

Mixed Signals: Rational-Choice Theories of Social Norms and the Pragmatics of Explanation

W. BRADLEY WENDEL*

TABLE OF CONTENTS

I.	INTRODUCTION	2
II.	THE "PROBLEM" OF SOCIAL NORMS	8
III.	"TO SEEM RATHER THAN TO BE": RATIONAL-CHOICE MODELS OF SOCIAL NORMS	12
	A. <i>Eclecticism—Elster</i>	12
	B. <i>Sanctioning—McAdams</i>	14
	C. <i>Signaling—Posner</i>	19
IV.	"THE CROOKED TIMBER OF HUMANITY": INTERNAL CRITIQUES	26
	A. <i>The Normativity of Norms</i>	26
	B. <i>Transaction Costs—Observability, Diversity, and Ambiguity</i>	31
	C. <i>Psychological Considerations— Intentionality and Disposition</i>	36
V.	"ADVANCED CAT MUSIC": EXTERNAL CRITIQUES	39
	A. <i>What Is Involved In Providing an Explanation?</i>	41
	B. <i>The Appeal of Economic Models</i>	49
	1. The Ideal of Freedom and the Rejection of Authority	49
	2. Science Envy and Value Neutrality	51
VI.	"SHADOWS OR THE FURNITURE ITSELF": ARBITRATING THE DISPUTE	56

[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 have become familiar with the new idea, however, after it has become part of our general stock of theoretical concepts, our expectations

* Assistant Professor, Washington and Lee University School of Law. The generous support of the Francis Lewis Law Center, Washington and Lee University, and Dean David Partlett, is gratefully acknowledged. A preliminary version of this paper was presented as a seminar in the philosophy department at Washington and Lee University; I am grateful to the faculty and student participants for their spirited questioning. I also appreciate the suggestions provided by the participants in a faculty workshop at Washington and Lee Law School. A subsequent version was presented at the international conference of the Law and Society Association, "Law in Action," in Budapest, Hungary, July 4-7, 2001. Special thanks to Greg Cooper, Quince Hopkins, Peter Huang, Lyman Johnson, Ron Krotoszynski, James Mahon, David Millon, Lad Sessions, and Jack Wilson for their insightful comments.

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 is¹

I. INTRODUCTION

The question of how societies secure cooperation and order in the absence of state enforced sanctions has long vexed law and economics scholars. Recently the concept of social norms—informally enforced rules of behavior—has occupied the attention of a large number of these theorists, who are concerned with understanding why economically rational actors would bother to follow rules whose costs seem to outweigh their benefits. Because of the prestige (or at least trendiness) of law and economics, it seems that now everyone in the legal academy is talking about social norms. Four symposia published in major law reviews have been devoted specifically to the economic analysis of social norms or the application of this analysis to particular settings,² to say nothing of the proliferation of books and independent articles on the same subject, some of which have already become quite influential.³ This burgeoning scholarship is closely related to a wider concern in law and economics, namely the attempt to develop a “behavioral” law and economics that integrates into economic models some of the findings from empirical social science

1. CLIFFORD GEERTZ, *THE INTERPRETATION OF CULTURES* 3-4 (1973).

2. Symposium, *Law, Economics, and Norms*, 144 U. PA. L. REV. 1643 (1996); Symposium, *The Legal Construction of Norms*, 86 VA. L. REV. 1577 (2000); Symposium, *Norms and Corporate Law*, 149 U. PA. L. REV. 1607 (2001); Symposium, *Social Norms, Social Meaning, and the Economic Analysis of Law*, 27 J. LEGAL STUD. 537 (1998). In addition, an issue of the *Wisconsin Law Review* dedicated to achieving some kind of rapprochement between the law and economics and law and society movements contained many articles alluding to the economic analysis of social norms. See Symposium, *Law and Society & Law and Economics: Common Ground, Irreconcilable Differences, New Directions*, 1997 WIS. L. REV. 375.

3. See DENNIS CHONG, *RATIONAL LIVES: NORMS AND VALUES IN POLITICS AND SOCIETY* (2000); ROBERT C. ELLICKSON, *ORDER WITHOUT LAW: HOW NEIGHBORS SETTLE DISPUTES* (1991) [hereinafter ELLICKSON, *ORDER WITHOUT LAW*]; RUSSELL HARDIN, *ONE FOR ALL: THE LOGIC OF GROUP CONFLICT* (1995); Lisa Bernstein, *Opting Out of the Legal System: Extralegal Contractual Relations in the Diamond Industry*, 21 J. LEGAL STUD. 115 (1992); David Charny, *Nonlegal Sanctions in Commercial Relationships*, 104 HARV. L. REV. 373 (1990); Robert D. Cooter, *Models of Morality in Law and Economics: Self-Control and Self-Improvement for the “Bad Man” of Holmes*, 78 B.U. L. REV. 903 (1998); Robert C. Ellickson, *Controlling Chronic Misconduct in City Spaces: Of Panhandlers, Skid Rows, and Public-Space Zoning*, 105 YALE L.J. 1165 (1996); Lawrence Lessig, *The Regulation of Social Meaning*, 62 U. CHI. L. REV. 943 (1995); Richard H. McAdams, *The Origin, Development, and Regulation of Norms*, 96 MICH. L. REV. 338 (1997) [hereinafter McAdams, *Origin*]; Cass R. Sunstein, *Social Norms and Social Roles*, 96 COLUM. L. REV. 903 (1996).

research.⁴ Behavioral analysis attempts to account for standard objections often raised to law and economics, such as the imperfect rationality of fallible human beings and the importance of considerations that cannot be reduced to self-interest—for example loyalty, solidarity, community, altruism, and other moral values. Although behavioral law and economics and the economic analysis of social norms address the problem from different directions, they are both concerned with the same jurisprudential issue, namely how a theory of human behavior ought to be developed that is relevant to the design of legal institutions.

The attempt by law and economics to account for social norms has sparked yet another round of an ongoing battle, between that group of theorists and critics from various other disciplines. Social scientists, members of the loosely defined law and society movement, legal theorists (generally identifying themselves as progressive or critical), moral and political philosophers, and even some scholars generally sympathetic to rational-choice approaches have long objected to the reductionist tendencies of law and economics.⁵ The economic mode of analysis has an almost pathological aversion to explanations that appeal to values, commitments, loyalties, relationships, or emotions. Indeed, one of its principal exponents, Richard Posner, recently excoriated moral philosophers for suggesting that moral values are relevant

4. See, e.g., BEHAVIORAL LAW AND ECONOMICS (Cass R. Sunstein ed., 2000); DONALD P. GREEN & IAN SHAPIRO, PATHOLOGIES OF RATIONAL CHOICE THEORY (1994); RICHARD H. THALER, THE WINNER'S CURSE: PARADOXES AND ANOMALIES OF ECONOMIC LIFE (1992); Russell B. Korobkin & Thomas S. Ulen, *Law and Behavioral Science: Removing the Rationality Assumption from Law and Economics*, 88 CAL. L. REV. 1051 (2000); Symposium, *The Legal Implications of Psychology: Human Behavior, Behavioral Economics, and the Law*, 51 VAND. L. REV. 1495 (1998); Thomas S. Ulen, *Rational Choice and the Economic Analysis of Law*, 19 LAW & SOC. INQUIRY 487 (1994).

5. For important contributions to the extensive critical literature, which varies in the degree of its sympathy with rational-choice theory, see, for example, ELIZABETH ANDERSON, VALUE IN ETHICS AND ECONOMICS (1993); JULES L. COLEMAN, MARKETS, MORALS AND THE LAW (1988); RONALD DWORKIN, *Is Wealth a Value?*, in A MATTER OF PRINCIPLE 237 (1985); MARK KELMAN, A GUIDE TO CRITICAL LEGAL STUDIES 110, 114-50 (1987); C. Edwin Baker, *The Ideology of the Economic Analysis of Law*, 5 PHIL. & PUB. AFF. (1975); Jane B. Baron & Jeffrey L. Dunoff, *Against Market Rationality: Moral Critiques of Economic Analysis in Legal Theory*, 17 CARDOZO L. REV. 431 (1996); Jules L. Coleman, *The Economic Analysis of Law*, in NOMOS XXIV: ETHICS, ECONOMICS, AND THE LAW 83 (J. Roland Pennock & John W. Chapman eds., 1982); Ronald Dworkin, *Why Efficiency? A Response to Professors Calabresi and Posner*, 8 HOFSTRA L. REV. 563 (1980); Jean Hampton, *The Failure of Expected-Utility Theory as a Theory of Reason*, 10 ECON. & PHIL. 195 (1994); Jean Hampton, *Rational Choice and the Law*, 15 HARV. J.L. & PUB. POL'Y 649 (1992); Jeffrey L. Harrison, *Egoism, Altruism, and Market Illusions: The Limits of Law and Economics*, 33 UCLA L. REV. 1309 (1986); Martha C. Nussbaum, *Flawed Foundations: The Philosophical Critique of (a Particular Type of) Economics*, 64 U. CHI. L. REV. 1197 (1997); Margaret Jane Radin, *Market-Inalienability*, 100 HARV. L. REV. 1849 (1987); Jeanne L. Schroeder, *The End of the Market: A Psychoanalysis of Law and Economics*, 112 HARV. L. REV. 483 (1998); Amartya K. Sen, *Rational Fools: A Critique of the Behavioral Foundations of Economic Theory*, 6 PHIL. & PUB. AFF. 317 (1977). For a summary of these competing approaches to the study of law, see Edward L. Rubin, *Law and the Methodology of Law*, 1997 WIS. L. REV. 521.

to the way in which judges ought to decide cases.⁶ Posner and others in the law and economics camp recommend instead that economic efficiency should be used as the sole extralegal (that is, not internal to legal texts like cases and statutes) criterion for justifying legal judgments. Critics of law and economics argue that if legal decisionmakers, such as judges and legislators, look outside the domain of legal texts and conventions, they should rely on a theory of human behavior that is better than rational choice theory.

Admittedly speaking as a critic of law and economics, I believe that seemingly endless battle between its proponents and critics conceals a trajectory traced by rational choice theory in the legal academy. The arc of law and economics perfectly fits the description offered in the passage from Clifford Geertz that opened this paper—it burst upon the scene like a bombshell; its adherents claimed it explains everything, from nuisance remedies (which it probably does),⁷ to sexual idiosyncracies,⁸ to racial discrimination,⁹ to holiday customs,¹⁰ to the whole universe of social norms; critics began to question these extravagant claims; and finally scholars not committed to one discipline or the other have tried to map the boundaries of the enterprise, to see where economic analysis reveals productive insights and where its methods of analysis are misplaced.

What no one has really made explicit in the law-and-economics-versus-everyone-else debate are the terms on which the dispute is supposed to be resolved. There are abundant implicit arguments about what counts as a good explanation, but they are not often made with sufficient clarity.¹¹ For example, consider an argument about why people display gratitude, through practices of writing thank-you notes, giving reciprocal gifts, and so on. A philosopher has advanced this critique of social science explanations of expressions of gratitude (which might be offered by rational-choice theorists, too):

An amateur sociologist would have little trouble pointing out the ways in which such displays reinforce dispositions to giftgiving and enlarge the

6. See RICHARD A. POSNER, *THE PROBLEMATICS OF MORAL AND LEGAL THEORY* 107-44 (1999).

7. See Guido Calabresi & A. Douglas Melamed, *Property Rules, Liability Rules, and Inalienability: One View of the Cathedral*, 85 HARV. L. REV. 1089, 1126 (1972).

8. See RICHARD A. POSNER, *SEX AND REASON* 173-82 (1992).

9. See, e.g., Richard H. McAdams, *Cooperation and Conflict: The Economics of Group Status Production and Race Discrimination*, 108 HARV. L. REV. 1003, 1033-82 (1995) [hereinafter McAdams, *Discrimination*]; Jody D. Armour, *Race Ipsa Loquitur: Of Reasonable Racists, Intelligent Bayesians, and Involuntary Negrophobes*, 46 STAN. L. REV. 781, 790-98 (1994).

10. See John L. Solow, *Is It Really the Thought That Counts? Toward a Rational Theory of Christmas*, 5 RATIONALITY & SOC'Y 506 (1993).

11. A noteworthy exception is Robert E. Scott, *The Limits of Behavioral Theories of Law and Social Norms*, 86 VA. L. REV. 1603 (2000). Richard McAdams has also suggested that the rational-choice theories should be accepted or rejected based on whether they reflect "sound principles of causal explanation or excessive reductionist zeal." Richard H. McAdams, *Signaling Discount Rates: Law, Norms, and Economic Methodology*, 110 YALE L.J. 625, 680 (2001) (book review) [hereinafter McAdams, *Signaling*]. This Article is largely an attempt to flesh out exactly this kind of argument.

degree of concession and concern people show one another. This, however, seems to me an unsatisfactory account of why reflective people seek to maintain the conventions of gratitude. Quite without regard to any further consequences, we care how people feel toward us and how they regard us Gratitude plays a role in the interrelationships precisely because it involves the demonstration of our feelings towards another. Thus it has value quite without regard . . . to any further actions it tends to produce.¹²

How are we to sort out these competing explanations? When it comes to social norms like the practice of showing gratitude, rational-choice theorists face the challenge of offering an explanation that provides an adequate account of these norms, while remaining faithful to the methodological commitments of law and economics. Expressions of gratitude are no doubt functional, in the sense that they contribute to social cohesion. But functional explanations slight the first-person perspective—that is, the reason that might be given by a person actually embedded in the social practice of expressing gratitude who is concerned with expressing her authentic feelings. Thus, one might require that an explanation of gratitude make reference to the kinds of reasons to which persons actually refer in deliberation and reflection. At the same time, however, rational-choice scholars are undoubtedly right that some degree of parsimony is necessary to prevent a theoretical model of social interaction from turning into a morass of ad hoc explanatory postulates and to enable empirical testing and verification of the model.

This Article is an attempt to evaluate the methodological claims of theorists of social norms. For this reason, its aim is primarily metatheoretical. It uses the tools of conceptual analysis—that is, moral and political philosophy and the philosophy of the natural and social sciences¹³—to arbitrate between a long-running dispute between two warring camps, rational choice scholars and their critics from the law and society movement and other disciplines. Both groups of legal scholars seek to understand and explain the interaction between formal, state-enforced legal rules and norms of cooperation that have their root in informal, everyday social interactions. Where the law and economics analysis has emphasized methodological tidiness and a minimum of initial assumptions about rationality, its critics have been willing, on occasion, to offer explanations that have a tendency to come off as question begging.¹⁴ For

12. Fred R. Berger, *Gratitude*, 85 *ETHICS* 298, 308 (1975).

13. See generally SCOTT GORDON, *THE HISTORY AND PHILOSOPHY OF SOCIAL SCIENCE* (1991); CARL G. HEMPEL, *ASPECTS OF SCIENTIFIC EXPLANATION* (1965); *READINGS IN THE PHILOSOPHY OF SOCIAL SCIENCE* (Michael Martin & Lee C. McIntyre eds., 1994) (reprinting many significant works that first appeared elsewhere); ALEXANDER ROSENBERG, *PHILOSOPHY OF SOCIAL SCIENCE* (2d ed. 1995); WESLEY C. SALMON, *CAUSALITY AND EXPLANATION* (1998) [hereinafter SALMON, *CAUSALITY*]; Wesley C. Salmon, *Four Decades of Scientific Explanation*, in *SCIENTIFIC EXPLANATION 3* (Phillip Kitcher & Wesley C. Salmon eds., 1989); GEÓRG HENRIK VON WRIGHT, *EXPLANATION AND UNDERSTANDING* (1971).

