8, at 271, 274 (reviewing study indicating that financial bonuses did not affect patterns of cross-cultural overconfidence bias).

476. SUNSTEIN, *supra* note 272, at 6; *see also* Kuhberger et al., *supra* note 17, at 1170 ("Our data pose a problem [for economists] because the framing effect is an economic anomaly, but it does not disappear with larger payoffs—if anything, it becomes stronger.").

477. Philip E. Tetlock & Barbara A. Mellers, *The Great Rationality Debate*, 13 PSYCHOL. SCI. 94, 97 (2002) (citations omitted).

478. *See* Mitchell, *Incompetence*, *supra* note 12, at 119-39.

479. Jolls et al., *supra* note 310, at 1489, *quoted in* Mitchell, *Incompetence*, *supra* note 12, at 120.

480. Mitchell, *Incompetence*, *supra* note 12, at 83-88.

equal incompetence. If they did, it would seem that Mitchell would be able to point to an erroneous policy prescription based on such a belief, but he cannot.[481]

Mitchell does make the point that because behavioral and cognitive research on human judgment and decision making proves that we live in a very complicated world, there are limitations on how far psychological research on behavior can take us. He argues that in the complexity/parsimony tradeoff, legal decision theorists will, by basing their policy prescriptions on greater realism, forfeit predictability.[482] But it is not necessarily true that the most parsimonious account of the facts leads to the best policy prescription.

We do not and cannot predict that all people will in all circumstances act in a self-interested fashion, but we can predict that the bigger the audit client, the more likely the audit firm will cave in

---

481. Content to launch broadsides against unspecified behavioralist policy prescriptions, Mitchell is unable to land a blow against any particular prescription and, indeed, mentions only a few even in passing. He does note that Sunstein suggests that employment and labor law issues be reexamined in light of behavioral research, Cass R. Sunstein, *Human Behavior and the Law of Work*, 87 VA. L. REV. 205 (2001), *cited in* Mitchell, *Incompetence, supra* note 12, at 86, and that Hanson and Kysar urge that all legal concepts that are premised on the assumption of a rational decision maker be reevaluated. Hanson & Kysar, *supra* note 395, at 634, *cited in* Mitchell, *Incompetence, supra* note 12, at 86. Mitchell provides absolutely no specific argument that would indicate these are not good ideas. He works hard to qualify assumptions of perfect irrationality that he (falsely, in my view) imputes to Sunstein and Hanson and Kysar, but he does not in any way show that these authors are wrong. Mitchell, *Incompetence, supra* note 12, at 86 nn.44-45.

Mitchell notes that Chris Guthrie has noted that regret aversion theory can help us to understand litigation behavior. Guthrie, *supra* note 411, at 43. Again, Mitchell has no critique of Guthrie's analysis or policy suggestions; he simply observes that "[b]y bringing regret aversion into the decision calculus, we complicate the model of litigation behavior but also may gain descriptive accuracy, for the avoidance and achievement of certain affective states appears to be an important part of many decisions, particularly settlement decisions." Mitchell, *Incompetence, supra* note 12, at 103-04. Believing that descriptive accuracy is a necessary predicate for useful policy prescriptions, this is a major goal of most legal decision theorists. Again, Mitchell can offer no substantive critique.

In an appendix to his article, Mitchell returns to Hanson and Kysar, mentioning that research regarding "[t]he potential ability of older adults to resist market manipulation bears directly on the claims of legal decision theorists Jon Hanson and Douglas Kysar, who call for an enterprise liability law based on the claim that companies successfully manipulate consumer preferences and spending by exploiting prevalent cognitive biases and errors in consumers." Mitchell, *Incompetence, supra* note 12, app. A at 158 (citing Hanson & Kysar, *supra* note 404). Again, Mitchell does not directly challenge Hanson and Kysar's well-supported argument, which was recently bolstered by one of Kahneman and Tversky's leading critics, Gerd Gigerenzer. *See* GIGERENZER, *supra* note 295, at 29-30 (explaining how the tobacco industry successfully misled the public regarding the dangers of smoking). This is not to say that the policy prescriptions of legal decision theorists cannot be challenged; they can. The point is that Mitchell apparently cannot challenge them on grounds that they do not take into account the variability of human thinking.

482. Mitchell, *Incompetence, supra* note 12, at 83.

when the client resists a suggested audit adjustment;[483] that the more important the client and the less the litigation risk, the less likely the audit firm will issue a qualified opinion;[484] that the more the auditor is worried about losing the client, the less likely it will issue a going concern qualification;[485] and that litigation risk does not tend to induce auditors to resign accounts *unless* the client is in poor financial shape.[486] Ultimately, behavioral research shows us that auditors' professional skepticism tends to be "counteracted when the client [is] important to the audit firm's practice development."[487] Universality and uniformity of action are not prerequisites to helpful policy insights. Every auditor need not act the same in every circumstance for us to realize that in cases of major audit failure, judicial presumptions that auditors are rational and would not endanger their reputations by stretching professional rules are questionable, and that restrictions upon non-audit services offered by audit firms are worthy of serious consideration.

Furthermore, we can predict that most consumers approaching a rental car counter will sign the form contract placed in front of them rather than bargain for their preferred level of risk or for an optimally efficient state as assumed by economists. Consider, for example, my arguments regarding contract-based defenses in securities fraud litigation.[488] In response to case law emanating from law and economics judges holding that investors should be allowed to contractually waive their right to sue for fraud,[489] I sought to use behavioral literature to undermine the economists' assumption that the reason investors or consumers sign such contracts is that they are voluntarily contracting for their desired level of risk.[490] I argued that

483. *See* Wright & Wright, *supra* note 425, at 22 ("Of concern was that a number of adjustments greater than planning materiality were subsequently waived (47.5%), suggesting the potential for financial reporting risk.").

484. *See* Krishnan & Krishnan, *supra* note 426, at 583.

485. *See* Paul Barnes & Hooi Den Huan, *The Auditor's Going Concern Decision: Some UK Evidence Concerning Independence and Competence,* 20 J. BUS. FIN. & ACCT. 213, 226 (1993) (concluding that economic pressures cause auditor reluctance to issue justified going concern qualifications); Thomas Kida, *An Investigation into Auditors' Continuity and Related Qualification Judgments,* 18 J. ACCT. RES. 506, 516 (1980) (finding that auditors can usually tell when a going concern qualification is justified).

486. *See* Scholz, *supra* note 429.

487. Michael K. Shaub & Janice E. Lawrence, *Ethics, Experience and Professional Skepticism: A Situational Analysis,* 8 BEHAV. RES. ACCT. 124, 155 (Supp. 1996).

488. Prentice, *supra* note 258.

489. *See, e.g.*, Rissman v. Rissman, 213 F.3d 381, 387 (7th Cir. 2000); Carr v. CIGNA Sec., Inc., 95 F.3d 544, 548 (7th Cir. 1996).