14. Robert Ellickson aptly summarizes the dispute:

To exaggerate only a little, the law-and-economics scholars believe that the law-and-society group is deficient in both sophistication and rigor, and the law-and-society scholars believe that the law-and-economics theorists are not only out of

example, Jon Elster, an economically-minded political philosopher, heaps scorn on those theories that adduce concepts like a “norm of cooperation” to explain cooperation, comparing them with the explanation of opium’s soporific properties in terms of its “dormitive principle.”¹⁵ Or, as Robert Ellickson, a rational-choice scholar, observes, functionalist sociologists assume that an observed norm is functional, and then spend a great deal of energy inventing explanations for it.¹⁶

At the same time, however, parsimony of assumptions and theoretical purity are no substitute for explanatory efficacy. As the British philosopher Bernard Williams has asked, in response to the arid, formal style of academic theorizing that has dominated the twentieth century, “what authority are theoretical tidiness or simplicity supposed to have against the force of concerns which one actually finds important?”¹⁷ Williams’s question is an excellent way to frame the issues that will be considered in the remainder of this Article: What are the respective claims to authority of methodological austerity and a plurality of explanatory concepts, where the object of study is the rich and varied domain of human behavior?

Regarding the structure of this Article, Part II briefly maps the conceptual terrain, showing the connection between social norms and problems within legal theory. Part III sets forth several models of social norms within the rational-choice tradition, including: Jon Elster’s eclectic account, which attempts to bridge the gap between economic explanations and those offered by other social science disciplines; Richard McAdams’s sanctions-based version, more clearly located within the domain of economic explanation; and finally Eric Posner’s signaling model, which attempts to outdo its competitors in fidelity to rational-choice methodology. The internal critiques in Part IV are principally directed at Posner’s theory, since it represents the apotheosis of the rational-choice method, although along the way I will make some critical observations directed at McAdams and Elster as well. The external critiques of Part V, by contrast, are aimed at the assumptions and methods of rational-choice theory generally. My ambition in this Part is to develop principles of theory acceptance—a metatheory—for theories of social norms as they relate to law. It is only by resolving questions about rational-choice theory at this level that it is possible to move beyond the seemingly interminable debate between law and economics scholars and legal theorists allied with other disciplines, such as the law and society movement, law and literature, and moral philosophy. Law and economics scholars are understandably frustrated by the “mere fulmination of the ‘People don’t behave that way’ school.”¹⁸ What I hope to offer here is something more than fulmination, a critique grounded in criteria of theory acceptance. Finally, Part VI concludes by applying the metatheory of social norms to the arguments offered by rational-choice scholars, by comparing the explanatory terms used in theories of social norms with the concepts used by rational-choice scholars to construct normative theories of law.

touch with reality but also short on humanity.

ELICKSON, ORDER WITHOUT LAW, *supra* note 3, at 7.

15. JON ELSTER, *THE CEMENT OF SOCIETY: A STUDY OF SOCIAL ORDER* 186 (1989).

16. ELICKSON, ORDER WITHOUT LAW, *supra* note 3, at 150.

17. Bernard Williams, *Introduction to ISAIAH BERLIN, CONCEPTS AND CATEGORIES* xi, xvii (Henry Hardy ed., 1978).

18. Thomas S. Ulen, *Firmly Grounded: Economics in the Future of the Law*, 1997 WIS. L. REV. 433, 457.

In this Article I will focus much of my critical attention on the signaling model of social norms, as developed by Eric Posner in a series of articles and his book, *Law and Social Norms*¹⁹ (hence the title of this Article). This emphasis is not a reflection of the success of the signaling model—quite the contrary, it is beset with significant weaknesses. Rather, the focus on signaling is intended to illuminate a fundamental paradox in the rational-choice analysis of social norms. To the extent rational-choice models of social norms seek theoretical purity, their fit with the reality of how human beings experience social interactions is compromised and, as a result, their normative authority is weakened. (As a purely positive matter, these theories may nevertheless enable accurate predictions to be made.) On the other hand, achieving greater fit with phenomenological reality necessitates relaxing some of the stringent assumptions of economic rationality. In other words, theories of social norms are useful approximations of reality only when they incorporate insights derived from social sciences, such as anthropology, psychology, history, and sociology, and the conceptual clarity that can be provided by philosophy. The signaling model shows this tendency clearly. It is constructed with the express aim of tracking closely the assumptions of economic rationality. Any concepts which cannot be reduced to self-interested utility maximization are jettisoned as irrelevant to the workings of the apparatus. By exalting theoretical tidiness over explanatory effectiveness, however, Posner's signaling account stands at odds with more pluralistic theories of social norms which are not so committed to methodological purity. Moreover, the theory fails on its own terms—as it continues to add new epicycles to explain recalcitrant observations, it creates internal inconsistencies that ultimately make the theory implausible. Because I think a similar fate awaits any theory which is too methodologically “thin” or parsimonious, I dwell on the signaling model at some length.

To briefly state my central claim, I will argue that if and to the extent rational-choice theories of social norms aspire to be normative vis-à-vis the legal system, they must be normative in the same way that the law aspires to be.²⁰ In other words, if the point of law and economics scholarship of social norms is to show that the legal system, or some particular rule or practice, can be improved, in the sense of being made more just or fair, by paying attention to informal processes of social ordering, then theory must develop its explanation of social norms in terms of considerations that also function in explanations of justice or fairness.²¹ Only if fairness or justice

19. ERIC A. POSNER, *LAW AND SOCIAL NORMS* (2000).

20. Cf. Barry Friedman, *The Counter-Majoritarian Problem and the Pathology of Constitutional Scholarship*, 95 NW. U. L. REV. 933 (2001) (calling for a more explicitly normative theory of constitutional decisionmaking); Pierre Schlag, *Normativity and the Politics of Form*, 139 U. PA. L. REV. 801 (1991) (calling into questions the dominance of normative legal thought).

21. See Denis J. Brion, *Norms and Values in Law and Economics*, in I ENCYCLOPEDIA OF LAW AND ECONOMICS 1041, 1042-52 (Boudewijn Bouckaert & Gerrit De Geest eds., 2000) (contrasting descriptive and prescriptive varieties of law and economics analysis). To the extent economic theories of social norms intend to be descriptive only, legal scholars can largely ignore them; law and economics becomes controversial in legal theory precisely when it attempts to make prescriptive claims about how legal institutions should be designed and what rules they should enforce.

concerns play no role in the normativity of law may a theory of social norms dispense with these concepts.²² There is considerable overlap here between the legitimacy and authority of law and the values on which the legal system is grounded.²³ If a rational-choice theory of law is objectionable because of its overriding concern with efficiency or aggregate wealth maximization, at the expense of fairness, then a rational-choice theory of social norms can be criticized in the same terms. Similarly, if legal rules should be expressive of substantive moral values, then theories of social norms must also reflect these values, to the extent that the theory has any bearing on the design of legal rules. The principles of theory acceptance, therefore, are derived from legal theory, which in turn derives these principles from conceptions of political legitimacy. In a nutshell, whether we have reason to accept the explanation of social norms offered by rational choice theorists is ultimately a jurisprudential question. This connection has been overlooked in the debate about law and economics, but I am hopeful that framing the debate in jurisprudential terms can help legal scholars move beyond the seemingly interminable squabbling between rational-choice theorists and their critics.

II. THE "PROBLEM" OF SOCIAL NORMS

How is cooperation possible? Since we perceive cooperation around us every day, one might wonder why legal scholars trouble themselves with that question at all. The answer, of course, is that the rational-choice vision of the human predicament has achieved unparalleled dominance in the legal academy in thinking about individual and social behavior, so explanations of informal ordering must jibe with the rational-choice picture. The standard economic story posits a multiplicity of self-interested utility maximizers, each with a set of preferences which he or she attempts to satisfy. Preferences may be altruistic or purely self-interested, but they are held to be exogenous to the model—they are deemed to arise independently of interactions with others; they are the cards which we are dealt, so to speak. Rationality consists of acting in such a way that one's preferences are satisfied at a minimal cost. Of course, no man is an island, so people interact with one another in a variety of settings. Interaction can be negative, as in the case of externalities, where one person imposes costs on another, or positive, when people engage in helping behavior.²⁴ In general, though, the actors would prefer their interactions to be mutually beneficial. Eventually, rational actors realize that they can best further their *separate* interests by agreeing to follow rules prescribing cooperation.²⁵

Because of the structure of interaction, however, the actors find themselves unable to reach cooperative solutions. In the absence of some kind of mechanism for enforcing cooperation, each actor decides what to do independently of the others. In the usual simplified game theory example known as the prisoner's dilemma, two

22. See, e.g., Louis Kaplow & Steven Shavell, *Fairness Versus Welfare*, 114 HARV. L. REV. 961, 1028-38 (2001).

23. See RONALD DWORIN, *LAW'S EMPIRE* 114-216 (1986); PHILIP SOPER, *A THEORY OF LAW* 8 (1984).

24. See ELSTER, *supra* note 15, at 11-12.

25. Robert Sugden, *Contractarianism and Norms*, 100 ETHICS 768, 786 (1990).

players have to choose what strategy to adopt—cooperate or defect.²⁶ The payoffs are structured in such a way that, in the absence of knowledge of the other player's strategy, the smart move is always to defect.²⁷ Of course, the other player is thinking the same thing, and accordingly also chooses the strategy of defecting. The result of this double defection is that both players are worse off than they would have been if they could have managed to secure cooperation.²⁸ Defection thus is the dominant strategy for the game, even though it is suboptimal from the standpoint of both players.

If both players could secure cooperation by agreeing to some kind of joint strategy, they would both be better off. The issue for individuals trapped in a collective-action problem that can be modeled as a prisoner's dilemma therefore becomes how to coordinate actions so that they can produce cooperative benefits. One familiar response is that the state, through the promulgation and enforcement of laws, channels behavior into cooperative enterprises.²⁹ By establishing civil remedies for breach of contract, criminal penalties for theft, and by permitting individuals to create collective entities such as corporations, whose rights and obligations are legally enforceable, the state coordinates what would otherwise be a chaotic welter of individual acts. In game theory terms, defection from the cooperative strategy is the trigger for legal sanctions. Because other players in the game know this, they can play using a cooperative strategy, assuming that the possibility of defection is controlled by state penalties.

As many observers of human societies observe, the state is not necessary to

26. See, e.g., ROBERT COOTER & THOMAS ULEN, *LAW AND ECONOMICS* 33-36 (2d ed. 1997). The typical prisoner's dilemma analysis is highly stylized, but it is a good representation of many real-world problems, such as arms races. See, e.g., THOMAS C. SCHELLING, *THE STRATEGY OF CONFLICT* 213-14, 214 n.4 (1960). Furthermore, if the game is set up as an *n*-person matrix, the result is the tragedy of the commons, a well-known problem in the economics of natural resources and other public goods. See Garrett Hardin, *The Tragedy of the Commons*, 162 *SCIENCE* 1243 (1968).

27. It hardly seems necessary to illustrate the payoff matrix from a prisoner's dilemma—every law professor (and most astute second-year law students) can draw this sort of diagram in her sleep. Nevertheless, for the sake of clarity, I will use the following payoff scheme, taken from ELLICKSON, *ORDER WITHOUT LAW*, *supra* note 3, at 160:

	B: Cooperate	B: Defect
A: Cooperate	3,3	0,5
A: Defect	5,0	1,1

Playing A: If B cooperates, A realizes 5 from defection rather than 3 from cooperation. If B defects, A realizes 1 from defection rather than 0 from cooperation.

28. See, e.g., ROBERT AXELROD, *THE EVOLUTION OF COOPERATION* 7-8 (1984); ELSTER, *supra* note 15, at 17-49; RUSSELL HARDIN, *COLLECTIVE ACTION* 23-24 (1982); EDNA ULLMANN-MARGALIT, *THE EMERGENCE OF NORMS* 18-19 (1977). Using the matrix above, the payoff for each player where both cooperate is 3, as compared with 1, in the case of mutual defection.

29. See generally MANCUR OLSON, *THE LOGIC OF COLLECTIVE ACTION* (1965).

prevent defection. Informal norms of cooperation exist independently of the state. It is part of our common experience that not all breaches of order are remediable by government sanctions: Adultery no longer gives rise to criminal penalties or a tort action for alienation of affections,³⁰ but it may certainly be the basis for informal responses. Family members may remonstrate with the straying spouse, and certain communities, such as religious congregations, may also play an active role in correcting such deviant behavior (these informal shaming sanctions give each spouse an incentive to “play cooperatively” rather than defecting from the marital relationship by cheating). Parties to long-term commercial relationships may similarly refrain from defecting despite the absence of a realistic threat of government penalties, as long as they believe there is some benefit to continued cooperation and as long as the other party to the transaction continues to cooperate. For example, commercial transaction partners may worry about acquiring a reputation as a cheat, which would drive up the cost of cooperating in the future.

In formal game theory analysis, the frequent interactions between players can be modeled as iterated games, where the players meet one another repeatedly.³¹ In an iterated game, each player can adjust his strategy to reflect the propensity of the other player to cooperate, for example, by playing “tit-for-tat.” Playing tit-for-tat means that a player will initially cooperate, but will continue cooperating only if the other player does so as well. As Robert Axelrod has demonstrated, tit-for-tat is the most successful strategy in an indefinitely repeated prisoner’s dilemma game.³² The threat of retaliation causes the other player to continue cooperating, knowing that defection will be punished by her adversary’s defection in a future game. Significantly, the iterated game structure makes cooperation the selfishly rational strategy.

Although social norms seem to solve the problem of cooperation, they pose a particular problem for rational-choice theorists because they seem to disconfirm many of the predictions of the theory; thus, they are often deployed as such by critics of law and economics. Absent a repeated interaction that can be modeled as an iterated game, complying with social norms seems to incur a cost, with no corresponding benefit to the actor. Thus, it is difficult to explain norm-compliance within a theory that takes as a foundational assumption the principle that people always act in their own self-interest, to maximize the satisfaction of their preferences.

For an example of this kind of puzzle, consider the problem of tipping.³³ A visitor to a city eats a sumptuous meal at a restaurant and is presented with the bill; she pays by credit card and receives a little paper receipt to sign, with a line for the server’s tip. If she walks out without signing the receipt, she is likely to be prosecuted for larceny or sued by the restaurant for conversion, but if she signs the slip without tipping the server, there can be no adverse legal consequences whatsoever. Rational-choice theory predicts that the customer will leave without tipping; since she is a tourist, she will suffer no reputational injuries as a result of stiffing the server, and since she can

30. See, e.g., W. PAGE KEETON, ET AL., PROSSER AND KEETON ON THE LAW OF TORTS § 124 (5th ed. 1984).

31. ELLICKSON, ORDER WITHOUT LAW, *supra* note 3, at 164; ELSTER, *supra* note 15, at 43-44.

32. See AXELROD, *supra* note 28, at 30-43.

33. See Saul Levmore, *Norms as Supplements*, 86 VA. L. REV. 1989, 1989-97 (2000) and POSNER, *supra* note 19, at 38-46 for a discussion of the tipping problem.

merely sign the credit card receipt and walk out the door without attracting attention, she may not even be forced to endure informal retaliation such as the glares of the staff. Assuming the customer is interested only in maximizing her utility, she seems to have little incentive to leave a tip. Moreover, the situation as described presents a collective action problem, since the server was probably motivated to provide good service by the prospect of receiving a tip. Because of these expectations, the customer can free-ride on the prevailing social practice of tipping, but secure the benefit of the server's solicitude without paying fifteen percent of a sizeable restaurant tab.

Despite the predictions of rational-choice theory, most people not only leave tips, but tip in appropriate amounts, according to established conventions. It is easy to attribute tipping to social norms, but of course that explanation is question begging if it is offered as an account of the customer's reason for tipping. Thus, rational-choice theorists must fit behavior such as tipping within the framework established by their presumptions about rationality, without resorting to ad hoc explanatory postulates like social norms that cannot be reduced to economically rational action. Some economic theorists adopt an expansive definition of preferences or individual welfare, anticipating problems such as this one. For example, Louis Kaplow and Steven Shavell in their book-length article (and forthcoming book) on welfare economics, define individual well-being as anything a person may value, including "personally held notions of fulfillment, sympathetic feelings for others . . . and [the avoidance of] anything else that the individual might find distasteful."³⁴ This definition begs the question of the problem of social norms in a similar way to the incantatory invocation of social norms to account for recalcitrant data. If any observation can be explained by positing an individual preference for the action (e.g. leaving tips), then it is hard to see how welfare economists can generate empirically testable, and falsifiable, hypotheses.

The tipping example brings up a further puzzle for rational-choice theorists, which is why anyone would bother to sanction someone for violating a norm, something which we know occurs regularly.³⁵ We can vary the example by supposing that the diner was part of a group of friends eating at the restaurant and that she did not leave a tip. After leaving, the person who paid the bill tells her friends that she had stiffed the server, and the friends react with horror. They chastise the bill payer, call her a cheapskate, and perhaps even suggest she should return to the restaurant and tip appropriately. Why would the friends behave this way? They do not know the server, and they received the benefits of attentive service at the restaurant. Moreover, they incur costs by informally sanctioning their friend—perhaps spoiling a pleasant evening with an acrimonious exchange over the tip. Their incentive also seems to be to free-ride on the prevailing norm of tipping, by enjoying their meal, and being quietly thankful that their bill was not inflated by fifteen percent or so. But any armchair sociologist knows this is not what happens, and this is the anomaly that a rational-choice theory must account for. The following section considers three attempts along these lines, to bring social norms within the explanatory apparatus of law and economics.

34. Kaplow & Shavell, *supra* note 22, at 980.

35. See McAdams, *Origin*, *supra* note 3, at 352.

III. "TO SEEM RATHER THAN TO BE":³⁶ RATIONAL-CHOICE MODELS OF SOCIAL NORMS

The concept of social norms has taken the law and economics world by storm, but it took some time for scholars allied with this discipline to construct robust theoretical models of social norms. In fact, the initial proliferation of references to social norms was soon followed by cautionary advice not to invoke "norms" reflexively whenever some inconvenient observation appeared to contradict the predictions of rational-choice theory.³⁷ Several economic theorists have taken up this challenge, and the result is now a number of competing economic models of social norms.³⁸ This Part examines three of the most prominent economic accounts of social norms, those offered by the political philosopher Jon Elster, and the theories advanced by law professors Richard McAdams and Eric Posner.

A. Eclecticism—Elster

Law and economics scholars in the legal academy seem to favor minimal or "thin" theories of social norms, which employ a restricted set of explanatory concepts, generally limited to utility maximization. It is important to point out, however, that there is no reason to believe that an account of human behavior must reside wholly either within or without the economic paradigm. Jon Elster's studies of social organization illustrate the possibility of reaching détente between rational-choice models and other social sciences.³⁹ Elster starts out with the basic problem that vexes rational-choice theory—how can self-interested agents avoid acting in ways that are individually rational but collectively disastrous?⁴⁰ He approaches this question in a somewhat idiosyncratic way, by proposing a conceptual distinction between rationality and social norms as motives for cooperation.⁴¹ This taxonomy is intelligible only if "rationality" is defined narrowly, as economists use the term, as instrumental, means-ends reasoning. In other words, to behave rationally is simply to act to maximize the satisfaction of one's preferences.⁴² Social norms, as we suggested

36. The motto of North Carolina is *esse quam videre*—to be rather than to seem. Ironically, rational-choice theories of social norms are fixated on the seeming, rather than the being, inherent in social life. Whether this represents a bleak and unappealing view of human nature, and whether that vision is correct, are questions to be taken up in Part V.

37. See Sunstein, *supra* note 3, at 945.

38. To be fair to Jon Elster, he has been studying social norms seriously long before the subject was trendy in the legal academy.

39. See ELSTER, *supra* note 15; Jon Elster, *Norms of Revenge*, 100 ETHICS 862 (1990).

40. ELSTER, *supra* note 15, at 17.

41. *Id.* at 35-36, 98.

42. See, e.g., Ulen, *supra* note 18, at 457. As Ulen lists them, the assumptions of rational choice theory are that "[d]ecision-makers are rationally self-interested; they have complete, transitive, and reasonably stable preferences; they can learn about and compute the costs and benefits of alternative courses of action; and they seek to maximize as many of their preferences as they feasibly can." See also ELSTER, *supra* note 15, at 98 ("Rational action is concerned with outcomes. Rationality says, 'If you want to achieve Y, do X.'"); HARDIN, *supra* note 3, at 46-

above, compel irrational behavior, because following them does not tend to improve the chances of achieving one's ends. In a different sense, however, there is nothing irrational about complying with social norms. Rationality may be defined differently, as something like "responding appropriately to reasons,"⁴³ in which case an actor would be rational if she followed a social norm for the appropriate reason. It is not necessary to enter into the debate over economic rationality at this point, because Elster's model of social norms does not really hinge on his definition of rationality, but other rational-choice theorists, including Eric Posner, do rely quite explicitly on the thin means-ends sense of rationality as a foundational premise in their explanations of social norms.

We can duck problems with the definition of rationality here because self-interest is not the only motivation allowed in the explanations offered by Elster:

I have invoked a larger variety of individual motivations [than most writers on rational choice]. Most writers try to make do with rational self-interest as a sole motivational assumption, while I have invoked a broader range of motives. Though I share their preference for a parsimonious explanation and their hesitation to get into a morass of ad hoc assumptions, I have concluded, with some reluctance, that there is no way in which the programme can be brought forward on this narrow basis. Ultimately, parsimony must take second place to realism.⁴⁴

Here Elster is alluding to criteria of theory acceptance. His account sacrifices some amount of parsimony—that is, explanatory unification in terms of the smallest possible number of independently acceptable assumptions⁴⁵—in favor of a plurality of concepts that are already familiar to students of social interaction, such as sociologists, psychologists, and philosophers. His theory is not necessarily incompatible with those that posit fewer independent explanatory concepts, such as Posner's. But there is a certain tension here, revealed by the pejorative label "ad hoc assumptions" applied to explanatory terms that cannot be reduced to the premises of economic rationality.

Unlike many rational-choice theorists, however, Elster is willing to concede that people are motivated by different types of reasons—including rough everyday moral concepts such as fairness, consequentialist considerations, and the "everyday" Kantian intuition that an action is not permissible unless the reasoning behind it can be generalized—and that selfish rationality alone is insufficient to explain many instances of successful cooperation. These norms are not outcome oriented; that is, they are obligatory for actors even though they do not tend to maximize the

47. I refer to this conception as means-ends rationality because it takes ends ("achieve Y") as given, and does not permit rational deliberation about ends.

43. See, e.g., T. M. SCANLON, *WHAT WE OWE TO EACH OTHER* 23 (1998) ("A rational creature is . . . one that has the capacity to recognize, assess, and be moved by reasons . . ."); Elizabeth S. Anderson & Richard H. Pildes, *Expressive Theories of Law: A General Restatement*, 148 U. PA. L. REV. 1503, 1509-10 (2000) (defining rationality as responding appropriately to reasons for action given by evaluative attitudes).

44. Elster, *supra* note 15, at 250.

45. See SALMON, *CAUSALITY*, *supra* note 13, at 69-70.

probability that an actor will achieve her ends.⁴⁶ To underscore this point, Elster does not argue that these reasons can be derived from the assumptions of economic rationality. In fact, he repeatedly emphasizes that “social norms provide an important kind of motivation for action that is irreducible to rationality or indeed to any other form of optimizing behavior.”⁴⁷ This claim puts him at odds with the majority of rational-choice theorists of social norms, as we will see.

Norms, as opposed to legal sanctions or considerations of economic rationality, are sustained by social approval or disapproval, either explicit and external, or as it is internalized in the form of sentiments of embarrassment, guilt, or shame.⁴⁸ It is this emotive impact of norms that motivates compliance, rather than the cognitive impact, weighing costs and benefits, that is associated with externally imposed sanctions. The key to the definition of social norms is that their violation generates evaluative responses such as blame or guilt.⁴⁹ In other words, the norms are *normative*—they serve as the basis for evaluation, whether praise or criticism, which can come from sources outside the actor (family, neighbors, coworkers, and so on), or can be internalized by the actor as the result of a lifelong process of socialization. There would be no point to criticizing someone for failing to maximize her own utility, but as we know, social norms do serve as the basis for normative judgments. Consider the evaluative terms we apply in daily life to people who fail to observe social norms: boor, redneck, oaf, philistine, jerk, litterbug, egomaniac, selfish bastard, and so forth. The ascription of these terms to a norm-violator would make no sense unless there was a generally accepted social understanding that norms ought to be followed, whether or not they tend to satisfy individual preferences. This basic quality of a norm, as opposed to a habit or a mere coincidence of behavior, is easily overlooked by rational-choice theorists who are concerned above all with shoehorning norm-compliance into economic rationality.

So where do norms come from? Elster is content to leave the existence of norms that cannot be reduced to self-interest as an exogenous “brute fact,” at least pending additional study.⁵⁰ Rational-choice theory need concern itself only with identifying the right mix of self-interested and normative motivations behind the action.⁵¹ This is fine as far as it goes—Elster is more concerned with alerting rational-choice theorists to the existence of social norms, and the impossibility of reducing them to self-interest, than he is to explaining their origin—but scholars who have followed him have thought it necessary to give an account of norm development that is congenial to the assumptions of economic rationality. The following subparts consider two of these attempts.

B. Sanctioning—McAdams

Richard McAdams’s account of social norms is something of a midpoint between Elster’s eclecticism and the austere signaling theory offered by Posner. In a series of

46. See ELSTER, *supra* note 15, at 98.

47. *Id.* at 15; *see id.* at 125, 133-34, 150.

48. *Id.* at 99.

49. *Id.* at 104-05.

50. *See id.* at 150.

51. *See id.* at 187.

articles, McAdams has sought to elaborate a rigorous model of social norms that is primarily indebted to economic analysis, but which does not ignore or seek to displace the insights of other disciplines.⁵² His goal is to provide more of a theoretical underpinning for social norms, within the framework of rational-choice assumptions, than Jon Elster's work offers.

Fundamental to McAdams's account is the preference most well-socialized people have for esteem—we want others to think well of us.⁵³ As a result of this preference, our peers can impose sanctions on us for violating social norms, simply by withholding esteem. It is critical that denying esteem is costless; otherwise people face a collective action problem.⁵⁴ Each person's disapproval only counts for so much. Standing alone, a look of disgust or whispered rebuke carries only so much disesteem. If people acted together, however, they could aggregate their disesteem and impose substantial costs on a norm-violator. McAdams imagines everyone in the neighborhood getting together and withholding approval from a person who litters, thereby driving up the cost of littering and giving the litterbug an incentive to carry his trash to the nearest receptacle. But since it requires the coordinated action of all the neighbors to shame the litterer effectively into putting trash in its proper place, any of the neighbors would have an incentive to free-ride on the sanctioning of others if sanctioning were costly⁵⁵ (provided that most other neighbors participated in sanctioning the litterbug, the individual would enjoy the benefit of clean streets without incurring any cost). McAdams argues, however, that there is no second-order collective-action problem associated with informal social sanctions, since conveying disapproval is costless.

There are good reasons to be skeptical of this argument. In the tipping example above, I assumed that the friends would, in fact, bear costs associated with sanctioning the cheapskate. At a very minimum, remonstrating with the tightwad would inject an element of negativity into an otherwise pleasant evening. In the worst case, the confrontation could spiral out of control, and mushroom into a fight that destroyed the friendship. In McAdams's littering example, there would certainly be costs attached to responding to the litterbug. If the person were a stranger, the neighbors would have to reckon with the possibility that they would be putting themselves in harm's way by chastising him for littering. (I lived in New York City, and cannot imagine confronting a stranger on the street who littered. The result would at least be a vitriolic argument, and there would be a nontrivial possibility of a violent response.) If, on the other hand, the person were an acquaintance, then the sanctioners would have to worry about spoiling their relationship with someone they must see frequently. McAdams might respond that the neighbors can costlessly withhold esteem, but it seems difficult to withhold esteem effectively without

52. See McAdams, *Discrimination*, *supra* note 9; Richard H. McAdams, *Group Norms, Gossip, and Blackmail*, 144 U.P.A.L. REV. 2237 (1996) [hereinafter McAdams, *Group Norms*]; McAdams, *Origin*, *supra* note 3; Richard H. McAdams, *Relative Preferences*, 102 YALE L.J. 1 (1992); McAdams, *Signaling*, *supra* note 11.

53. The discussion in this paragraph is drawn from McAdams, *Origin*, *supra* note 3, at 355-75.

54. See HARDIN, *supra* note 3, at 52-53; Robert Axelrod, *An Evolutionary Approach to Norms*, 80 AM. POL. SCI. REV. 1095 (1986).

55. See ELSTER, *supra* note 15, at 41, 132-33.

exhibiting some kind of negative attitude toward the miscreant. Perhaps neighbors can work behind the scenes, by spreading gossip,⁵⁶ but this also carries costs. No one wants to acquire a reputation as a scold or a busybody, and spreading gossip can backfire if the object discovers its source and retaliates against the person spreading rumors.⁵⁷

McAdams recognizes another difficulty inherent in explaining norm development through sanctioning. The problem is that a person will go out of her way to sanction a norm-violator only where it is already clearly established that there is a norm on point.⁵⁸ There is considerable risk of embarrassment attached to publicly expressing disapproval of another's actions. Even if one accurately gauges the violator's transgression, the sanctioner has to worry about appearing to be a sanctimonious intermeddler, like Dana Carvey's "Church Lady" character from "Saturday Night Live." But the real downside is misapprehending the relevant standards of behavior and becoming an object of ridicule for bungling the local norms. For example, suppose a person jaywalks in front of an observer. The observer, concluding that there must be a norm against jaywalking, because it is destructive of community and civility, publicly berates the jaywalker. In Seattle, this sanction would in fact reinforce a norm against jaywalking, along the lines McAdams suggests. (On several occasions I have observed pedestrians there waiting for the light to turn green at 1:00 a.m. in the rain.) People would conclude that the observer is a real Seattlite, and probably ask her for directions to the nearest latte cart. By contrast, in Manhattan, jaywalking is perfectly acceptable, at least by social-norms standards (the police department's feckless attempt to write tickets for jaywalking failed in the face of widespread acceptance of the practice by pedestrians). The observer would therefore look ridiculous if she expressed disapproval of jaywalking.