490. *See, e.g.*, *Session Three: Discussion of Paper by George L. Priest, Yale University,* 10 CARDOZO L. REV. 2329, 2339 (1989) (quoting the statement of Peter Huber), *cited in* Prentice, *supra* note 258, at 342.

there are many more plausible explanations for why people sign form contracts that waive their right to sue for fraud,[491] including:

- Rational Ignorance. Much evidence suggests that people rationally believe that it will not be cost effective for them to read most form contracts before they sign them. The forms are time-consuming to read, hard to understand, and the seller's agent probably has no authority to alter them anyway.[492]
- Overoptimism, Overconfidence, and the Illusion of Control. People tend toward irrational optimism and overconfidence, including when they act as investors or consumers. Therefore they underestimate the extent to which they are at risk of being defrauded. This phenomenon is exacerbated by the illusion of control, people's belief that they can exert control over purely random events. Empirical studies show that people tend to believe that the terms of the contracts they sign are more favorable than they truly are.[493]
- Probabilities and Future Events. The behavioral literature also indicates that people are not good at calculating probabilities in general and specifically tend to ignore low probability risks, such as the risk of being defrauded.[494]
- False Consensus Effect and Personal Positivity Bias. Much psychological evidence indicates that people tend to believe that others see the world as they do (false consensus effect) and generally perceive other people in a positive light, often naively. Therefore, people who are honest tend to believe that others are treating them honestly. When the concept of cognitive dissonance (the tendency to suppress information inconsistent with positions taken in order to preserve psychological consistency)[495] is factored in, people are especially reluctant to reach the conclusion that they have made a mistake in deciding to trust a promoter or stock broker

491. *See* Russell Korobkin, Bounded Rationality and Unconscionability: A Behavioral Theory of Policing Form Contracts (UCLA Law & Econ. Research Paper No. 03-02, Jan. 2003) (exploring through a behavioral lens why people sign form contracts that waive their right to sue for fraud), http://papers.ssrn.com/abstract=367172.

492. Prentice, *supra* note 258, at 358-62.

493. *Id.* at 362-63.

494. *Id.* at 363-64.

495. *See, e.g.*, LEON FESTINGER, A THEORY OF COGNITIVE DISSONANCE (1957); PLOUS, *supra* note 135, at 22-30 (explaining concept generally).

who has sold them stock or a developer who has sold them a house.[496]

- Inability to Detect Deception. Again, there is substantial empirical evidence that people are unable to detect when they are being deceived, but, worse still, inaccurately believe that they can do so. This leaves investors and consumers particularly prone to being defrauded when they run into real crooks.[497]
- Insensitivity to Source. Psychological evidence shows that one of the reasons that people are poor lie detectors is that they have trouble disregarding information even when they learn that its source is questionable. Once they decide to trust a stockbroker, they tend to continue to trust long after suspicious facts arise.[498]
- Salience of Oral Communications. Because people's minds are more attuned to oral communications than written communications, the oral promises made by stockbrokers or car salesmen have more impact than the written contract provisions that disclaim those promises.[499]
- Status Quo Bias. Because people prefer the status quo, when form givers (merchants/stockbrokers) offer adhesion contracts to form takers (consumers/investors), the form takers will be reluctant to attempt to alter what they view as the status quo, particularly because a dense form contract has an "authoritative legality" about it.[500]
- Social Proof. Social proof is the notion that people tend to take their cues for proper behavior from the actions of those around them, and therefore they will be hesitant to reject a form contract that they know most people in their situation typically sign without even reading.[501]
- Anchoring and Adjustment. Related to the status quo bias is the anchoring and adjustment heuristic—the tendency to be heavily influenced by initial information and then to fail to adjust sufficiently to account for new

496. Prentice, *supra* note 258, at 364-65.

497. *Id.* at 366-67.

498. *Id.* at 367-69.

499. *Id.* at 369-71.

500. *Id.* at 371-73 (quoting G. Richard Shell, *Fair Play, Consent and Securities Arbitration: A Comment on Speidel*, 62 BROOK. L. REV. 1365, 1368 (1996)).

501. *Id.* at 373.

information. Thus, when investors or consumers have their expectations set by oral representations made by sellers, their views will likely not sufficiently adjust toward the reality spelled out in the fine print of the written contracts they are presented.[502]

- Regret Theory. Because substantial psychological evidence shows that (a) regret is an unpleasant emotion, (b) people desire to avoid regret, (c) people factor that desire into their decision making, and (d) people suffer more regret when bad consequences result from their active decisions than from passive decisions and from abnormal situations than from normal situations, anticipated regret inclines decision makers to passively accept the normal situation which is embodied in the typical form contract they are presented.[503]

Certainly there is variation from person to person and situation to situation in the impact of these behavioral inclinations and cognitive illusions. My argument simply does not depend on all of them affecting all people equally in all situations. These behavioral and cognitive considerations combine, I believe, to provide a much more compelling explanation of why most people sign form contracts that waive their right to sue for fraud than the economic explanation, which assumes that they are rationally bargaining for their desired level of risk. Ultimately, the proof is in the pudding. I demonstrated my point that these problems are systematic by showing that contracts for the sale of new automobiles (before the courts began applying the strict liability theory)[504] essentially disclaimed all meaningful liability. The same was true of contracts for the sale of securities (before Congress intervened by passing the federal securities laws).[505] No matter what people's risk profiles were, they tended to sign such waivers.

Consider that it is much safer and minimally inconvenient to wear a seat belt. Chicago Man would buckle up every time.[506] Yet, for

502. *Id.* at 373-74.

503. *Id.* at 374-78.

504. *Id.* at 388.

505. *Id.* at 388-89.

506. *See* Carol M. Ostrom, *The Risk Takers—In Search of the Next Thrill: Bravery and Biochemistry,* SEATTLE TIMES, Oct. 20, 1996, at 20 ("If fear of risks were rational, . . . [w]e'd never, ever lapse on buckling our seatbelts."), 1996 WL 3687564; *National Press Club Luncheon with Cynthia Trudell,* FED. NEWS SERVICE, Aug. 18, 1999 (quoting Ms. Trudell, Chair and President of Saturn Corporation, as saying that wearing seatbelts is so "logical and rational," yet when they first came out it was difficult to induce people to wear them).

behavioral reasons, most people did not wear seatbelts until the law required them to do so.[507] The behavioral factors (overconfidence, overoptimism, illusion of control, and the like) do not have to act equally upon all decision makers in all circumstances for behavioral insights to be helpful and for rules requiring the wearing of seatbelts to be well justified.[508]

Simply put, universal and uniform use of decisional heuristics and vulnerability to cognitive biases are not required for research in judgment and decision making to lend valuable insights to legal policy questions.

2. Mitchell's Suggestions

Apparently Mitchell is not ready to completely give up on legal decision theorists, for he proceeds to suggest ways that they can improve their application of behavioral decision theory to legal issues.[509] All are unobjectionable "best practices" sort of suggestions.[510] Mitchell builds to the conclusion that:

---

507. Don Colburn, *Seat Belts, Survival and the Law*, WASH. POST, Feb. 13, 1985, at Z07 (noting that before seatbelt laws were passed, "[d]espite all the grim evidence and multimillion dollar safety campaigns with slogans like 'Buck up for safety' and 'Lock it to me,' seat belt use nationally . . . hovered below 15 percent").

508. *See Serfdom and Seat Belts*, NEW REPUBLIC, June 3, 1985, at 4, 42 ("[F]ew people wbo choose not to buckle up have made a rational calculation that the added risk to them is not worth the inconvenience.").

509. Mitchell, *Incompetence, supra* note 12, at 125-132 (instructing legal decision theorists to: (1) consult original sources and avoid overreaching in your cbaracterization of results from this work; (2) follow American Psychological Association guidelines for reporting methods and results of empirical work; (3) if not trained in the social sciences, avoid unaided application of behavioral theory to the law; and (4) use multiple methodologies as much as possible when doing empirical research).

510. One of Mitchell's suggestions hits close to home. He argues that although law professors may be "wonderful autodidacts," those untrained in the social sciences should "refrain from behavioral decision theory's unaided application to the law" but may be encouraged to engage in interdisciplinary research projects. Mitchell, *Incompetence, supra* note 12, at 127. As an autodidact, I have taken Mitchell's suggestion to do interdisciplinary research with those who are trained in the social sciences. *See, e.g.*, Prentice & Koehler, *supra* note 405. However, in my other articles cited throughout this article, I have not refrained from attempting to apply behavioral decision theory to the law. I have no doubt that I am not as qualified to examine this area as is Mitchell, and I am certain that my lack of formal training has led or will lead me to make mistakes that those better trained would avoid (although the only mistake Mitchell accuses me of so far is his rather than mine, *see supra* notes 440-446).