56. See McAdams, *Origin*, *supra* note 3, at 362.

57. McAdams states that gossip is "a consumption good, a pastime rather than a burden," citing numerous anthropological and sociological studies that reveal a widespread love of gossip among people of different cultures. McAdams, *Group Norms*, *supra* note 52, at 2244. I have not delved into these sources specifically, but I do know from other work on the anthropology of honor and shame that interfering with another's reputation is a dangerous game, and can easily backfire if the object of gossip or innuendo learns of an effort to undermine his or her good name. See, e.g., KENNETH S. GREENBERG, *HONOR & SLAVERY* (1996); FRANK HENDERSON STEWART, *HONOR* (1994); WILLIAM IAN MILLER, *BLOODTAKING AND PEACEMAKING: FEUD, LAW, AND SOCIETY IN SAGA ICELAND* (1990); HONOUR AND SHAME: THE VALUES OF MEDITERRANEAN SOCIETY (J.G. Peristiany ed., 1966); BERTRAM WYATT-BROWN, *SOUTHERN HONOR: ETHICS AND BEHAVIOR IN THE OLD SOUTH* (1982). I cannot dispute McAdams's contention that some people enjoy gossiping, but I also find it implausible that it is a completely costless activity in any social setting. As Elster notes, no one wants to be known as a goody-two-shoes, since "[d]o-gooders often make others feel bad." ELSTER, *supra* note 15, at 189 (emphasis omitted). (McAdams does later concede that gossip is discouraged where it is merely the passing of damaging information *without adequate reason*. See McAdams, *Group Norms*, *supra* note 52, at 2281-82. Adequate reason here would have to be sufficient to overcome the stigma of being regarded as sanctimonious or meddling.) The truth is probably somewhere in between, and too complicated to be captured by a simple model—some people enjoy gossip some of the time, but have the good sense not to gossip when it is dangerous, while other people do not enjoy gossip at all or generally disapprove of gossip.

58. See McAdams, *Origin*, *supra* note 3, at 362-64.

To continue the example, suppose we take a person who is unfamiliar with the culture of either city and drop her into one or the other at random. Will she, observing an act of jaywalking, publicly show disapproval? Probably not, recognizing as McAdams does that “[w]here there is doubt about what the consensus is, there are esteem risks to expressing what may turn out to be a minority opinion.”⁵⁹ No one wants to look like a rube, so before sanctioning a jaywalker, the observer will hold her tongue and watch what others do, to see whether there is a norm against jaywalking in the city.⁶⁰

McAdams is surely correct that potential sanctioners of norm-violations first “test the waters” to make sure there is a consensus before publicly taking a position.⁶¹ But after making this concession, the sanctioning model can explain only ongoing enforcement of existing norms, not the development and evolution of norms in the first place. Suppose we are wondering whether a new norm will develop, requiring people to let passengers off the subway before forcing their way on. If everyone is testing the waters, by looking around to see whether others express disapproval of passengers who shove their way onto the subway, they will find that there is no existing norm to enforce. What, then, would cause people to begin expressing disapproval of aggressive subway passengers? Here is where McAdams makes another crucial concession, one that prevents his theory from being as methodologically pure as Posner’s signaling account, but one that ensures that his model will not fall into the pitfalls I will discuss below regarding the signaling model. McAdams notes that people make evaluative judgments of one another “*independent of and prior to any [social] norm*—individuals . . . are not utterly indifferent to all traits and behaviors.”⁶² People care about things like civility, fairness, public safety, and an environment free of litter, and it is these values that motivate people to create and enforce social norms against littering, jaywalking, stiffing waiters, and pushing their way onto the subway. Social norms are parasitic on other norms and values whose sources are not simply the approval and disapproval of others. Rather, they are grounded in (to borrow a term from political philosophy) conceptions of the good—visions of what is worthwhile in human life. When people express disapproval of littering, they are enforcing a norm that is ultimately grounded in values such as respect for the environment and an aesthetic appreciation for litter-free surroundings.⁶³

59. *Id.* at 362-63.

60. Travelers can consult many guidebooks containing explanations of tipping customs in foreign countries not only to understand tipping practices, but to know when it is appropriate to express disapproval of one’s companion’s failure to tip. An informal version of a guidebook might be the presence of a few members of the public who are willing to bear the risks of being out of step with prevailing customs, and therefore sanction violators of an evolving norm. See McAdams, *Group Norms*, *supra* note 52, at 2251 (proposing a modification of Robert Cooter’s internalization account of norm development).

61. McAdams, *Origin*, *supra* note 3, at 363-64.

62. *Id.* at 358 (emphasis in original). I added the bracketed word “social” because evaluative judgments are based on norms, albeit not “social norms,” but norms of ethics, etiquette, and so on.

63. McAdams refers to these considerations—such as “be a loyal friend” or “do one’s share”—as “meta-norms” or “abstract norms,” perhaps to avoid biting the bullet and talking

It may be true that the first person to publicly express disapproval of a new or evolving norm may have to be a "hero,"⁶⁴ but it is significant that the ascription of heroism to someone who is ahead of the curve of social norm development is also not arbitrary. We call people heroes or prophets who are the first to articulate a norm that may have been inchoate in or peripheral to the community's practices, but we do not label as heroes the people who misidentify a norm. Martin Luther King, Jr., was a hero; Ted Kaczynski was not. The difference between King and Kaczynski is not just the fact that King's vision attracted followers and Kaczynski's did not; it is that King articulated an authentic vision of the good for humans, while the Unabomber advocated a conception of society that excessively devalued the dignity of humans relative to the value of nature.

This example also suggests that social norms would not coalesce around actions that have no significance whatsoever, in terms of values that matter to people. There would be no point in making an ostentatious public show of disapproval of another person's decision which was a matter of indifference from the point of view of ethics, etiquette, or aesthetics. To borrow Ronald Dworkin's illustration, suppose your neighbor prefers rum-raisin ice cream to vanilla. No one is going to sanction her choice of ice cream flavors, because the preference for rum-raisin is arbitrary from the standpoint of anything that makes a difference to a well-lived human life. (In the same manner, some arguments for toleration of gays and lesbians seek to establish that sexual preference is similarly not a matter of moral significance.) If, on the other hand, your neighbor has a preference for torturing people, or playing loud music at 2:00 A.M., or littering, or racial discrimination, she would be subjected to the disapproval of others. This disapproval would be directed at the norm-violator because her neighbors think it is worth their time, and the cost associated with sanctioning, to bring her into line with community standards. Why? Not because those community standards just happen to be a social norm, but because they matter, for some reason that is quite independent of their being the object of social consensus.

Let me be the first to applaud an economic model of social norms that makes room for noneconomic values. I have no quarrel with McAdams's contention that specifically *social* norms, as opposed to moral or aesthetic norms, may be the product of publicly observable sanctioning behavior by others. McAdams is not concerned with asking why people have evaluative opinions about others, but that is not a fault of his model. He is interested in the mechanism by which norm-enforcement occurs. But in order to expand his account into a full *explanation* of human behavior, it would be necessary to inquire into the sources of value which ground the presocial-norm evaluative judgments he identifies. That is the task for some discipline other than economics—perhaps art, literature, philosophy, social criticism, or psychology. As long as rational-choice theorists are content to live harmoniously alongside other humanistic disciplines whose insights are relevant to explaining social life, they are to be welcomed. But some members of the law and economics circle have overweening ambitions to explain even the sources of value that underlie social

about moral values like loyalty and fairness, of which these abstract norms are surely an instantiation. See *id.* at 383-85, 395; McAdams, *Group Norms*, *supra* note 52, at 2252 & n.46.

64. See McAdams, *Origin*, *supra* note 3, at 369-70.

norms. It is when economic theories approach this kind of methodological imperialism that their critics turn particularly fierce.

C. Signaling—Posner

Eric Posner's analysis of social norms is animated by an almost obsessive concern with methodological purity (as well as an inordinate fondness for game theory jargon). Where Jon Elster is content to explain collective action on the basis of a wide variety of motivations, Posner denounces such a pluralistic framework as "phenomenologically attractive but methodologically sterile."⁶⁵ And where Richard McAdams admits that norms tend to arise where there is some underlying value at stake, Posner wishes to eliminate "messy" explanatory concepts like values from his model. As I will argue in the following Part, the division between Posner, on the one hand, and such law and economics scholars as Elster, McAdams, Ellickson, Sunstein, and Bernstein, on the other, illustrates the predicament of rational-choice theories: either they accommodate a wide range of human motivations, but risk complicating neat theoretical assumptions, or they posit a desiccated conception of human nature in order to remain faithful to the principle of parsimony. As Robert Scott rightly observes, richer analysis comes with a price, in the form of loss of some degree of empirical falsifiability, but it brings with it the benefit of enabling a more satisfying explanation of social life.⁶⁶

Posner's model of social norms is provocative not only because of its theoretical austerity, but because it turns the "problem" of social norms on its head. For rational-choice theorists, social norms are a puzzle. Why would someone leave a big tip when there is no legal or nonlegal sanction for failing to do so? Why do people buy a new set of "in fashion" clothes every year when they have perfectly good ones from last year that still fit? These behaviors seem to contradict the assumption that people are rational utility maximizers. For Posner, however, these actions are explicable as means to solve a problem relating to cooperation—namely, how we can find suitable partners for long-term cooperative arrangements, relationships which, if maintained over the long run, will yield a cooperative surplus in excess of the utility that each partner could have achieved on his or her own. Social norms, Posner argues, are a communication or "signaling" device, by which reliable transaction partners can find each other and enter into cooperative relationships.

Imagine a person looking for a transactional partner, keeping in mind Posner's broad definition of "transaction"—this person could be seeking a spouse, a business associate, an employee, or just a buddy. This person needs a way to separate the sheep from the goats, so to speak. She must ascertain which persons out of a pool of strangers are likely to defect from cooperative arrangements and which are likely to stick around, incurring short-term losses for the sake of realizing a long-term cooperative surplus. It would be useful if everyone in such a marketplace had information about the disposition of the others, but we are assuming that the participants in the market are strangers to one another. In a small community, where

65. POSNER, *supra* note 19, at 235 n.5. He uses the same epithet to describe the argument that people internalize social norms. *See id.* at 43.

66. *See* Scott, *supra* note 11, at 1607-08.

everyone knows everyone else, there is no need to acquire information about the cooperativeness of others—it is already widely known. This situation is represented in game theory as an iterated game, in which the players can adjust their strategies to the previous moves made by the others. Someone who is cooperative and reliable can communicate this disposition to others by declining to cheat when the opportunity presents itself. The player thus acquires a reputation for cooperativeness, and attracts additional transactional partners. In a large, impersonal community, characterized by the absence of information obtained from previous interactions, potential cooperators can always try to communicate their reliability to one another, but it seems quite easy to mimic signs of suitability as a transactional partner.⁶⁷ Con artists might utter the right soothing words, and then abscond with their partners' money after the second party had committed to the transaction. In game theory terms, the disposition of potential transacting partners is private, nonverifiable information.⁶⁸ It cannot be known directly, and it is easy to fake, since talk is cheap—in the absence of some kind of formal agreement, reassurances of one's reliability are not binding.⁶⁹

It is important at this juncture to mention a significant work in another discipline which presents the same kind of problem that Posner regards as fundamental. Moral philosopher David Gauthier attempted a project that was even more ambitious than Posner's attempt to explain social norms using economic methodology; Gauthier's task was to derive the content of the rules of *morality* from the assumptions of economic rationality.⁷⁰ In *Morals By Agreement*, Gauthier sought to explain both the rationality of agreeing to substantive principles of morality and the content of the principles themselves, by reference only to a thin conception of rationality as utility maximization. As many commentators pointed out, this is not an entirely original project; Hobbes, for one, had much the same agenda.⁷¹ But Gauthier's attempt was

67. See Alan Nelson, *Economic Rationality and Morality*, 17 PHIL. & PUB. AFF. 149, 156 (1988).

68. See, e.g., DOUGLAS G. BAIRD, ET AL., *GAME THEORY AND THE LAW* 122 (1994).

69. See *id.* at 303 (defining game theory concept of "cheap talk"). Posner occasionally argues that talk alone can function as a signal of discount rate. See, e.g., POSNER, *supra* note 19, at 194-97. The examples he gives are of incommensurability claims—for example, someone's claim that he would be offended if a friend offered him \$20 to make amends for missing lunch together. As the discussion below indicates, however, Posner's theory depends on an action being costly in order to function as a signal, and there seems to be very little cost associated with making incommensurability claims. People unquestionably say things like, "You are infinitely valuable to me," to spouses and family members, and these comments do function to maintain stable cooperative relationships, but they do not reflect the speaker's discount rate, as Posner's model requires.

70. See DAVID GAUTHIER, *MORALS BY AGREEMENT* (1986). Judging by the citations in major law and economics works, Gauthier's work also seems relatively unknown in rational-choice circles. Considering the amount of energy that rational-choice scholars spend fending off criticism from philosophers, it is surprising that few, if any, have examined the treatment that Gauthier's work has received at the hands of his critics.

71. See, e.g., Jody S. Kraus & Jules L. Coleman, *Morality and the Theory of Rational Choice*, 97 ETHICS 715 (1987). Rawls also attempts to derive substantive conceptions of (political) justice from assumptions about rationality. See JOHN RAWLS, *A THEORY OF JUSTICE* (1971). Unlike Gauthier, however, Rawls does not begin with economic rationality. Rather, his "thin theory of the good" is a more nuanced kind of rationality. See *id.* § 25, at 142-50. For

unusual in the austerity of its initial assumptions about human society and rationality, and in its strong conclusion that not only are morality and self-interest compatible, but that morality is derivable from self-interest. (Hobbes's arguments related more to the political domain, as opposed to individual morality.) Along the way, Gauthier encountered exactly the same problem as Posner grapples with—namely, how people with a disposition to cooperate can match up with one another, instead of getting stuck in long-term relationships with people who are inclined to cheat or defect.

Gauthier divides the world up into "straightforward maximizers" ("SMs"), those who are disposed to maximize their utility thinking strategically only from their own point of view (that is, assuming that others will also straightforwardly maximize), and "constrained maximizers" ("CMs"), persons who are disposed to maximize their utility in light of the utilities of others with whom they interact.⁷² In other words, CMs base their actions on a joint strategy—on the possibility of achieving a cooperative surplus—and can therefore obtain the benefits of cooperation that are unavailable to SMs.⁷³ This is where Gauthier poses the same problem that Posner claims to solve—how CMs can find one another, so that they do not interact with SMs, expecting to receive cooperation in return for cooperation, and find themselves exploited by the SMs. Gauthier does not have a promising solution to this problem, except to suggest rather optimistically that CMs can recognize others' dispositions to cooperate "not with certainty, but as more than mere guesswork."⁷⁴ As several of his reviewers have pointed out, Gauthier's theory is somewhat embarrassed by its resort to some contrived assumptions about the probabilities of CMs being able to recognize one another in a large, impersonal marketplace.⁷⁵

One strength of Posner's account is that it offers a more promising solution to the problem of imperfect information about others' dispositions. He suggests that participants in the marketplace can engage in actions that are costly—and thus difficult to fake—which reveal vital information about the reliability of the actor. These actions function as *signals*, again using a game theory term, meaning that they are actions which convey private, nonverifiable information to other "players."⁷⁶ In the context of seeking partners for long-term commercial or social relationships, the most important piece of data to communicate, Posner argues, is discount rate. Some people are willing, as a matter of psychological disposition, to incur short-term losses (or at least not realize short-term advantages) in pursuit of longer-term cooperative gains.⁷⁷ Borrowing another game theory term, Posner calls these "good types," and

example, the parties in the original position are assumed to be capable of a sense of justice, so that agreements reached behind the veil of ignorance will be respected. *Id.* at 145. Note that even this "thin theory of the good" is subject to the objection, which has been made by several critics of Rawls, that the outcome of deliberations in the original positions is loaded up by a contestable definition of rationality. *See, e.g.,* Thomas Nagel, *Rawls On Justice*, in *READING RAWLS: CRITICAL STUDIES ON RAWLS' "A THEORY OF JUSTICE"* 11 (Norman Daniels ed., 1989). Thus, the results are not in fact independent of anyone's conception of the good.

72. GAUTHIER, *supra* note 70, at 166-67.

73. *Id.* at 170.

74. *Id.* at 174.

75. *See, e.g.,* Nelson, *supra* note 67, at 160.

76. *See* POSNER, *supra* note 19, at 122-23.