Just as psychologists can prevent lawyers from making errors regarding matters of psychology, lawyers' understanding of legal issues and institutions enables them to offer contributions that those trained only in psychology cannot. *See* Blumenthal, *supra* note 1, at 35 (suggesting that "social scientists are often equally ignorant about fundamental legal issues that render their research if not irrelevant, then difficult for the law to apply"). Still, I admit that the contributions will likely be of a higher quality when made by academics trained in the area, such as my occasional co-author Jonathan Koehler at Texas, Jeff Rachlinski at Cornell, and, yes,

> Given the applied nature of legal decision theory, the primary goal should arguably be to explain and predict behavior in discrete situations to the greatest extent possible rather than seek to build an overarching theory of legal behavior whose variables explain a small percentage of variability in many situations but a large percentage in none. For instance, rather than being concerned about how the hindsight bias may operate across legal cases in general, we should focus on identifying what particular legal judgments and what legal actors are most prone to the hindsight bias in what types of cases.[511]

This is the trajectory that legal decision theory has taken since Jolls, Sunstein and Thaler wrote their initial survey article,[512] and it is the course that I expect we shall continue to see in the future.

3. First, Do No Harm

Mitchell implies that cognitive psychologists and legal decision theorists should not make any public policy suggestions unless they are certain that their programs will not be counterproductive.[513] The "first do no harm" principle is generally good advice for physicians in light of the self-healing properties of the body, but it is questionable in the public policy context unless one holds the dubious belief that the law has similar self-correcting properties.[514]

Mitchell goes on to make a surprising concession—that some modest improvements in the law might well be made based on behavioral insights. For example, Mitchell suggests that we might (1) train judges and bureaucrats in statistics, because research from psychology demonstrates that often a relatively small amount of training can dramatically improve probabilistic reasoning;[515] (2) require that questions and evidence be presented to judges in the form of frequencies rather than probabilities, because evidence from psychological experiments shows that humans often handle frequencies better than probabilities;[516] (3) ask legal decision makers to explicitly consider alternative or opposing evidence and arguments, because psychologists' studies show that this process can help debias some of the prominent K-T shortcomings in human reason;[517] and (4) ask legal actors to explain their choices, because requiring subjects in

---

Gregory Mitchell at Florida State. Just as law schools began hiring more faculty trained in economics as the law and economics movement bloomed, so they will likely hire more faculty with psychology training in the near future.

511. Mitchell, *Incompetence, supra* note 12, at 130.

512. Jolls et al., *supra* note 310, at 1471.

513. Mitchell, *Incompetence, supra* note 12, at 132.

514. I thank Jeffrey Rachlinski for pointing this out to me.

515. Mitchell, *Incompetence, supra* note 12, at 132.

516. *Id.* at 133.

517. *Id.* at 133-34.

psychology experiments to provide such rationales has been shown to reduce framing effects.[518]

Mitchell is completely right that all of these suggestions are worth exploring. He is, on the other hand, unduly timid in failing to embrace numerous other policy prescriptions offered by legal decision theorists. Mitchell's stated basis for his reticence is reasonable enough. In his mind, the battle between Chicago Man and K-T Man breaks down to an "unavoidable ideological debate."[519] On the one hand, law and economics represents a conservative, free market point of view. Behavioralists represent a more liberal, paternalistic/activist government approach.[520] Only raw politics can break the deadlock, Mitchell implies.

This dichotomy is arguably unfair to both the Chicago Man advocates and those in the K-T Man camp. Admittedly, the policy ramifications of a law and economics approach tend to counsel against government involvement in the affairs of men, but academics of all political stripes find some versions of economic analysis appealing.[521] Even less than law and economics, legal decision theory seems to belong to no particular political camp. As Jolls and her coauthors pointed out, their view is that behavioral analysis is not so much paternalistic as it is simply "anti-antipaternalism."[522] Behavioral analysis refutes the arguments that economists have used to challenge existing governmental paternalism. At this stage in its development,

---

518. *Id.* at 134-35.

519. *Id.* at 137.

520. *Id.* Mitchell cites Tetlock who recently noted, regarding the fundamental attribution error (which is "the tendency to attribute behavior to the actor's dispositions and to ignore powerful situational determinants of the behavior." RICHARD NISBETT & LEE ROSS, HUMAN INFERENCE: STRATEGIES AND SHORTCOMINGS OF SOCIAL JUDGMENT 31 (1980)) that:

> Authoritarian conservatives deem it prudent managerial practice to communicate to subordinates a low tolerance for justifications and excuses that invoke situational causes for conduct that falls short of organizational expectations. People will be more motivated to behave properly if they believe that improper behavior will almost automatically tarnish their reputations—a social variant of the doctrine of strict liability. From an authoritarian-conservative perspective, failing to hold people responsible for outcomes that they could have controlled is arguably a more serious error than holding people responsible for outcomes that they could not control. By contrast, the anti-authoritarian egalitarians see the fundamental attribution error as punitive, not prudent. They disagree with conservatives about both the frequency with which subordinates will invent specious justifications and fictitious excuses for substandard performance and the relative importance of avoiding Type I errors (condemning the innocent) versus Type II errors (acquitting the guilty), deploring the former error to a greater degree).

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, 320-21 (2000).

521. *See* Wetlaufer, *supra* note 26, at 38 (comparing conservative versus liberal views of the use of economic analysis).

522. Jolls et al., *supra* note 310, at 1545.

it may seem to be paternalistic and to support politically liberal points of view, but only because thus far it has been used primarily in reaction to Chicago School teachings. As it becomes less reactive and more constructive,[523] as I think it will, legal decision theory will seem less ideological. For example, Kahneman and Sunstein, with others, recently made a persuasive psychology-based case for punitive damages reform,[524] an agenda item clearly dearer to conservatives than liberals. And Camerer and his colleagues presented an "asymmetric paternalism" rationale designed to tailor behavioralism's application in such a way as to appeal even to conservatives.[525] Indeed, after several years of writing conservative policy prescriptions in the securities field, some of which suffered from a lack of insight into the psychology of human decision making,[526] Stephen Choi and Adam Pritchard recently published a stinging critique of the Securities and Exchange Commission, reasonably pointing out that governmental decision makers are likely to suffer from tbe same limitations as private decision makers.[527] Perhaps we're all behavioralists now.

D. Final Insights

There are many useful insights in Mitchell's second article, but legal decision theorists do not hold the belief that he imputes to them—that people are universally and uniformly irrational. Nor do their policy prescriptions reflect such a belief. Nor need such be the state of the world for psychological research to usefully inform legal analysis.

---

523. Cunningham admits that if behavioral theory simply settles for refuting law-and-economics arguments, it will not fulfill its promise; he uses behavioral finance theory to furnish a positive theory of market behavior with important implications for corporate and securities law. *See* Lawrence A. Cunningham, *Behavioral Finance and Investor Governance*, 59 WASH. & LEE L. REV. 767, 837 (2002).

524. *See generally* CASS R. SUNSTEIN ET AL., PUNITIVE DAMAGES: HOW JURIES DECIDE (2002).

525. *See* Colin Camerer et al., *Regulation for Conservatives: Behavioral Economics and the Case for "Asymmetric Paternalism,"* 151 U. PA. L. REV. 1211, 1212-13 (2003) ("A regulation is asymmetrically paternalistic if it creates large benefits for those who make errors, while imposing little or no harm on those who are fully rational.").