77. *See id.* at 19-21.

they seem to be roughly equivalent to Gauthier's CMs. Others are apt to defect from cooperative endeavors whenever they can obtain short-term benefits from withholding cooperation. Posner labels this second category, which resemble Gauthier's SMs, "bad types."⁷⁸ He is at pains to underscore that the adjectives good and bad do not refer to moral evaluations; they are simply convenient labels chosen to mark off a rough distinction between discount rates.⁷⁹ If the two types do have any moral quality, it is in the difference between cooperative and opportunistic dispositions, but Posner (somewhat dubiously) insists that this distinction can be captured by the concept of discount rate.⁸⁰ Good types thus endeavor to show potential transactional partners that they have a low discount rate, by acting in ways that reveal their disposition. These actions function as signals in the multiplayer game of trying to match up with suitable cooperative partners. It is worth going through the bother of sending signals, because bad types, or SMs, will be excluded from cooperative arrangements in which both actors can increase their expected utility beyond that which they could achieve by following an individual strategy.

In order to function as a signal, an action must be costly, observable, and arbitrary.⁸¹ Costly, because they must reveal the disposition of the actor to incur short-term costs where they are necessary to realize long-term surpluses from cooperation, and because they must not be capable of being simulated by bad types looking to bamboozle good types. (Remember that talk is cheap.) "Signals reveal type if only the good types, and not the bad types, can afford to send them, and everyone knows this."⁸² Observable, because signals communicate information to strangers who otherwise lack information about the disposition of the actor. And arbitrary, because observers in the marketplace must satisfy themselves that the action is taken in order to send a signal; if the actor does something because he has a preference for doing so, it says nothing about his discount rate. An altruist may enjoy giving gifts, and may even derive a net benefit from her generosity, so observing her magnanimity does not provide information about her willingness to incur short-term costs.⁸³ If a putative signal does not satisfy these criteria, the result may be a "pooling equilibrium," in which the good types and bad types are lumped in together as potential transaction partners (or no one matches up at all), rather than a "separating equilibrium," in which good types only match up with other good types.⁸⁴

Posner asserts that a bewildering variety of actions can function as signals: A merchant who is new in town might invest in expensive office space, to convince lenders, suppliers, and customers that she is committed to remaining in the community, because she could make enough money to pay for the cost of the

78. See *id.* at 18-20. Gauthier uses the terms "constrained maximizer" and "straightforward maximizer," respectively, to express the distinction between good types and bad types. See GAUTHIER, *supra* note 70, at 15.

79. See POSNER, *supra* note 19, at 25.

80. See *id.* at 18.

81. *Id.* at 22.

82. *Id.* at 19.

83. See *id.* at 22.

84. *Id.* at 19-20. For concise definitions of these terms, see BAIRD, *supra* note 68, at 312, 314.

expensive facilities only if she stayed in business for a long time.⁸⁵ We follow an elaborate, albeit unwritten, code of social etiquette: chewing with our mouths closed, uttering the right social platitudes (for example, answering “fine” when asked how we are doing, rather than “rotten!”), pretending to listen when our coworkers tell boring stories, speaking without a regional accent, grooming ourselves correctly, and so on.⁸⁶ Giftgiving is a particularly rich example of signaling behavior, in both business and social settings. Employers throw big holiday parties, bosses give flowers to their secretaries on Secretaries’ Day, and professionals take clients out to lunch, where business is discussed only after the requisite period of small talk on personal matters.⁸⁷ People give one another gifts, some in a perfunctory or ritualistic way, like the proverbial Christmas fruitcake,⁸⁸ others with great solemnity, like engagement rings.⁸⁹ Americans fly flags and join in rallies and parades, more zealously during times of national crisis, and some pathetic fraction of eligible voters actually makes it to the polls to vote.⁹⁰ Posner actually argues (and he seems quite sincere about this) that *all* of these patterns of behavior can be explained as attempts to communicate one’s discount rate to potential transactional partners. I have no doubt that many of these actions do have significance as signals, in addition to whatever underlying nonsignaling purpose they may serve. Bosses do not mind if they are observed being nice to their secretaries, for it makes employees more apt to trust them and work hard. My objection, rather, is Posner’s argument that norms are essentially arbitrary and unconnected with any other kinds of reasons the actor may have for engaging in the behavior in question.

Posner has to resort to a highly strained account of how certain actions become signals, because of his insistence that the ascription of signaling to an action be arbitrary. Imagine that everyone at a workplace went jogging during their lunch hour because they were all fitness buffs. Posner would say that lunchtime jogging is not a signal of a low discount rate (even though it is costly and observable), and therefore not a social norm, because of the presence of a nonsignaling reason, namely the shared desire to better one’s health. Social norms, in the Posnerian sense, can arise only with respect to actions that people have no reason to engage in other than signaling discount rate. So, at this same workplace, wearing a particular color of

85. See POSNER, *supra* note 19, at 20-21.

86. *Id.* at 22. As Dave Barry would say, I am not making this up—these are really Posner’s examples.

87. *Id.* at 23-24, 151.

88. *Id.* at 52, 169-71. Posner’s illustration is the Chinese custom of giving “moon cakes,” which apparently are “quarter-pound cakes . . . filled with a paste of mashed lotus root, sugar and oil . . . [and are] a rare rib-sticking treat that would keep well into the icy winter months when most people subsisted on cabbage.” *Id.* at 169 (quoting Craig R. Smith, Moon Cakes: Gifts That Keep on Giving and Giving and . . ., WALL ST. J., Sept. 30, 1998, at A1). Despite this mouth-watering description from the *Wall Street Journal*, Smith, *supra*, Posner reports that no one really likes to eat moon cakes. *Id.* at 171. For what it’s worth, however, Peter Huang told me that he likes moon cakes. This by itself does not disconfirm Posner’s theory, of course, but it does serve as a reminder that many of his blanket statements about people’s preferences and reasons should be taken with a grain of salt.

89. *Id.* at 71.

90. *Id.* at 113.

jogging shorts while exercising might be a signal. This sounds implausible, but Posner suggests that there may be “norm entrepreneurs,” who engage in public actions and hope that others will follow.⁹¹ If public behavior does indeed coalesce around the entrepreneur’s action, then a norm develops (if imitation fails to develop, the norm entrepreneur becomes an object of ridicule). Posner seems to imagine a person showing up one day wearing chartreuse running shorts, hoping that others will start copying the fashion. A social norm would result only if others wore the chartreuse shorts, having no reason to do so other than their desire to show others that they have a low discount rate—they were willing to absorb the cost of going to the store to buy a new pair of shorts because that is the fashion. Of course, this is a bizarre example, but one is forced to come up with odd hypotheticals in order to settle on an action that is completely arbitrary from the point of view of the preferences and tastes of the actors.

Because in Posner’s system a norm cannot be supported by nonsignaling reasons, there seems to be no reason for others to emulate the public action taken by the putative entrepreneur. It is unclear why anyone should risk exposing himself to ridicule by copying the public action of a wannabe norm entrepreneur, when there is no reason to take the action other than the vague hope that the action may eventually evolve into a signal of discount rate. Unlike the case with McAdams’s “heroes,” Posner cannot explain the success of his norm entrepreneurs by pointing to some underlying value that they expressed and made socially salient. Again, Posner has unwisely abandoned a previously plausible position, in order to hew to an austere rational-choice model. In a previous article, Posner, like McAdams, agreed that norms generally originate around actions that further the community’s “sense of justice or of the general good” or are based in commonly shared moral values.⁹² The “village gossips” who enforce norms do not direct their disapproval arbitrarily, but aim it at actors who transgress some community value. In Posner’s Coasian example, the gossips retaliate against railroads whose trains emit sparks that catch on fire the stacks of flax by the tracks. I think the example is slightly misleading, but not because the village scolds would not direct their ire at the railroad. Rather, their disapproval would probably be underwritten by some rough sense of corrective justice (the railroad should pay for what it damages) rather than a full-blown economic-efficiency argument about maximizing the aggregate value of the farmers’ crops and the

91. *Id.* at 30-32.

92. See Eric A. Posner, *Law, Economics, and Inefficient Norms*, 144 U. PA. L. REV. 1697, 1713-14, 1720 (1996) [hereinafter Posner, *Inefficient Norms*]. Compare Russell Hardin’s analysis of Cubs fans, who use their loyalty to the team to signal *something*—being a local resident, having ties to the Chicago area (for Cubs fans living elsewhere), or even admiring scrappy, tenacious teams. HARDIN, *supra* note 3, at 53-55. People use Cubs loyalty as a means of lowering the transaction costs that would otherwise make it difficult to coordinate with people with similar interests. Identifying with the Cubs has a signaling function here, but what Hardin recognizes, and Posner refuses to concede, is that the signaling behavior does not coalesce around actions which are arbitrary; critically, actions become signals because they have an interpretive or meaningful dimension. Cubs fans want to hang out with Cubs fans, not because they are willing to make a public display of their loyalty to some arbitrary totem, but for the reason that they are from Chicago, have similar rituals, customs, and tastes, and can therefore more easily find common interests.

railroads' investment. This objection will be developed further in the concluding sections of this paper, where I take issue with the psychological assumptions of rational-choice theories of social norms.

Posner does say that norm entrepreneurs are too valuable to be shunned—they are often entertainers or politicians with great fame and power—but this response simply begs the question of why they are rich and prominent in the first place. Britney Spears, Jennifer Lopez, Madonna, or whoever, is not popular because she was too important to be shunned; rather, she *became* popular after the actions she adopted proved to be persuasive, in the sense of being successful in the marketplace of potential norms. Unlike Bill Gates, Madonna became rich by suggesting norms for adoption by teenagers; the interesting question is why the proposed Madonna norms took off, while the fashions, hairstyles, and attitudes offered by some forgotten B-movie actress did not.

Unfortunately, Posner has nothing at all to say about the reasons some proposed norms are persuasive and others are not. If chartreuse shorts become fashionable, perhaps it is because the first person to wear the gaudy shorts was known as a person with an uncanny ability to anticipate the next fashion trend. This reason would be helpful as part of an explanation of why the trend took off, but it has nothing to do with discount rates, or the cost of shunning the norm entrepreneur. More significantly, I think that Posner's insistence on the essential arbitrariness of signaling behavior is misplaced. It blinds him to the varied reasons that people have for following norms, many of which are not related to their desire to signal discount rate. The desire to be "cool," by wearing clothes that are in fashion, a feeling of solidarity with one's fellows (the lunchtime jogging group), or simply mindless conformism are likely explanations of wearing chartreuse shorts, and are not taken into account by Posner's model.⁹³

In the following Parts, I will consider two related lines of objection to rational-choice theories of social norms. The first criticisms are "internal" to the signaling theory, in the sense that they reveal ways in which Posner's account is internally contradictory. They are like the objection raised above, to Posner's "norm entrepreneur" conjecture. Although here they are directed specifically at *Law and Social Norms*,⁹⁴ the internal objections of this section could be raised at any thin rational-choice theory of social norms that is so committed to methodological purity that it must add new epicycles to explain empirical observations within the assumptions of economic rationality, instead of admitting explanatory concepts from other sources. Part V broadens this critique by focusing on "external" criticisms, that is, objections that assume the validity of Posner's arguments (or those of any other rational-choice model of social norms) but which nevertheless show that economic accounts of social behavior should be abandoned in favor of a different kind of theory.

93. There is some ambiguity in Posner's theory, because he refers repeatedly to normative motivations that have nothing to do with signaling. See McAdams, *Signaling*, *supra* note 11, at 681-84 (summarizing these passages and arguing for a possible alternative reading of Posner). Ultimately, however, I agree with McAdams that Posner is committed to a model that, formally at least, denies that anything like moral values may be invoked as an explanation for social norms.

94. POSNER, *supra* note 19.

IV. "THE CROOKED TIMBER OF HUMANITY": INTERNAL CRITIQUES

I take the title of this section from one of Isaiah Berlin's favorite lines from Kant: "Out of the crooked timber of humanity no straight thing was ever made."⁹⁵ Berlin takes this to mean that the behavior and motivations of humans are too complex to be reduced to some kind of master value, from which all other reasons for action can be derived, or shoehorned into one kind of virtue, excellence, perfection, or form of life. The arguments of this Part take a similar tack, with respect to rational-choice explanations of social norms. In general, by flattening out the diversity and variety of reasons for action, rational-choice theories misinterpret the phenomenon of social norms. The arguments of this section are primarily internal, in the sense that they show how a particular theory fails to cohere. Part V carries this critique further, and shows that even if a rational-choice model is internally consistent, it nevertheless should be rejected in favor of one which makes use of a wider variety of explanatory terms and concepts.

The focus of this Part is predominantly Eric Posner's signaling model, which posits that social norms are the result of signaling behavior among potential transaction partners. Again, this should not be taken as a sign that Posner's account is the strongest rational-choice model of social norms. The emphasis on Posner is rather a function of his methodological commitment to explain social norms while making as few assumptions as possible about the things that people value and their reasons for action. Many of the problems addressed in this Part could be fixed by including explanatory terms (such as moral values, personal attachments and relationships, loyalties, and so on) beyond economic rationality, but Posner steadfastly refuses to make these adjustments to this theory. As a result, we are left with an object lesson in the dangers of overreliance on economic methodology. Posner's book is carefully argued and is based on a substantial body of scholarship; thus, it seems fair to conclude that the kinds of difficulties that Posner's theory must surmount would presumably face *any* rational-choice model that aspires to theoretical purity. And though it is possible to address these problems by moving away from strict rational-choice assumptions, this move is costly in terms of the advantages that law and economics scholars believe their intellectual framework offers to legal scholars. The wider implications of this dilemma are spelled out in the next Part; for now, I will look closely at some of the inconsistencies, lacunae, and unsolved puzzles in Posner's account.

A. The Normativity of Norms

Paradoxically, a strict rational-choice approach risks stripping norms of their normativity. In other words, the signaling model denies any evaluative or action-

95. See the epigraph to ISALAH BERLIN, *THE CROOKED TIMBER OF HUMANITY: CHAPTERS IN THE HISTORY OF IDEAS* xi (Henry Hardy ed., 1990). Kant, of course, was famously in need of a firm-handed editor, and was nothing like the elegant and lucid writer that Berlin was; thus, his aphorism could be rendered more literally as: "Out of timber so crooked as that from which man is made nothing entirely straight can be built."

guiding significance to what ordinary observers, not steeped in the mysteries of rational-choice theory, would regard as social norms. Posner insists on an implausibly strong definition of norms: “[S]ocial norms describe the behavioral regularities that occur in equilibrium when people use signals to show that they belong to the good type. Social norms are thus endogenous: they do not *cause* behavior, but are the labels that we attach to behavior that results from other factors.”⁹⁶ Thus, if a person brushes his hair in the morning, uses respectful forms of address, and chews with his mouth closed, we can say he is following a social norm only if he is trying to send signals that he belongs to the good type. If his explanation for his actions was instead based on considerations like “manners,” we would have to conclude that he is not following a norm (what is he doing then?). The same is true for all manner of explanations that are offered for ordinary social behaviors like refraining from littering, voting, displaying the flag on the Fourth of July, and similar examples which Posner tries to bring within his model. If a person flies the flag because she “loves her country,” Posner would have to say she is not adhering to a social norm. Finally, Posner would have a hard time accounting for norm-following when no one is present to observe the action or when the observers are not people whom the actor cares to impress. As Jon Elster notes, “I do not pick my nose when I can be observed by people on a train passing by, even if I am confident that they are all perfect strangers whom I shall never see again and who have no power to impose sanctions on me.”⁹⁷ According to Posner, Elster is not following a social norm in his example, no matter what reason he might give for refraining from picking his nose.

Posner has it exactly backwards. If someone acts intending to send a signal, she is *by definition* not following a social norm, for a norm is something that one complies with because one is supposed to, not for instrumental reasons. It is certainly possible that a person may mimic norm compliance in order to send a signal (probably a misleading signal) of that person’s reliability as a potential transactional partner, but in the great majority of cases where people adhere to norms, they must do so for reasons unconnected with signaling. Otherwise, the regular pattern of activity would not be a *norm*. The mainstream of rational-choice scholarship on social norms accepts that people feel obligated to follow norms, either because of fear of being punished by their peers, or, more commonly, through internalization of the norm.⁹⁸ Posner

96. POSNER, *supra* note 19, at 34 (emphasis in original).

97. ELSTER, *supra* note 15, at 131.

98. See McAdams, *Origin*, *supra* note 3, at 340, 350; McAdams, *Signaling*, *supra* note 11, at 679; Scott, *supra* note 11, at 1610-11. The principal proponent of internalization as the explanation for norms is Robert Cooter. See Robert D. Cooter, *Decentralized Law for a Complex Economy: The Structural Approach to Adjudicating the New Law Merchant*, 144 U. PA. L. REV. 1643 (1996) [hereinafter Cooter, *Decentralized Law*]; Robert D. Cooter, *Law and Unified Social Theory*, 22 J.L. SOC’Y 50 (1995); Robert D. Cooter, *Structural Adjudication and the New Law Merchant: A Model of Decentralized Law*, 14 INT’L REV. L. & ECON. 215 (1994) [hereinafter Cooter, *Structural Adjudication*]. Interestingly, Posner appears to have backed away from his earlier position that a norm gives a third party license to impose sanctions on a person who engages in the undesirable behavior, for example that a norm is normative. See Eric A. Posner, *Law, Economics, and Inefficient Norms*, 144 U. PA. L. REV. 1697, 1699 (1996) [hereinafter Posner, *Inefficient Norms*]. See also *id.* at 1709 (“[N]orms have moral force . . . [g]ame theory does not explain these phenomena. Explaining them requires

might respond that this argument is merely an attempt to define “norm” by fiat as behavior engaged in for noninstrumental reasons, thus begging the question against him. But this is not merely a semantic argument. The concept “norm” functions in an explanation of social behavior and is therefore subject to the principles of sound concept formation and theory acceptance that apply to any attempt to give an account of observed phenomena. To give a rough analogy, if one observes mercury rising in a column, there is a difference between two possible explanations—one given in terms of atmospheric pressure and the other in terms of the maxim that “nature abhorreth a vacuum.”⁹⁹ One explanation appeals to concepts that are admissible in the discourse community; the other does not. Scientists admit concepts like atmospheric pressure because they are useful, in light of what scientists are trying to do. Anthropomorphized pictures of nature, by contrast, must be excluded, because they violate criteria of relevance that are generated by the scientific community’s purposes. Similarly, only a soundly formed concept of a social norm may be deployed in a social science explanation.

This sounds a bit abstract, but let us try to get a handle on the term “norm.” The word refers to an action-guiding principle, the deviation from which subjects the actor to criticism in the terms appropriate to the domain.¹⁰⁰ There are moral norms (“don’t exploit people”), norms of prudence (“look both ways before crossing the street”), norms of etiquette (“if you are invited to a wedding, send a gift”), norms of civility (“don’t throw a tantrum if you don’t get your way in a faculty meeting”), aesthetic norms (“don’t paint your house pink”), formal legal norms (“drive 65 miles per hour on Interstate 81”) as well as informal legal norms that account for factors like enforcement (“you can drive 72 miles per hour on I-81 without being stopped”), norms that are generally considered aspects of tradition or custom (“take your hat off in church”), and highly context-specific norms that are incidental to particular well developed social practices (“don’t clap between movements of a symphony, but you can applaud at the end of an aria in an opera”). Violating these norms results in the appropriate kind of criticism or evaluation by one’s peers. Breaching a norm of etiquette leads to being thought of as a boor, failing to respect norms of civility results in acquiring a reputation as a jerk, violating moral norms causes one to be thought of as an evil person, and so on. Legal norms are unique, in that their violation can form the basis for state-sponsored punishment.

A norm is distinguished from a mere observed behavioral regularity *by its normativity*, by its function as a basis for evaluation by others of the action. Patterns of behavior may be normative, or they may not be. Wearing one’s hair in the prescribed manner may be normatively neutral (a colleague of mine used to have dreadlocks and no one at our university cared), or it may be highly normatively

a psychological theory.”).

99. For the latter, see BARUCH SPINOZA, *THE ETHICS* (1677), reprinted in *THE ETHICS AND SELECTED LETTERS* 42 (Samuel Shirley trans., Seymour Feldman ed., 1982). The Shirley translation of Spinoza’s *Ethics* has “there is no vacuum in Nature,” but the passage is popularly rendered in the more colorful language used in the text. See, e.g., BARTLETT’S FAMILIAR QUOTATIONS 277 (Justin Kaplan ed., 16th ed. 1992); James D. Gordon III, *Oh No! A New Bluebook!*, 90 MICH. L. REV. 1698, 1703 (1992) (book review).

100. See, e.g., *Normative*, in *THE OXFORD COMPANION TO PHILOSOPHY* 626 (Ted Honderich ed., 1995).

charged, as during the Vietnam War in the late '60s and early '70s, when wearing long hair was an unambiguous political statement and the principle "Cut your damn hair!" was normative in many American households and workplaces. One cannot know whether a pattern of behavior is normative without observing the responses of others to deviation from that standard.¹⁰¹ If we look around, we can recognize a vast array of normative principles—standards that guide our actions and subject us to specific kinds of criticism and evaluation—that are not simply conventions or more or less arbitrary focal points on which individuals' actions have converged, but are rather occasions for informal social responses when they are violated.

Significantly, if you were to ask someone why she enforced a genuine social norm, the answer would probably not be merely "because it is the way things are done"—at least the discussion would not end there. In the case of an authentic norm, it is possible to push the chain of justification back through additional steps, until the argument is grounded in a value whose significance lies in something other than mere social consensus.¹⁰² The answer to the question, "Why not litter?" would be given in terms of respect for the environment and for other people who wish to see clean roadsides and parks. Similarly, the answer to the question, "Why cut my hair?" was formerly given in terms of respect for authority and traditional values; when those sources of normativity faded during the '60s and '70s, the reasons for wearing short hair became detached from moral values. Now, wearing short hair is normative only in certain domains, like the military or business, where there are reasons of loyalty and solidarity for enforcing conformism. It is not *socially* normative to have short hair, even though short hair may be fashionable from time to time.

Strict economic rationality, by contrast, really recognizes only one action-guiding principle: "Maximize the satisfaction of your preferences." This is the only norm allowed on the playing field. Posner's theory elaborates this principle somewhat by focusing on the possibility of achieving cooperative surpluses, but it is still essentially a monistic account of action. The only motivation that is admissible in his explanatory account of human behavior is utility maximization; thus, the only criticism that could possibly be offered of a person in Posner's world is that she failed to act in a way that tended to maximize her preference satisfaction. It is simply incoherent, for Posner, to object that someone failed to follow a principle of action that would not be expected to lead to the greatest utility, as compared with the other available options. Why would an economically rational agent regard another as irrational who simply maximized her expected utility in a given situation? But if economic rationality means never criticizing someone for maximizing their expected utility, the implication is that failure to follow a social norm can never be criticized in terms appropriate to social life—that is, as a departure from standards of morality, etiquette, civility, or whatever. This is why scholars in other disciplines often object to the imperialist tendencies of economic analysis, because of this tendency to translate the explanatory terms

101. Compare Lawrence Lessig's concept of "social meaning." Lessig, *supra* note 3.

102. Compare the criticism of emotivism and existentialism (and their position that values are arbitrarily chosen) in MARY MIDGLEY, *BEAST AND MAN: THE ROOTS OF HUMAN NATURE* 177-84 (1978). Even a mainstream economist who accepts that norms may have signaling functions concedes that norms arise around actions that signify something of importance about the actor, in terms of something other than that person's discount rate. See Robert Axelrod, *An Evolutionary Approach to Norms*, 80 *AM. POL. SCI. REV.* 1095 (1986).

employed by humanistic and social science disciplines into a few concepts that are deemed admissible in economics, such as “preferences” and “expected utility.”

The question is then whether “preference-satisfaction” or some other account in terms of, say, respect, decency, civility, or ethics, is a better explanation of some observed behavioral regularity that we wish to call by the name “social norms.” In the natural sciences, we favor explanations which employ concepts like atmospheric pressure, and reject accounts such as “nature abhorreth a vacuum.” There are good reasons for doing so—concepts like barometric pressure enable us to make predictions, formulate testable hypotheses and general laws, extend knowledge from one domain to another, and intervene in the natural world when scientific knowledge is applied practically, through engineering. Economic theorists have similar aspirations: They believe that their models are empirically testable, enable predictions, and help guide the intervention by political and legal actors in social life. Critics of law and economics, therefore, must point to something that economic models cannot do, some desideratum of an explanation that is missing, in order to displace economic accounts from the field of social norms. Otherwise the argument is doomed to go round and round in a circle, with Posner defining norms as “signaling,” and his critics defining norms as “actions not motivated by utility maximization” and contending that Posner has not accounted for them. The battle of the labels will eventually take on the character of the old Loony Toons cartoon in which Bugs Bunny and Daffy Duck keep changing the signs from “duck season” to “r(w)abbit season” and back, with a befuddled Elmer Fudd looking on, unsure whom to shoot.

I believe the dispute can be resolved, but it is not as simple as saying “people don’t act like that.” They *do* act like that, at least if action is understood behavioralistically, solely in terms of observable manifestations. What is missing from the economic models is an account of the first-person phenomenology of social life—that is, the experience of being one of these individual actors described by the theoretical models. In some contexts, this is not a defect of economic theories. For example, it may be quite useful for policymakers to be able to predict behavior in the aggregate, and as long as a model enables this prediction, it doesn’t really matter whether the theory accurately captures the experience of living a life. In other situations, however, it matters critically whether the theory is radically at odds with one’s subjective experiences. In these contexts, we may have reason to reject the rational-choice model in favor of one that employs explanatory concepts that are congruent with those actually deployed by the deliberating individual.

This line of argument has implications beyond Posner’s signaling model. It is directed at the whole methodology of rational-choice theory. Of course, if an economic model of social norms admits explanatory concepts from other disciplines, as Elster’s (and to a lesser extent McAdams’s) does, then the theory is less vulnerable to this kind of attack. But, as Posner clearly recognizes, admitting terms and concepts from outside the domain of economics imperils the methodological distinctiveness of rational-choice theory, and risks alienating followers who are attracted to economic analysis for these distinguishing features. This dilemma for economic theorists will be one of the subjects of Part V, on theory acceptance.

B. Transaction Costs—Observability, Diversity, and Ambiguity

According to Posner, signals help good types find each other in a crowded marketplace, where most participants are strangers. People could match up and form transactional pairs or larger groups at random, but then of course there is a risk that the bad types will defect from the cooperative endeavor, robbing the good types of a share of the cooperative surplus. Thus, the problem for the good types is finding other good types to match up with. The result should be a “separating equilibrium,” where good types match up with good types and shun bad types. The cost of sending signals must be sufficiently low, however, because there are already means in place to ascertain whether a potential cooperative partner is reliable. Parties can order credit reports, use professional networks as information-gathering devices, read articles in the trade press about the stranger (if that person or entity is sufficiently high profile to have attracted coverage), and in extreme cases, hire private investigators to find out about the potential cooperative partner’s background. Let us call these “formal” methods of finding suitable transactional partners.

Signals appear to provide a natural way to avoid the transaction costs associated with formal methods, but notice the problem that Posner refers to as “pooling”: if signals aren’t sufficiently costly, the bad types can also send them, mimicking the signals sent by good types and thereby luring other good types into transactions, from which the bad types will eventually defect. To avoid being stuck in a pool with the bad types, the good types must send expensive signals. In order for this separating equilibrium to occur, the cost of sending signals must be high enough to prevent the bad types from aping the good types’ signals. Some of Posner’s examples, such as renting expensive office space, giving pricey gifts, or throwing lavish parties, are explicable as costly signals of reliability. As these actions become more and more expensive, however, it becomes relatively less costly in the aggregate for seekers of transactional partners to rely on formal methods of ascertaining reliability, such as obtaining credit reports. In other words, the cost of sending sufficiently expensive signals to create a separating equilibrium threatens to swallow up the savings in transaction costs promised by the signaling mechanism.¹⁰³

Because of his emphasis on reducing transaction costs, Posner tends to concentrate on relatively inexpensive behaviors as examples of signals: wearing one’s hair neatly, eating with one’s mouth closed, showing up on time for meetings, feigning interest in cocktail party chitchat, and not being hooked on drugs, to name only a few.¹⁰⁴ These are all undoubtedly social norms, but it is difficult to imagine how they could be used to avoid a pooling equilibrium caused by bad types jumping on the bandwagon. Is there really a bad type who is so impulsive or addicted to instant gratification that he could not chew with his mouth closed, use the socially approved forms of address, or stand around and pretend to care about someone’s boring story at a cocktail party, in order to lure another person into a transaction, from which the bad type would subsequently defect? Moreover, the availability of a wide variety of inexpensive signals of reliability means that good types must invest even more in

103. “There is little point in collective action if the gains are wiped out by the costs of inducing people to participate.” ELSTER, *supra* note 15, at 40.

104. POSNER, *supra* note 19, at 21-22.

signals to avoid pooling with the bad types. Suppose that many or all the bad types are sufficiently interested in matching up with the good types to invest in easy-to-emulate signals, such as keeping their hair brushed and arriving at work on time. The good types looking to stand out from this crowd are faced with the problem of taking actions that are not only more costly than doing nothing, but more costly than sending this first barrage of signals. The background noise of relatively routine actions that can also be intended as signals drives up the cost of sending unambiguous signals and, therefore, makes pooling equilibria more likely.

Paradoxically, actions intended as signals can actually *increase* transaction costs if similar actions are difficult to interpret properly. Transaction costs increase where signals are ambiguous, because the recipient of a message must expend effort decoding it. In most cases, there are “signaling reasons” and “nonsignaling reasons” for engaging in a particular pattern of behavior. Nonsignaling reasons can include habit, socialization, fear of sanctions, deference to authority, moral reasons, or simply idiosyncratic preferences.¹⁰⁵ Posner admits that some actions are sufficiently fraught with ambiguity that they do not function effectively as signals.¹⁰⁶ Furthermore, Posner acknowledges that observers must have at least approximate information about the tastes of others, so that they can filter out the actions resulting from habit, and concentrate solely on costly, observable actions that are related to the actor’s discount rate.¹⁰⁷ Even this modest concession shows the fragility of the signaling model of social norms, because so many public actions can be based on nonsignaling as well as signaling reasons. If a large number of perceived behavioral regularities can be explained by nonsignaling reasons, then it must also be true that they would not be effective signals. Posner would then be in the position of denying the status of “norm” to these regularities, even though in ordinary language these actions would be deemed norms.

Take a trivial example: My wife is no longer a practicing Roman Catholic, but upon entering a church she crosses herself reflexively. She assures me that this is a deeply ingrained involuntary response, equivalent to a behavioral tic, which conveys absolutely nothing about her religiosity or anything else of substance. At the same time, it is exactly the kind of action that might be taken by others to be a signal of a low discount rate, at least according to Posner.¹⁰⁸ (I think this sort of behavior has nothing whatsoever to do with discount rates, but that is a criticism to be developed later.) Now let a stranger, who is looking for a transaction partner, enter the picture. In order for him to know how to evaluate my wife’s crossing behavior, the stranger would have to have a rough idea about the strength of her religious commitments, so that the observer could hazard a guess as to whether my wife was crossing herself in order to send a signal, out of religious devotion, or purely out of habit. Of course, if the observer is not a stranger, and knows my wife well enough to know that her

105. *See id.* at 22-23, 135.