526. *See, e.g.*, Stephen Choi, *Regulating Investors Not Issuers: A Market-Based Proposal*, 88 CAL. L. REV. 279, 333-34 (2000) (suggesting that securities professionals be virtually totally deregulated and that investors be regulated instead); A.C. Pritchard, *Markets as Monitors: A Proposal to Replace Class Actions with Exchanges as Securities Fraud Enforcers*, 85 VA. L. REV. 925 (1999) (suggesting that the stock exchanges could be efficient regulators of securities fraud).

527. Stephen J. Choi & A.C. Pritchard, *Behavioral Economics and the SEC*, 56 STAN. L. REV. 1 (2003).

If legal decision theorists do, as Mitchell charge, "proceed[ ] on the basis of inaccurate understandings of judgment and choice,"[528] then their policy prescriptions should be bogus and easily debunked. However, while Mitchell is quick to cherry pick random claims of legal decision theorists that he thinks are overblown, he does very little in either of his articles to persuade readers that any particular position or policy prescription suggested by the legal decision theorists is ill-founded.

V. LEGAL DECISION THEORY: AN ASSESSMENT

This article is already far too long for me to insert a lengthy projection of the future of legal decision theory. Because psychology and economics tend to focus on different aspects of behavior and to use different sources of data,[529] both have a role to play in policy analysis. In 1986, Harrison argued that "the current application of economic analysis to law should be regarded as an interim step toward the integration of law with the behavioral, natural, and social sciences."[530] Introduction of cognitive and behavioral literature from the psychology field is another, and a positive, step toward that integration. Generally, legal decision theorists seek "not to displace law and economics but rather to supplement it so that it serves as a more useful analytic and predictive tool. . . ."[531]

Given the ubiquity of law and economics in legal scholarship, the goal of improving economic analysis is hardly an unimportant one, but it remains relatively modest.[532] Although Mitchell criticizes me for a relatively ambitious statement that I made in one article—a claim that "behavioral research can improve upon standard law and economics analysis almost across the board"[533]—I think that the statement is accurate in the context in which it was made. My argument there, and here, is that whenever it is important to know why people do what they do and how they make decisions, behavioral

528. Mitchell, *Incompetence, supra* note 12, at 73 (quoting Sunstein, *Behavior Analysis, supra* note 10, at 1194)).

529. *See* Mikhail Myagkov & Charles R. Plott, *Exchange Economies and Loss Exposure: Experiments Exploring Prospect Theory and Competitive Equilibria in Market Environments,* 87 AM. ECON. REV. 801, 802 (1997).

530. Jeffrey L. Harrison, *Egoism, Altruism, and Market Illusions: The Limits of Law and Economics,* 33 UCLA L. REV. 1309, 1314 (1986).

531. Lee Anne Fennell, *Death, Taxes, and Cognition,* 81 N.C. L. REV. 567, 571 (2003).

532. *See* Jeffrey J. Rachlinski, *The Uncertain Psychological Case for Paternalism,* 97 NW. U. L. REV. 1165, 1173-74 (2003) (noting that "psychologists studying judgment and choice have been particularly well aware of their competition with economics" and therefore their work has emphasized "errors in judgment more so than the psychologists studying perception or memory").

533. Prentice, *supra* note 36, at 135, *quoted in* Mitchell, *supra* note 12, at 71 n. 7.

analysis can improve upon a simple assumption of *homo economicus*.[534] The premise of a rational actor that is fundamental to much law and economics theory is simply wrong,[535] meaning that this scholarship often generates explanations that are unpersuasive,[536] conclusions that are unverifiable,[537] predictions that are unreliable,[538] and policy prescriptions that are unsatisfying.[539]

Unlike law and economics in its more extreme forms, legal decision theory does not purport to explain why the law is as it is or to provide a universal normative standard.[540] It does not purport to offer

---

534. *See* Fanto, *supra* note 311, at 1341, 1342 (noting that the Chicago Man model does not describe how people actually act and arguing that the more realistic behavioral account of human behavior "is critical for legal studies because laws and legal rules are often designed to affect and modify behavior, and this goal can hardly be achieved without the best available understanding of the behavior itself").

535. *See supra* notes 27-37 and accompanying text; *see also* Ronald J. Allen & Brian Leiter, *Naturalized Epistemology and the Law of Evidence,* 87 VA. L. REV. 1491, 1517 (2001) (noting that most versions of the Chicago Man model are "relatively poor predictors of behavior").

536. *See, e.g.*, KAUSHIK BASU, SEXUAL HARASSMENT IN THE WORKPLACE: AN ECONOMIC ANALYSIS WITH IMPLICATIONS FOR WORKER RIGHTS AND LABOR STANDARDS POLICY (MIT, Dep't of Econ., Working Paper No. 02-11, Feb. 2002) (using more than a little creativity means in order to justify on economic grounds rules against sexual harassment in the workplace), http://papers.ssrn.com/abstract_id=303184.

537. *See, e.g.*, Paul Hirsch et al., *"Dirty Hands" Versus "Clean Models": Is Sociology in Danger of Being Seduced by Economics?,* 16 THEORY & SOC'Y 317, 331-32 (1987) ("On the micro level, economists' assumption of rationality can be restated as psychological hedonism, at which point the proposition becomes irrefutable. If a person chooses a job with lower pay, the economist will add that his or her utility function must include variables besides pay—you just have to include them in formulas to show that utility was maximized."); Wendel, *supra* note 5, at 54 ("[A]s many critics, even those who are generally sympathetic with economic analysis, have observed, claims of empirical verifiability have largely not been borne out." (citations omitted)).

538. *See infra* notes 545-553 and accompanying text.

539. *See* Samuel Bowles, *Endogenous Preferences: The Cultural Consequences of Markets and Other Economic Institutions,* 36 J. ECON. LITERATURE 75, 103-04 (1998) (noting that "experiments in economics, sociology, and psychology have raised serious doubts about the behavioral accuracy of the minimalist conception of homo economicus" and that "economics pays a heavy price for its self-imposed isolation from the other behavioral sciences"); Hirsch et al., *supra* note 537, at 320 ("By precluding attention to non-rational elements of human behavior, economists leave themselves no mechanism for *learning* about the crude and messy empirical world that so defies their models. Economists pay a heavy price for the very simplicity and elegance of their models: empirical ignorance, misunderstanding, and, relatedly, unrealistic and bizarre policy recommendations.").

Two examples of law and economics-generated reform proposals that are questionable include Stephen Choi's suggestion that professionals in the financial markets should be totally deregulated and investors, on the other hand, should be regulated, Choi, *supra* note 526, at 333-34, and proposals that insider trading laws be gutted. *See* FRANK H. EASTERBROOK & DANIEL R. FISCHEL, THE ECONOMIC STRUCTURE OF CORPORATE LAW 253-75 (1991). I have explained by disagreements with Choi's proposal. *See* Prentice, *supra* note 317, at 1399.

540. *See* Nussbaum, *supra* note 15, at 1198 (Law and economics typically "presents itself as explanatory/predictive; but through a certain characteristic use of the concept of rationality, it ends up making normative judgments as well. Thus Richard Posner, for example, both characterizes (most) human behavior as rational in the precise descriptive sense he gives to that

an entirely new paradigm.[541] In this sense, legal decision theory has aspirations much more realistic than Mitchell ascribes to it.