106. *Id.* at 27-28.

107. *See id.* Later he asserts that astute observers can distinguish between behaviors engaged in for signaling and nonsignaling reasons. *See id.* at 38. This distinction is probably based on information the observer already has about the person under observation. But as the observer builds up a more complete picture about the other person, the other person has less need to send signals, because the relevant information (discount rate) will already be known.

108. *Id.* at 18-27.

crossing has no significance, then there is no need of signaling; he probably also knows her disposition to cooperate in, or defect from, transactions. Absent personal knowledge, which would anyway make signaling superfluous, this kind of guesswork quickly becomes complicated. First, it is unclear how one is to ascertain the probability that someone is crossing herself for nonsignaling reasons. Must one have information about the religious commitments of people in a given geographic region? (Perhaps people in San Francisco are generally less religious than people in San Antonio.) An ethnic group? (Latinos as a group are more heavily Catholic than Anglos.) An age group? (Young people tend to be less committed to traditional religions than middle-aged or older people.) This little exercise in thumbnail demographic analysis can become rather time-consuming, substantially driving up the transaction costs associated with interpreting social behavior as signals.

This interpretive difficulty arises even for the paradigmatic case of signaling behavior among strangers. Imagine some kind of social event, which Abigail has attended solely for the reason of locating a transactional partner. She is looking around for a good type, and notices Benjamin listening intently to a tedious account of someone's trip to Europe. Intrigued, she sidles up to the conversation. Benjamin nods politely, interjects all of the appropriate "active listening" comments, and appears to be genuinely absorbed in the story. Abigail concludes that Benjamin may be a suitable person to do a deal with, because he is willing to put up with the cost of listening to someone prattling on and on; thus, he must also be willing to defer gratification in business transactions. Then a terrifying thought grips her—maybe Benjamin is just as boring as the other guy! Maybe there's no cost at all to listening to the story; in fact, Benjamin may be having the time of his life. Because of the presence of plausible nonsignaling reasons for Benjamin's actions, his behavior at the party really tells Abigail nothing about his suitability as a transactional partner. This is obviously an oversimplified example—no one decides to do a deal with another based solely on one observation—but even if Abigail followed Benjamin around for a week, she would still not be able to say with confidence that his socially agreeable behavior was the result of acting on signaling reasons or nonsignaling reasons. In any event, skulking around at cocktail parties eventually becomes costly for Abigail in terms of opportunity costs and out-of-pocket expenses, and at some point will become more expensive than resorting to formal methods of verifying Benjamin's reliability.

Posner is of course right that we are constantly under observation by our neighbors and associates, and that many of our actions do tend to send messages about what kind of people we are. The problem is that these messages may be garbled by the complexity and diversity of human personalities, by the crooked timber of humanity. For instance, Posner argues that people fly flags on the Fourth of July because they wish to show their neighbors that they are willing to put up with the expense and hassle of displaying the flag in order to augment their reputation.¹⁰⁹ "Because the person who displays a flag could only do so for reputational gains if he has a low discount rate, that person must belong to the good type."¹¹⁰ As Posner concedes, however, there may be different reasons for displaying the flag—nonsignaling reasons, as I have termed them. People may be genuine patriots, even jingoists, or

109. *Id.* at 115-17.

110. *Id.* at 116.

they may have a keen aesthetic appreciation for red and white stripes. Similarly, people may refrain from displaying the flag because they are bad types (because they cannot possibly make the short-term investment in time and money required to hang out the flag), or simply because they are cantankerous, and strongly resist little rituals of conformity, although they may otherwise be excellent transaction partners.

Although Posner admits the existence of nonsignaling reasons for flying flags or refusing to display them,¹¹¹ and grants that the good types may eventually move on to another kind of signal to prevent pooling equilibria,¹¹² he does not consider the extent to which *all* of the norms he uses as examples are supported by nonsignaling reasons, and are therefore subject to interpretive ambiguities. He considers norms one-by-one, not globally, and so always has a last resort available, to the effect that good types will abandon that signal and take up another. But if all signals are beset with the sort of interpretive difficulties attendant to the flag example, there may be a global pooling equilibrium, with good types unable to separate themselves from bad types by using any signal. Potential transaction partners will then abandon informal discount-rate verification procedures and adopt formal methods, causing the whole signaling house of cards to collapse. The product would be no social norms. Since we know social norms exist, the explanation must be either that Posner's signaling mechanism is the wrong account, or objections to the costliness of signals are misplaced. I admittedly do not have anything more than armchair empiricism to back up this argument, but it does seem likely that potential signals are caught between a *conceptual* rock and a hard place: Either they are inexpensive, and the bad types can easily mimic them, or they are so expensive that they gobble up the transaction cost savings realized from using informal methods of verifying discount rates.

Some of Posner's other examples cannot be explained by his signaling model because the actions in question are not within the volitional control of the actor. He posits that not being obese or refraining from speaking with a regional accent is an effective signal of one's low discount rate.¹¹³ It is certainly true that in some regions one's economic and social prospects will be diminished by the "wrong" regional accent, whether a Brooklyn accent in Mississippi or a dialect filled with "y'all's" in Philadelphia,¹¹⁴ and it is doubtless true that obese people face discrimination in employment. But accents and obesity, or their lack, are effective as signals precisely because potential transacting partners assume that they are *not* within the speaker's control. Because an accent cannot easily be assumed or cast aside, it is thought to be a reliable indicator of something relevant, such as education or social class. Speaking in a standard flat, Midwestern, Tom Brokaw-type accent would be a signal in Posner's terms only if it were possible to turn one's accent on and off, and it were somewhat costly to do so. (Perhaps it would require a period of intense concentration, like method-actor preparation, prior to any social encounter.) Other than a few talented mimics and professional actors, people generally do not have the ability to maintain an accent other than their own for an extended period of time.¹¹⁵ Similarly,

111. *See id.* at 128.

112. *See id.* at 120.

113. *See id.* at 21-22.

114. There are many wonderful moments in the movie *MY COUSIN VINNY* (20th Century Fox 1992) that play off of this kind of incongruity.

115. Even some professional actors have a hard time with accents. For instance, think of Rob

obesity for many people is a medical condition having nothing to do with willpower and self-control, which is presumably the way it is (mis)interpreted by potential transaction partners. Posner's fallback position is that any observed behavioral regularity that arises because of coincidences in preferences, budget constraints, and (one might add) physiological traits is a mere behavioral regularity, not a social norm.¹¹⁶ Because of the ambitiousness of this theoretical model, however, he fails to carry through this distinction to many of his illustrations of social norms.

The potential ambiguity of signals is heightened by social heterogeneity in the relevant marketplace. In many cases, the population of senders and receivers of signals is too diverse for certain actions to count as signals. Anthropological accounts and cultural sensitivity courses are full of stories about the divergent meaning of, say, making eye contact in various cultures.¹¹⁷ Moreover, even a relatively homogeneous culture may recognize competing values, so that a signal may be ambiguous even within that culture. For example, in Posner's discussion of why some people violate social norms, he states that "[m]ovie, pop [music], and sports stars can flout middle-class norms with impunity because their services are in such high demand."¹¹⁸ The response, of course, is that these entertainers are in high demand in part because the audience has a taste for the *outré* in addition to a desire to see conformity. Even solid, reliable bourgeois citizens like people who flout convention on occasion. That is one explanation for the popularity of movies and television shows about the Mafia, like *The Sopranos*. Posner might respond that this kind of internal diversity is irrelevant, because when someone is looking to establish a long-term, stable transactional relationship, she is not seeking someone like Dennis Rodman or Tony Soprano as a partner.¹¹⁹ To the extent this is true, it means that a person might not seek to enter into a cooperative relationship with Dennis Rodman himself, but it does not mean that the same person would not seek as a partner someone who thought Rodman was cool. But one must admit that it is possible to be a respectable member of the bourgeoisie and also to admire real or fictional public figures who enjoy transgressing middle-class values.

Another way in which signaling may not reduce transaction costs, as compared with using formal methods, is where the behavior mandated by a norm is not readily observable. Recall that in Posner's account one of the prerequisites for an action counting as a norm is its observability. Where Posner seeks to explain voting (and solve the "voting paradox") with reference to the signaling function of norms, it is fair to query him on how voting can serve as a signal of one's willingness to incur

Morrow's dreadful Boston accent in *QUIZ SHOW* (Buena Vista Pictures 1994), Tom Cruise's attempted Irish brogue in *FAR AND AWAY* (Universal Pictures 1994), or Kevin Costner's mangled (and subsequently forgotten altogether) Southern dialect in *JFK* (Warner Bros. 1991).

116. POSNER, *supra* note 19, at 41-42.

117. See, e.g., ANNE FADIMAN, *THE SPIRIT CATCHES YOU AND YOU FALL DOWN: A HMONG CHILD, HER AMERICAN DOCTORS, AND THE COLLISION OF TWO CULTURES* 65 (1997) ("Doctors could also appear disrespectful if they tried to maintain friendly eye contact (which was considered invasive), touched the head of an adult without permission (grossly insulting), or beckoned with a crooked finger (appropriate only for animals).").

118. POSNER, *supra* note 19, at 28.

119. Entering into a transaction with Tony Soprano raises a whole host of other complications.

short-term costs unless voters ostentatiously announce that they have indeed voted. In the absence of the little “I voted” stickers sometimes given out at polling places, how are strangers (who after all, are the people to whom signals of reliability are directed) to know that someone voted? Posner responds that close associates of the voter are aware of that person’s having gone to the polls,¹²⁰ but there are a couple of problems with this response. First, it is not always true. People may duck out of work for a half hour on election day, or arrive at work late, either because they voted or because they went to McDonald’s. Second, people close enough to a person to know that she voted on election day have already had abundant opportunities to observe her, and draw conclusions about whether she is a good type or a bad type. Voting is, at best, a redundant signal where close associates are involved.

Another example of norms that do not lead to observable behavior comes from a recent electronic “book” published by Stephen King.¹²¹ Essentially King offered a serialized book on a massive, collective honor system.¹²² Each reader could download a chapter from a Web site and was asked, but not required, to pay one dollar for downloading the file.¹²³ If seventy-five percent or more of readers paid the dollar, King agreed to write the subsequent chapter.¹²⁴ It worked, and the second chapter will be published on the Web site.¹²⁵ This example poses a challenge for Posner’s model, because the observed payment rate of more than seventy-five percent can surely be attributed to a social norm (“pay your fair share” or the like), but the action mandated by the norm is completely opaque to observers. When someone arranges to pay a dollar to Stephen King in the privacy of her home or office, it is difficult to explain that action based on the desire to seem reliable to one’s peers, yet it occurred in over three-fourths of the cases in which a person could either cooperate or free ride.

C. Psychological Considerations—Intentionality and Disposition

A preliminary, and relatively superficial, criticism of the signaling model is that it assumes an implausible psychology of social behavior. Posner wants to explain a wide variety of actions as the result of communicating discount rates to potential transaction partners. At the same time, he insists that actions taken for nonsignaling reasons are not taken in response to social norms, although these actions may coincide with norms if others do deliberately engage in these actions in order to send signals. His theory thus requires an element of intentionality to distinguish actions taken for nonsignaling reasons from those taken for signaling reasons. But is it reasonable to assume that people who follow the vast majority of social norms do so because they *intend* to communicate information about their discount rate? If people vote because they want their voice to be heard, if they pay Stephen King for his online book chapter out of a sense of fairness, if they give flowers to their secretaries

120. POSNER, *supra* note 19, at 122-23.

121. STEPHEN KING, THE PLANT, at <http://www.stephenking.com/download.html>, described on *All Things Considered: The Street Performer Protocol* (National Public Radio broadcast, Aug. 29, 2000).

122. *Id.*

123. *Id.*

124. *Id.*

125. *Id.*

out of a sincere sense of appreciation for hard work, or if they carry their fast food wrappers to a trash can because they wish to express respect for the environment, they are not deliberately sending signals, although of course observers would draw conclusions about these individuals' character from their behavior. It is a matter of common sense psychology that people act out of mixed motives much of the time. We give to charities not only because we feel morally obligated to help others in need, but also because it makes us feel good, and because of the positive reputational effects that accompany giving. We certainly do not mind that others see us carrying litter to a trash can, even though many people do refrain from littering even when they are not being observed. Speaking loosely and nonempirically, it is difficult to imagine a situation in which someone is following a social norm purely for signaling reasons. More plausible rational-choice theories of social norms accept that norms can operate on a subconscious level, so that the actor need not be aware of complying with them, let alone intending to act in such a way that will be perceived as a signal.¹²⁶

A more serious objection relates to the conceptual possibility of conjoining economic rationality and a long-term cooperative disposition (what Posner calls a low discount rate) in a single agent. One of the critics of David Gauthier's rational-choice theory of morality raised an extremely interesting objection, which to my mind is potentially fatal to Gauthier's project.¹²⁷ It may raise similar problems for the signaling theory of social norms. Nelson argues that we need to know more about the disposition to cooperate in a long-term relationship. It cannot be a disposition to make a *conscious* decision whether or not to defect from the cooperative solution, because if the actor were in a position to deliberate consciously, utility maximization would sometimes demand defection. In most cases the disutility from defection would be

126. See ELSTER, *supra* note 15, at 106. Elster's example is the social norm of not standing too close to someone when talking to them. *Id.* It is unlikely that anyone is ever conscious of following this norm, although we are certainly aware of departures from it, *see id.*, as the *Seinfeld* episode about the "close talker" shows. Some of Posner's examples are downright bizarre, such as his claim that gays and lesbians come out of the closet because they discount heavily the future payoffs from cooperation. POSNER, *supra* note 19, at 24. "When he stops trying to conceal his behavior, people shun him because he is showing that he does not care what they think about him, and *this can only be* because he discounts the future greatly . . ." *Id.* (emphasis added). Coming out can be motivated by the desire to be true to oneself, or to strike a blow for individual liberty, or as a public statement that homosexual attraction is not shameful. None of these reasons has anything to do with discount rate. Similarly, Posner tries to explain teenage obnoxiousness as a signal, arguing that teenagers show they are good types (with respect to other teenagers) by engaging in behavior that cuts off the possibility of interaction with adults, thereby raising the stakes associated with getting expelled from their own peer group. *Id.* at 101-02. This explanation is risible. Teenagers make themselves obnoxious to adults because they are trying on a new, separate identity, and not because they are engaging in an elaborate attempt to rig the costs of nonconformity in favor of cooperating with their peers. In addition, it is probably not too much of an exaggeration to say that teenagers aren't interested in cooperating with adults anyway, so they don't incur substantial costs, in terms of inability to satisfy their preferences, by driving adults crazy. The deeper point of these objections is that despite his assiduous attempts to avoid making assumptions about human psychology, Posner is assuming a contestable psychological thesis and not adducing any evidence for it.

127. See Nelson, *supra* note 67, at 149.

substantial—the loss of future opportunities to cooperate, for example. But a situation may come along in which the payoff from defection is so high that a rational utility maximizer would defect, even in light of the long-term reputational costs. In order to overcome the incentive to defect in that case, a person, to become a Posnerian “good type,” must bind himself so strongly that he can overcome the inclination to defect where the payoff is sufficiently high.