Although in an ideal world the psychological literature would produce a persuasive, all-encompassing theory of human behavior comparable to that which some economists have erroneously claimed for law and economics, this seems unlikely to occur in the near future.[542] Legal decision theory has been termed a form of legal pragmatism because of its atheoretical focus.[543] One reason that legal decision theorists generally do not pretend to have the one right answer to all legal issues is that psychology itself, similar to biology and geology, "does not have large-scale unifying theories of the Einsteinian type. . . . [M]ost psychologists view the possibility of a theory that would unify the entire discipline as highly unlikely."[544]

Milton Friedman famously argued that it is unimportant that economic reasoning is based on assumptions that do not reflect the real world so long as those assumptions produce predictable results.[545] Whereas economics generates predictions from its broad, simplified assumptions, psychology generates predictions from its empirical findings in experiments conducted in the laboratory and in the field. The charge that legal decision theory lacks a unified, overarching theory "is not fatal [because] [n]eeded support for predictions about behavioral responses to legal rules can come indirectly, from

---

term, and then, shifting over to a normative use of the same term, blames certain other agents for not conforming their behavior to those standards."); Wendel, *supra* note 5, at 4 ("Posner and others in the law and economics camp recommend . . . that economic efficiency should be used as the sole extra-legal . . . criterion for justifying legal judgments.").

541. *See* 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.").

542. Thus, Mitchell sketches the future of legal decision theory as I see it when he suggests that "[a]n alternative future for legal decision theory lies in placing a greater emphasis on careful research into discrete problems or in incrementally trying to improve the predictive power of law and economics by identifying those irrational tendencies that seems strongest and most prevalent across similar legal settings," so that "the rational actor assumption can be relaxed or supplemented in specific, manageable ways." Mitchell, *Incompetence*, *supra* note 12, at 87.

543. Korobkin & Ulen, *supra* note 10, at 1057.

544. STANOVICH, *supra* note 64, at 127; *see also* Donald G. MacKay, *The Theoretical Epistemology: A New Perspective on Some Long-Standing Methodological Issues in Psychology*, *in* HANDBOOK FOR DATA ANALYSIS, *supra* note 98, at 229, 229 (noting that the failure of psychology "to develop general and plausible theories [is the discipline's] greatest shortcoming").

545. Milton Friedman, *The Methodology of Positive Economics*, *in* ESSAYS IN POSITIVE ECONOMICS 3, 14-16 (1953). In other words, Friedman is claiming that it is okay to be right for the wrong reasons. Hirsch et al., *supra* note 537, at 324.

defensible inferences from experimental results."[546] Theories that come after observation, as they tend to in psychology, can be much more reliable than theories that economics produces ex ante, for in economics the theorists have a much stronger incentive to interpret the facts to fit the theory, whether they comfortably do so or not.[547]

Therefore, Judge Posner overstates the case with his claim that legal decision theory's lack of an all-encompassing theory means that "it is profoundly unclear what 'behavioral man' would do in any given situation."[548] As noted earlier,[549] economic theory predicts that Chicago Man will approach a rental car counter and actively bargain for his desired level of risk. Legal decision theory predicts that K-T Man will sign whatever contract is put in front of him.[550] Neither view is universally correct, but which more accurately describes the typical transaction?

Chicago Man purchases insurance coverage that best fits his needs. K-T Man tends to purchase whatever is the default coverage even when he has other choices available. Empirical evidence of actual consumer decisions confirms that many people act more like K-T Man than like Chicago Man.[551]

Economist judges predict that an audit failure will not be due to an auditor's reckless auditing, because it would be irrationally injurious to his or her reputation for an auditor to act recklessly.[552] Legal decision theorists predict that a plausible cause of an audit failure is reckless auditing because auditors, like everyone else, are subject to bounded rationality, rational ignorance, the confirmation bias, the self-serving bias, and a host of other cognitive limitations and behavioral predispositions.[553] Ask the Enron shareholders whether it is better to simply accept the economists' absolute assumption, or to examine evidence to determine whether the auditors might have been reckless.

---

546. STEVEN D. WALT, LIQUIDATED DAMAGES AFTER BEHAVIORAL LAW AND ECONOMICS 20 (Univ. of Va. Law & Econ. Research Paper No. 01-18, Dec. 2002), http://papers.ssrn.com/sol3/papers.cfm?abstract_id=353260.

547. *See* Neel P. Parekh, Note, *Theorizing Behavioral Law and Economics: A Defense of Evolutionary Analysis and the Law*, 36 U. MICH. J.L. REFORM 209, 211-12 (2002) ("While searching for information consistent with a theory, researchers often skew the data they obtain.").

548. Posner, *supra* note 11, at 1559.

549. *See supra* note 489 and accompanying text.

550. *See supra* notes 490-503 and accompanying text.

551. *See supra* note 87.

552. *See* DiLeo v. Ernst & Young, 901 F.2d 624, 629 (7th Cir. 1990).

553. *See* Prentice, *supra* note 36, at 152-81.

Furthermore, it appears that Kahneman and Tversky's prospect theory[554] does have the potential to unite a broad number of psychological concepts. Prospect theory assumes that people making decisions are trying to maximize outcomes (not necessarily wealth maximization), but often fail to do so in predictable and systematic ways. In making decisions under uncertainty, for example, people tend to deviate from the Chicago Man model in at least four respects:

> First, people evaluate decision options relative to some reference point, generally the status quo. When choosing between options that appear to be gains relative to that reference point, people tend to make risk-averse choices; when choosing between options tbat appear to be losses, people tend to make risk-seeking choices. . . .
>
> Second, individuals' risk preferences tend to reverse when they are faced with *low-probability* gains and losses. Individuals tend to make risk-seeking choices when selecting between options that appear to be low-probability gains and risk-averse choices when selecting between options that appear to be low-probability losses. . . .
>
> Third, individuals tend to value losses more heavily than gains of the same magnitude. . . .
>
> Fourth, and finally, individuals tend to overvalue certainty.[555]

In a recent paper, Guthrie summarized the legal decision theory scholarship applying prospect theory. He makes a persuasive case that the new literature has added valuable insights into questions as diverse as why litigation is settled, why legal penalties sometimes do not deter negligence, whether consent in medical treatment cases is truly informed, how consumers are vulnerable to being manipulated by the sellers of products, why people frequently do not act as Chicago Man in negotiating contracts, why lawyers often get caught up in the criminal activities of their clients, why people pay taxes even though it is unlikely they will be audited, when corporate managers will be excessively risk seeking, when stockbrokers will be likely to take advantage of their customers, and when companies will attempt to engage in predatory pricing.[556]

---

554. Prospect theory has many facets, but it stresses importantly that expected utility theory does not adequately describe decision making under uncertainty in part because it does not take into account tbe fact that people's decision making is heavily influenced by reference points. For example, people are risk averse for gains and risk seeking for losses of bigh probability, yet risk seeking for gains and risk averse for losses of low probability. *See* Daniel Kahneman & Amos Tversky, *Prospect Theory: An Analysis of Decision Under Risk,* 47 ECONOMETRICA 263 (1979); Tversky & Kahneman, *supra* note 462, at 297.

555. Guthrie, *supra* note 1, at 1118-19 (citations omitted) (emphasis added). Thus, the four key aspects to prospect theory are framing of ordinary gains and losses, framing of low-probability gains and losses, loss aversion, and the certainty effect. *Id.*

556. *Id.* at 1134-36, 1139-55.

Thus, prospect theory has the potential to add valuable insights to legal policy across a wide range of issues.[557] However, even its strongest proponents do not envision for it the type of grand, overarching role that many law and economics scholars once aspired to for their work.[558] This is okay, for as Allen and Rosenberg have recently pointed out, vast areas of the law, including tort theory, do not seem suitable for all-encompassing "top down" theories.[559] Yet, for the last two decades, many economists have been trying to fit messy facts into their simple overarching theory.[560] Their success has been spotty. Furthermore, Gigerenzer has observed that "the comparatively recent arrival of experimental economics and game theory . . . [has] compelled economists to examine the demands their models place on the people whose behavior they are trying to describe."[561] Those economists who have incorporated psychological evidence about how people actually make decisions into their models have produced more satisfying insights than those who have not.[562] For those who resist adding this element of realism to their analysis, such as Judge Posner in his recent examination of the rules of evidence,[563] Allen and Leiter note that "[r]ather plainly, what matters is how people and the system behave in fact, not how they are predicted to behave by the application of formal tools, no matter how

557. *See* ROBYN M. DAWES, RATIONAL CHOICE IN AN UNCERTAIN WORLD 44 (1988) ("Prospect theory is a successful descriptor, however, not just because it incorporates irrationality, but because it predicts the *direction* of irrationality when it occurs."); PLOUS, *supra* note 135, at 95 (observing that prospect theory is the most widely accepted alternative proposed to replace expected utility theory).

558. *See* Guthrie, *supra* note 1, at 1163 ("Prospect theory's central insight, not unlike the central insight of rational choice theory, is a rather blunt tool of analysis. As such, it cannot explain the way all actors make decisions in all contexts. Nonetheless, it represents a valuable refinement to the maximization assumption and should inform law teaching, legal scholarship, and policymaking.").

559. Allen & Rosenberg, *supra* note 32, at 683.

560. *See* Hirsch et al., *supra* note 537, at 331 (noting that for economists, "[t]he question becomes not *whether* the data fit their assumptions, but how they can be *made* to fit the model").

561. GIGERENZER ET AL., SIMPLE HEURISTICS, *supra* note 294, at 347.

562. For example, in the field of criminal law David Dana's recent examination of the "puzzle" of escalating penalties (the fact that deterrent effect does not necessarily escalate along with an increase in penalties) included behavioral analysis to supplement his primarily economic point of view, and his conclusions were more satisfactory than they would have been otherwise. *See* Dana, *supra* note 34, at 733. Dana looked at the influence of the availability heuristic and overoptimism on criminal behavior. *Id.* at 759-63. In another recent article, contractarian Richard Painter clearly recognized that the status quo bias could lead parties to accept default rules that are not optimal, and his proposals for the proper rules governing the lawyer-client relationship became more realistic than they would have been without behavioral input. *See* Richard W. Painter, *Rules Lawyers Play By*, 76 N.Y.U. L. REV. 665, 687 (2001).

563. Posner, *Evidence*, *supra* note 34.

elegant, for which there is substantial discomfiting, if not disconfirming, data."[564]

Policy prescriptions based on complicated but very real facts have more promise than those based on elegant but very wrong theory. Imitating economists, some legal decision theorists will, no doubt, attempt to find a simple overarching theory to fit psychology's messy facts.[565] I predict only partial success. The world is just too complicated. Mitchell says that "while empirical research can provide better answers than we currently have, it will not provide incontestable or simple answers about legal rationality for prescriptive use."[566] I agree. The policy prescriptions offered by legal decision theorists will never be incontestable.[567] They will seldom be simple. However, for K-T Man to have more descriptive, explanatory, predictive, and prescriptive power than Chicago Man, people need only be systematically (not universally and uniformly) subject to the various heuristics and biases discussed in the literature. And they are.

Despite its limitations, legal decision theory carries the promise of significant impact. In light of Cass Sunstein's writings,[568] future discussions of risk regulation cannot comfortably proceed without taking into account how people actually feel about risk.[569] In

---

564. Allen & Leiter, *supra* note 535, at 1518.

565. For example, Owen Jones's time-shifted rationality, which seeks to explain modern behavioral scholarship through principles of evolutionary psychology, has broad theoretical implications, but seems unlikely in the near term to uncover principles that are sufficiently specific to generate much predictive power across a broad range of issues. *See* Owen D. Jones, *Time-Shifted Rationality and the Law of Law's Leverage: Behavioral Economics Meets Behavioral Biology*, 95 NW. U. L. REV. 1141 (2001); *see also* Parekh, *supra* note 547, at 221-27 (offering other evolution-based theories to bolster behavioral law and economics).

566. Mitchell, *Incompetence*, *supra* note 12, at 14.

567. *See, e.g.*, Hanson & Kysar, *supra* note 404, at 1560 (using behavioral insights to argue that because consumers are prone to misperceive risks and manufacturers will inevitably take advantage of them, the legal system should impose enterprise liability on manufacturers). *But see* 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 (2000) (disagreeing with Hanson and Kysar regarding their conclusion as to the appropriate policy response).

568. *See, e.g.*, CASS R. SUNSTEIN, RISK AND REASON: SAFETY, LAW AND THE ENVIRONMENT (2002).

569. As noted, this article is already too long for me to undertake an extensive discussion of my vision of the future of legal decision theory, but a common theme will likely be how lawmakers, educated as to the heuristics and biases literature, can improve upon decision making, often overcoming the errors of the uneducated, lay public. Sunstein's work in risk regulation may well be a prototype. He has explained in detail how the heuristics and biases lead the citizenry to overestimate many environmental and similar risks and underestimate others. For example, the availability heuristic induces citizens to overestimate vivid, headline-grabbing risks and to underestimate more serious but less publicized risks. Unless we are careful, this leads to anecdote-driven "pollutant of the month" legislation. *See* Jolls et al., *supra* note 310, at 1518; Timur Kuran & Cass R. Sunstein, *Availability Cascades and Risk Regulation*,

light of the work of Russell Korobkin, Jeffrey Rachlinski, and Chris Guthrie, future discussions of contract law cannot cavalierly omit evidence regarding how real people actually think when signing contracts.[570] In light of Donald Langevoort's substantial scholarship,[571] future policy prescriptions regarding corporate governance or stockbroker regulation cannot safely ignore behavioral and cognitive evidence regarding how people actually respond to incentives. In light of my own work[572] and, much more importantly, the Enron debacle,[573] I suspect fewer courts in cases of audit failure will admit in writing to a presumption that the auditors in an inherently conflicted situation must have acted rationally to preserve their long-term reputational capital.[574] After this article was accepted for publication, Mitchell himself made public an article suggesting a roadmap for properly bringing a behavioral research agenda to the field of evidence law.[575] For my money, Mitchell's agenda holds more promise for bringing insight to the field of evidence law than we are likely to gain from even the best economic analyses,[576] but we are probably better off gaining the benefit of both. When looking for dimes, it is better to have an economic street lamp *and* a behavioral street lamp shining in the area, although if I must settle for one I will take the latter.

That does not mean that behavioral analysis will always get it right, or that legal decision theorists will always agree on the proper

---

51 STAN. L. REV. 683, 698 (1999). Legislators and regulators schooled in psychology literature have the potential to provide more rational risk regulation than public outcry would demand.

570. *See supra* note 10.

571. *See supra* note 10.

572. *See* Prentice, *supra* note 36; Prentice, *supra* note 211.

573. The conflicts between audit work and consulting work have become increasingly apparent as the Enron saga unfolds. At this writing, evidence of conflicts between audit work and tax work is coming to the fore. *See, e.g.*, Ken Brown & John D. McKinnon, *IRS Later Opposed Tax Strategies Sold by Auditor*, WALL ST. J., Feb. 6, 2003, at A3 (noting that controversy over aggressive tax shelters "is certain to revive the controversy over whether accounting firms should be permitted to provide both auditing and consulting services to big customers"), 2003 WL-WSJ 3958684; Jeremy Kahn, *Do Accountants Have a Future?*, FORTUNE, Mar. 3, 2003, at 115 (noting that the accounting industry's potential liability for bogus tax shelters "is likely to run into the billions"), 2003 WL 8253597.

574. Indeed, the last case to cite *DiLeo v. Ernst & Young*, 901 F.2d 624, 629 (7th Cir. 1990), for the proposition to which I objected—that courts should not assume auditors would act irrationally by doing anything to endanger their reputation for honesty—was *Reiger v. Price Waterhouse Coopers*, 117 F. Supp. 2d 1003 (S.D. Cal. 2000), back in October of 2000. Since then, courts are more likely to say that when auditors violate GAAP, they forfeit the presumption that it is "irrational" for them to risk their reputations. *See In re* SCB Computer Tech., Inc., 149 F. Supp. 2d 334 (W.D. Tenn. 2001).

575. GREGORY MITCHELL, MAPPING EVIDENCE LAW (Fla. State Univ. Coll. of Law, Pub. Law & Legal Theory, Working Paper No. 75, Mar. 2003), http://papers.ssrn.com/sol3/papers.cfm?abstract_id=386860.

576. *See* Posner, *Evidence*, *supra* note 34.

resolution of an issue any more than economists always so agree.[577] Nor, of course, does it ensure that, even if the legal decision theorists have it right, courts and legislatures will readily embrace this new scholarship.[578]

VI. CONCLUSION

Clifford Geertz observed that

> [C]ertain ideas burst upon the intellectual landscape with tremendous force. They resolve so many fundamental problems at once that they seem also to promise that they will resolve all fundamental problems, clarify all obscure issues. . . .
>
> After we bave become familiar with the new idea, however, after it has become part of our general stock of theoretical concepts, our expectations are brought more into balance with its actual uses, and its excessive popularity is ended. A few zealots persist in the old key-to-the-universe view of it; but less driven thinkers settle down after a while to the problems the idea has really generated. They try to apply it and extend it where it applies and where it is capable of extension; and they desist where it does not apply or cannot be extended . . . . [I]t still explains something; and our attention shifts to isolating just what that something is. . . .[579]

This statement well fits the trajectory of the law and economics movement, and may also describe the eventual track of legal decision theory. Ardent supporters of legal decision theory have no doubt painted an enthusiastic picture of the potential that it has for

---

577. *See* John C. Moorehouse et al., *Law & Economics and Tort Law: A Survey of Scholarly Opinion,* 62 ALB. L. REV. 667, 670 (1998) (reporting that a survey of law and economics scholars found no consensus as to the efficiency of common law rules in general, a bedrock principle for many leaders in the field).

578. I do not expect courts and legislatures to jump eagerly on board the legal decision theory train, no matter how valuable its potential contributions appear to be. Blumenthal notes that "it takes prodigious effort and confidence for a judge to overturn precedent based on empirical findings." Blumenthal, *supra* note 1, at 72. The good news is that he has also noted that "[c]ourts' receptivity to social science knowledge may be increasing . . . ." Jeremy A. Blumenthal, *The Reasonable Woman Standard: A Meta-analytic Review of Gender Differences in Perceptions of Sexual Harassment,* 22 LAW & HUM. BEHAV. 33, 51 n.6 (1998).

However, legal decision scholars should not be overly optimistic. Given the prominent judicial position of law and economics scholars, Circuit Judges Easterbrook, Posner, and (formerly) Bork, and the number of judges who were processed by Henry Manne's economics boot camps, *see* Henry N. Butler, *The Manne Programs in Economics for Federal Judges,* 50 CASE W. RES. L. REV. 351, 352 (1999) ("By 1990, approximately forty percent of the sitting federal judges had completed Manne's flagship program—*the Economics Institute for Federal Judges.*"), it is surprising that law and economics has found so little purchase in judicial opinions. It is also chastening for advocates of legal decision theory, who are unlikely to have three leaders in their field appointed to such prominent judicial positions and are similarly unlikely to find corporations and conservative foundations that will fund psychology camps for judges. *See* Anita Bernstein, *An Old Jurisprudence: Respect in Retrospect,* 83 CORNELL L. REV. 1231, 1239 (1998) (noting the failure of law and economics to have much impact on the law in spite of Henry Manne's efforts, and arguing that "[l]ife-tenured, politically vetted federal judges have little incentive to work at revising what they know" (citation omitted)).

579. CLIFFORD GEERTZ, THE INTERPRETATION OF CULTURES 3-4 (1973).

informing legal analysis, and some may have fallen into the overadvocacy trap.[580] Articles such as those Gregory Mitchell has recently published[581] are valuable to bring perspective to the area. Legal decision theory's attempts to inject into legal scholarship the concepts of behavioral psychology, cognitive science, and related fields must be able to survive challenges such as Mitchell's in order to prosper in the long run. His articles, then, are a welcome addition to the debate about the usefulness of this new research. They deserve and demand a response that I have attempted to provide.

Mitchell claims that psychological research "does not prove that experimental subjects—much less real-world legal decision makers—systematically violate norms of rationality when forming judgments and making decisions."[582] Despite the limitations of social science research that Mitchell explores in detail, the debate over whether the economists' Chicago Man or the psychologists' K-T Man better describes reality is over; the psychologists won.[583] As Nobel Prize-winning economist Daniel McFadden recently said:

> When one looks at the whole body of experimental studies of cognition and choice over the past twenty-five years, what stands out is that humans fail to retrieve and process information consistently, and this generates a variety of cognitive anomalies, including behavior that makes consumers vulnerable to exploitation in markets. Available, salient information looms too large, and beliefs are distorted because attention to new information is selective. These failures may be fundamental, the result of the way human memory is wired. I conclude that perception-rationality fails, and that *the failures are systematic, persistent, pervasive, and large in magnitude.*[584]

---

580. Lee Sechrest and Richard Bootzin argue that

> There are many pitfalls for psychologists who attempt to influence the content of public policy. One of the major problems is what Campbell (1969) called the *overadvocacy trap*. The task of getting new policies adopted is so difficult that the psychologist, as a policy advisor, is almost certain to make exaggerated claims about the degree and likelihood of a policy's effectiveness. As Campbell (1972) stated, social scientists who attempt to help develop policy too often "speak with a certainty unjustified by the validity of their science."

Sechrest & Bootzin, *supra* note 238, at 388 (citing Donald T. Campbell, *Reforms as Experiments*, 24 AM. PSYCHOLOGIST 409 (1969) and Donald T. Campbell, *Comments on the Comment by Shaver and Staines*, 27 AM. PSYCHOLOGIST 164 (1972)).

Mitchell says that legal decision theorists "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." Mitchell, *Pessimism*, *supra* note 12, at 1911. I suspect he is right; it is certainly consistent with the behavioral research that behaviorists, along with everyone else, would have a tendency to do this.

581. *See* Mitchell, *Pessimism*, *supra* note 12; Mitchell, *Incompetence*, *supra* note 12.

582. Mitchell, *Pessimism*, *supra* note 12, at 1912.

583. Markman & Medin, *supra* note 195, at 414; *see also* Shafir, *supra* note 295, at 277 (noting that the evidence from the K-T tradition "point[s] to the fact that the normative theory is irreconcilable with the ways in which people make choices").

584. McFadden, *supra* note 6, at 96 (emphasis added).

Any sting caused by Mitchell's attacks on Daniel Kahneman's heuristics and biases work will perhaps be eased somewhat by the Nobel Prize in economics that Kahneman was awarded in the fall of 2002.[585]

Mitchell's second main point, that legal decision theory "cannot lay claim to empirical validity superior to that of the perfect rationality assumption,"[586] is true only if one accepts Mitchell's inaccurate characterization of legal decision theory's position as one of "equal incompetence." Rather than believing that all people are universally and uniformly irrational in all settings, legal decision theorists simply hold to the view, expressed by McFadden, that various heuristics and biases limit human decision making in a systematic way. If one looks instead at how legal decision theorists actually apply the empirical evidence from psychology and related fields, it becomes clear that useful policy insights can be gained to a wide variety of legal issues even if decision makers are not universally and uniformly irrational in all settings.

Attempts to paint the heuristics and biases literature as the product of parlor tricks arising only from psychology laboratory experiments involving college sophomores will become less and less persuasive as studies of the brain demonstrate physically these phenomena. Neuroscientists using CAT scans, PET scans, MRIs, MRAs, and other techniques can now view brain functions without invading the skull.[587] Using such techniques, scientists have seen displayed in brain wave recordings direct evidence of loss aversion[588]

---

585. *See* Jon E. Hilsenrath, *Nobel Winners for Economics Arc New Breed,* WALL ST. J., Oct. 10, 2002, at B1 (noting that Kahneman was awarded the Prize "for research showing how quirks in human behavior, such as a tendency to avoid risk or to be over-confident, lead people to behave in ways economists would consider irrational or that don't always bring positive outcomes"), 2002 WL-WSJ 3408400.

Mitchell's reservations to the contrary, the heuristics and biases literature is receiving more and more acceptance, and even acclaim, in the economics field. The American Economic Association recently awarded its prestigious John Bates Clark medal for leading economists under 40 to behavioralist Matthew Rabin, and the MacArthur Foundation recently gave a genius award to MIT behavioral economist Sendhil Mullainathan. *Id.* at B3.

Coincidentally, Kahneman shared that Nobel Prize with Vernon Smith who won for his work in experimental economics, work which has repeatedly demonstrated that markets do not operate as traditional economics assumes. *See* Chris Giles, *Economists Get the Idea,* FIN. TIMES, Oct. 15, 2002, at 13 (noting that Professor Smith's work showed that in both simple markets and complicated markets such as auctions, people do not behave in accordance with rational man theory), 2002 WL 101374225.

586. Mitchell, *Incompetence, supra* note 12, at 72.

587. *See* RICHARD RESTAK, THE NEW BRAIN 3 (2003).

588. *See* William J. Gehring & Adrian R. Willoughby, *The Medial Frontal Cortex and the Rapid Processing of Monetary Gains and Losses,* 295 SCI. 2279, 2279-81 (2002).

and the gambler's fallacy.[589] They have direct evidence that every time people make decisions that affect their own lives, the portions of the brain that produce emotions are involved, even though the decision would seem to call for straight forward rationality.[590] Other brain work indicates that "the brain has special sectors for emotions, and that some types of emotions, including some fear-type reactions, can be triggered before the more cognitive sectors become involved at all."[591]

Still other experiments using an FMRI (Functional Magnetic Resonance Imaging) provided direct evidence from brain functions[592] to support the existence of the omission bias that Koehler and I discussed in a recent article.[593] And neuroimaging studies support the existence of time-delay traps[594] that I have discussed in connection with the causes of crime.[595] Still other experiments eliminate doubts some have had that when experimental subjects trust more than Chicago Man it is because they do not understand the game correctly. FMRI scans of trusters find activity in parts of the brain called Brodman's areas 8 and 10 that have been associated with mental thinking about the motivations of others and delayed gratification.[596] Nobel Prize winning economist Vernon Smith and his coauthors recently reported these and similar results, calling for a new research push in this field they term neuroeconomics so that we may produce "better models of human behavior and consequently a better understanding of legal problems."[597]

Mitchell does a valuable service by reminding us that legal decision theory holds no magic key, and that easy answers will remain elusive.[598] Just as in the physical sciences where what we "know"

---

589. *Id.*

590. *See* RESTAK, *supra* note 587, at 111-12 (citing the work of Professor Dean Shibata).

591. SUNSTEIN, *supra* note 568, at 44-45 (citing JOSEPH LEDOUX, THE EMOTIONAL BRAIN: THE MYSTERIOUS UNDERPINNINGS OF EMOTIONAL LIFE (1996)).

592. *See* RESTAK, *supra* note 587, at 113-14 (citing the work of Professor Joshua Greene).

593. *See* Prentice & Koehler, *supra* note 405.

594. *See* RESTAK, *supra* note 587, at 117-18 (citing the work of psychologist Laura L. People).

595. *See* Prentice, *supra* note 36, at 177-78.

596. Tim Harford, *In Search of the Inside Story of Economics*, FIN. TIMES, Sept. 30, 2003, at 12 (noting that such brain activity is not present in nontrusters or in those who know that they are playing a game with a machine rather than with another human), 2003 WL 63746636.

597. TERRENCE CHORVAT ET AL., LAW AND NEUROECONOMICS 45 (George Mason Univ. Sch. of Law, Law & Econ. Research Paper No. 04-07, 2004), http//ssrn.com/abstract= 501063. *See generally* PAUL GLIMCHER, DECISIONS, UNDERCERTAINTY, AND THE BRAIN: THE SCIENCE OF NEUROECONOMICS (2003).

598. Ultimately, behavioral theory will likely be only a part of the ultimate picture, explaining how man makes decisions. That picture will include institutional economics, positive economics, experimental economics, organizational theory, cognitive science, psychoanalytic theory, perhaps memetics. *See generally* ROBERT AUNGER, THE ELECTRIC MEME: A NEW THEORY

today about the origins of the universe[599] and the behavior of black holes[600] is different that what we "knew" two decades ago, our knowledge of the cognitive functions of humans and their decision making behavior in society will continue to evolve. Psychology's understanding of K-T Man will be different ten years from now than it is today, but legal decision theorists will follow the research where it leads. We do not now have final answers to the most difficult questions regarding human judgment and decision making. Likely we never will. Nonetheless, "at any one time we have available a body of research bearing on the nature and determinants of behavior, and we ought to make of it the best we can. Policy may be better formed out of the best evidence available than out of the prejudices of individual decision makers or out of thin air."[601] Or, I suggest, out of the false premise that man is a rational maximizer of his utilities.

OF HOW WE THINK (2002) (describing the new field of memetics—a complex theory about how all information is transmitted). Ultimate answers will be complex and, as Mitchell valuably points out, context specific.

599. *See, e.g.*, Faye Flam, *One Big Universe Deserves Another; 2 Experts Say Replacement Due in a Trillion Years*, PHILA. INQUIRER, Apr. 28, 2002, at A3 (describing new theory regarding the origin of the universe proposed by physicists at Princeton and Cambridge).

600. *See* Michael Alicea, *Scientists Pulled in By Mystery of Black Hole*, PALM BEACH POST, June 9, 2002, at 7D (describing physicists' new theories regarding black holes); Ann Schrader, *Scientists Reverse Black Hole Theory: CU 'Hunter' Helps Find Energy Leaving Object*, DENVER POST, Oct. 23, 2001, at B2 (same).

601. Sechrest & Bootzin, *supra* note 238, at 381.

***

Tendencies Versus Boundaries: Levels of Generality in Behavioral Law and Economics

Gregory Mitchell 56 Vand. L. Rev. 1781 (2003)

In this reply to Professor Prentice's article, Professor Mitchell offers some additional thoughts in favor of a modest approach to revising the law's assumption of rationality, as compared to the bolder approach argued for by Professor Prentice. After discussing how much of the evidence on human rationality can be used both to attack and defend the rational actor assumption due to ambiguities in this evidence, Professor Mitchell turns to the larger question of whether legal decision theorists describe behavior at too general a level to be useful in the formulation of legal policy. Professor Mitchell argues that legal decision theorists have placed too great an emphasis on finding and describing behavioral tendencies toward irrationality, without due regard for the boundary conditions on these supposed tendencies. As a result, much of the interesting and important information about the constraints on rational versus irrational behavior is consigned to ceteris paribus clauses and treated as "noise" that should be controlled and ignored rather than elucidated and understood.

***