Because of the possibility that defection may be in the actor’s self-interest in the future, however, it would not be economically rational to adopt this strongly binding disposition never to defect. At most, an economically rational agent might try to become a “fair-weather good type,”¹²⁸ cooperating most of the time, unless the payoff from defecting were sufficiently attractive to overcome the reputational costs associated with defecting. Even if it were economically rational to do so, however, Nelson questions whether it is psychologically plausible to imagine people who are capable of essentially brainwashing themselves into cooperativeness¹²⁹ (once again, rational-choice theory cannot avoid engaging with assumptions about human psychology, even in its most methodologically pure incarnations). Finally, even assuming that one could “dispositionalize” herself never to defect when it is expedient, the fact that it is difficult to do so means that there would probably continue to be a number of fair-weather good types running around in the marketplace. The signals sent by thoroughgoing good types would therefore have to distinguish them not only from the bad types, but also from the fair-weather good types. That is a lot of pressure to put on the signaling mechanism, and it seems unlikely that social norms could function as sufficiently fine-grained signals to create separating equilibria between good types and bad types, and between fair-weather good types and thoroughgoing good types.

A final problem with Posner’s focus on signaling one’s discount rate is that he needs to assume a principle of “unity of discount rates,”¹³⁰ perhaps akin to the classical doctrine of the unity of the virtues.¹³¹ It could be the case that someone has a disposition to cooperate only in certain kinds of relationships. Someone could be a faithful spouse but a lousy business partner, or vice versa. I detest civic and social rituals and would not spend a dime to do something that was compelled by one, like fly a flag on the Fourth of July, but I like to think of myself as cooperative and reliable in other endeavors. In a famous illustration of this kind of disjunction, consider that Bill Clinton was widely reputed to be fiercely loyal to his political friends, but was of course spectacularly disloyal to his wife. To take a less dramatic example, imagine a law-abiding, morally conscientious dog lover who is also dedicated to protecting the environment, who is confronted with a local ordinance forbidding him from taking his dog on a nature trail, which he had used with his dog for years without complaint.¹³² The dog lover’s willingness to abide by or violate norms is not something that can be expressed as a single value that holds constant across the diverse commitments that are implicated in this case. Perhaps he will

128. This is my term, not Nelson’s term.

129. See Nelson, *supra* note 67, at 157.

130. See POSNER *supra* note 19, at 18-20.

131. See generally ALASDAIR MACINTYRE, *AFTER VIRTUE* 165-78 (2d ed. 1984).

132. This example is from Scott, *supra* note 11, at 1608-12.

ignore the ordinance and walk his dog on the trail, but this may say nothing whatsoever about his willingness to defect from other cooperative arrangements. It may be the case that he values his dog more than he values avoiding the ire of his neighbors, but it would be risky indeed to predict from this observation that he would cheat on his taxes, commit adultery, or engage in any other behavior which would seem to indicate a uniformly high discount rate. The reasons he values his dog more than his neighbors' approval do not necessarily carry through as reasons for defecting from other cooperative arrangements.¹³³

We must be careful here, because a theory of social norms can offer an adequate *prediction* of human behavior even if it uses terms that are not part of the deliberation of actors, or omits concepts that do function in practical reasoning. "No one thinks that baseball players consciously solve quadratic equations whenever they throw the ball, but we can nevertheless predict the path the baseball takes using quadratic equations."¹³⁴ This is true as far as it goes, but it assumes that the only useful thing to do with a theory is to generate predictions or hypotheses which can be tested empirically. But a theory may also be called upon to explain an observation, in which case it may need to include additional explanatory terms. To take a classic example from the philosophy of science, falling barometric pressure can *predict* the coming of a storm, but the mercury in the barometer does not *explain* the storm—both the dropping barometric pressure and the storm are explained by a common cause.¹³⁵ The following Part considers how criteria of theory acceptance can help sort out the debate between rational-choice theorists, who emphasize methodological parsimoniousness and empirical testability, and their critics, who accept a wider variety of explanatory terms, at the price of a certain amount of simplicity and theoretical elegance.

V. "ADVANCED CAT MUSIC": EXTERNAL CRITIQUES

Human behavior is simply not fully explicable in terms that correspond to the assumptions of economic rationality. I realize objections like this to the law and economics project are made constantly, and on the whole, rational-choice scholars have not found them persuasive. Perhaps with the benefit of hindsight, the rational-choice theorists will be judged to have had the better arguments. In that case, my protests, and those of others, will sound to future readers like the complaints of a contemporary music critic who referred to *Tristan und Isolde* as "advanced cat music."¹³⁶ Maybe we will all be rational-choicers someday, just as Western musicians

133. For this reason, McAdams suggests as a modification to Posner's theory that people can be understood as signaling "values, commitments, or beliefs" to one another, and not just discount rates. McAdams, *Signaling*, *supra* note 11, at 685-86. That certainly seems like a sensible modification to me, although it violates Posner's principle of strict methodological purity. See POSNER, *supra* note 19, at 18. Posner insists on thin sets of explanatory concepts because the resort to values and beliefs as terms in an explanatory argument risks the distinctiveness of economic models of social norms. See *id.* at 19.

134. BAIRD, *supra* note 68, at 125.

135. See, e.g., SALMON, CAUSALITY, *supra* note 13, at 111.

136. NICOLAS SLONIMSKY, LEXICON OF MUSICAL INVECTIVE 235 (W.W. Norton & Company 2000) (1953). The allusion here is to the aleatory notes produced by a cat walking down a

eventually accepted and then embraced Wagner's harmonic innovations. Moreover, as Eric Posner notes, competitors to rational-choice theory have not been as influential in the legal academy as the law and economics crowd.¹³⁷ But I don't think critics of rational-choice theory need to worry about appearing to be dinosaurs in the legal academy. Simply put, the amount of insight that can be squeezed out of a theory which treats morality, emotions, commitments, and departures from ideals of economic rationality as peripheral concerns is limited, and this limitation is imposed not by naysayers in the academy but by the nature of legal reasoning itself.

Posner is aware that his stripped-down conception of rationality and social interaction, in which all social norms can be explained as signals of discount rate, is bound to be criticized as grossly mischaracterizing the reasons for human behavior. But he thinks he can exclude considerations like commitments, values, and emotions from his model and still produce a useful theory:

Readers will object that racial discrimination, patriotism, ceremonial gift-giving, clothing fashions, and other complex social phenomena . . . involve more than the efforts of members of a group to signal to each other that they belong to the good type. . . . My response is that this book reflects a methodological commitment. My claim is that rational choice can shed light on social norms by focusing on the reputational source of behavioral regularities to the exclusion of their cognitive and emotional sources. I do not claim that rational choice theory can offer a complete explanation of social norms or of cooperation. Cognition and emotion are not irrelevant. They are just not well enough understood by psychologists to support a theory of social norms, and repeated but puzzled acknowledgments of their importance would muddy the exposition of the argument without providing any offsetting benefits.¹³⁸

I think Posner is wrong in this passage, but wrong in an interesting way, for he at least makes clear the terms on which we should accept or reject a rational-choice explanation of social norms. He acknowledges that a theory can move from one that is highly parsimonious in its assumptions to one which is richer, more detailed, but potentially messier (Posner explicitly worries about "muddy[ing] the exposition of the argument"). The question is why we should prefer one to the other. That is the question I wish to take up in this Part.

Imagine a spectrum of explanations of some domain of social behavior, from the

piano keyboard, although other critics have employed differing feline similes in attacking composers they disliked, referring for example to amorous cats yowling in the night, *id.* at 117 (quoting an article critical of Liszt in the London newspaper ERA, Feb. 25, 1882), "crazy cat music," *id.* at 150 (quoting an article by Hugo Leichtentritt critical of Schoenberg in Berlin's SIGNALE, Feb. 7, 1921), and my personal favorite, a cat with catarrh, *id.* at 249 (quoting an article critical of Webern in the BOSTON EVENING TRANSCRIPT, Apr. 17, 1913).

137. POSNER, *supra* note 19, at 38-39.

138. *Id.* at 46; see also Scott, *supra* note 11, at 1604-05 ("[T]he analyst might continue to treat values, moral character, and preferences as exogenous, not because these phenomena are unimportant, but because her analytical tools don't allow her to say anything systematic about them.").

most “thin” or parsimonious in terms of explanatory terms or concepts, adding additional terms and concepts through increasing degrees of “thickness,” or what Posner calls muddiness:

Thinnest |— (1) ——— (2) ——— (3) ——— (4) ——— (5) —|Thickest

On this continuum, (1) is the position occupied by Posner’s theory of social norms, for it posits no explanatory concepts other than economic rationality (self-interested preference satisfaction) in its explanation of social behavior. Position (2) represents a thicker set of explanatory terms, but is still recognizably in the rational-choice camp. Richard McAdams’s model of social norms fits here, as do other economic accounts such as Robert Ellickson’s. Jon Elster’s work fits in position (3), right on the boundary of what can still be considered rational-choice theory. Positions (4) and (5) are out of the rational-choice world entirely, and belong to other disciplines such as sociology, psychology, literature, or philosophy.

To underscore the thesis of this Article, any theory that lawyers, judges, or legal scholars use to illuminate human behavior must make use of the same conceptual terms that are relevant to understanding the authority of law. To fail to make this linkage consigns a theory to irrelevance, vis-à-vis the legal system. It may be the product of a fascinating intellectual exercise, like proving Fermat’s last theorem, or it may produce testable hypotheses, but it is worthless from the standpoint of planning legal intervention in human affairs. A methodologically thin rational-choice theory, in the neighborhood of (1) on the above continuum, simply fails to connect with the concerns that animate legal actors such as judges and academic critics of the law, who reason about the legitimacy and justice of legal rules based on values far more multifarious than economic rationality. The remainder of this Article is an argument that any theory of social behavior that purports to be relevant to law must be constructed using a sufficiently thick set of explanatory terms.

A. What Is Involved In Providing an Explanation?

Consider a number of examples of observed phenomena to be explained by some sort of theoretical analysis:

- The mercury in a barometer falls before a storm hits the area.
- Planets follow an elliptical orbit, with the sun at one of the foci.
- Two days after being exposed to someone with the measles, a person exhibits symptoms of measles.
- People consistently express attitudes of disapproval toward certain actions, such as adultery, racial discrimination, or hitting.
- In many circumstances people tend to behave cooperatively even in the absence of formal legal sanctions.

If we are asked to supply a *theory* to *explain* these observations, it should be clear

that not just any account will do. The story we tell must take a certain form, appeal to particular kinds of considerations, provide valuable insight, and be linked with wider concerns (such as verifiability, knowability, and criteria of truth) in some way. Philosophers of the natural and social sciences have made tremendous progress in clarifying the form that an adequate theoretical explanation must take, but this work has not yet had much of an impact on the debate between economic theorists and their critics. I am hopeful that this situation will be rectified and offer this Article as the beginning of that process, because these principles of theory acceptance promise the best way out of this seemingly intractable debate.

One kind of explanation common in the natural sciences is a *covering-law* account, in which “with the aid of suitable initial conditions, an event is explained by subsuming it under one or more laws of nature.”¹³⁹ So, for instance, Kepler’s laws of planetary motion can be explained by subsuming them under Newton’s laws of physics. The form of explanation is deductive, meaning that as a matter of logical form, the observation will follow necessarily from the natural law and suitable initial conditions.¹⁴⁰ It is critical that the explanation include at least one law of general applicability in order to link the event to be explained with regularities that occur in nature.¹⁴¹ The general law introduced as part of the explanation must satisfy several additional criteria. First, it must have implications that can be tested empirically. Natural laws generally support counterfactual conditionals—that is, they can be adduced in support of a statement like, “If this substance had been water it would have boiled at 100 degrees Celsius.”¹⁴² One of the most powerful reasons for accepting a covering-law model is that it enables prediction, and prediction requires an explanation to accommodate counterfactuals.¹⁴³ Second, a law of nature must have

139. SALMON, CAUSALITY, *supra* note 13, at 51 (emphasis omitted). Hempel is the most influential exponent of this model. See, e.g., CARL G. HEMPEL, PHILOSOPHY OF NATURAL SCIENCE 49-54 (1966); Carl G. Hempel, *The Logic of Functional Analysis*, in READINGS IN THE PHILOSOPHY OF SOCIAL SCIENCE, *supra* note 13, at 349.

140. See SALMON, CAUSALITY, *supra* note 13, at 303.

141. See *id.* at 334. Jon Elster criticizes the lack of a generally applicable law in the explanatory argument offered by some critics of rational-choice theory. Jon Elster, *Introduction*, in RATIONAL CHOICE 24-26 (Jon Elster ed., 1986). Non-rational-choice explanations provide only a “thick description.” *Id.* at 26. Although I am advocating a thicker explanation of social norms, I do agree with Elster that a theory cannot simply be a richly detailed anecdotal account of an event. See also ELLICKSON, ORDER WITHOUT LAW, *supra* note 3, at 7 (criticizing Clifford Geertz). To be fair to Geertz, who is largely responsible for the popularity of the term “thick description” in the social sciences, his interpretations of culture do employ general principles, albeit not stated in terms of natural kinds or natural laws. Geertz’s explanatory concepts are things like beliefs, desires, values, intentions, and purposes—in short, concepts that give *meaning* to human activity. These concepts are not readily susceptible to testing and falsification, however, which is why Elster, Ellickson, and others regard him as insufficiently scientific.

142. See HEMPEL, *supra* note 139, at 56; Brian Fay & J. Donald Moon, *What Would an Adequate Philosophy of Social Science Look Like?*, in READINGS IN THE PHILOSOPHY OF SOCIAL SCIENCE, *supra* note 13, at 21, 26-28.

143. One frequently encounters the argument that parsimonious models, like rational-choice theory, are to be preferred over theories that invoke a greater variety of explanatory terms. The thin theories are better as grounds for making predictions than their rivals. See, e.g., Scott,

explanatory relevance, in the sense that the general law offers good grounds for believing that the event did, or will occur. As a Nobel Prize winning physicist and noted commentator on science put it, the general law must account for what is going on in the minds of scientists when they say, "Aha!"¹⁴⁴ This is an informal description of what is involved in providing an explanation that is *relevant* to the question asked. The criterion of explanatory relevance is the key, I think, to getting beyond the interminable metatheoretical debates over law and economics.

With humans, things get quite a bit more complicated. For one thing, human behavior tends to be susceptible only to statistical generalizations, as opposed to deductive explanations. In a scientific explanation using statistical laws, the conclusion cannot be inferred deductively; rather, the conclusion is supported inductively, and the support given the conclusion varies in strength according to the probability of the event.¹⁴⁵ To take a frequently discussed example, one can explain Jane Doe's measles symptoms as a result of Jane's having been exposed to measles, but the explanation is not deductive, because not all persons exposed to measles will contract the disease.¹⁴⁶ The tricky part in constructing an inductive explanation of some observed event is including all of the relevant evidence as part of the explanation, while not loading up the premises of the argument so that the explanation ends up being trivial or circular.¹⁴⁷ The premises in the argument that explain Jane's case of the measles can include the fact of her exposure to someone who was suffering from the measles, and the general law that there is a high probability that someone exposed to the measles will catch the disease, but not the observation that Jane caught the measles.

Another significant difference between explanations in the natural and social sciences is the necessity, in the latter, of accounting for the intentional, meaningful, or semantic dimension of behavior, as it has been variously described.¹⁴⁸ Imagine that we observe a person deviate from her normal route to work, stop and go into a building, step inside a partitioned booth and pull a lever, and then return to work. It is impossible to explain this behavior without referring to certain concepts, such as intentions, rules, institutional structures, beliefs, dispositions, expectations, duties, and values. The observed behavior has *meaning* only in light of social conventions, institutions, and rules. It is true that materialist and behaviorist theories have been constructed that attempt to dispense with these categories. For example, it may be possible to give an explanation of the action by translating these concepts into dispositional terms, by correlating certain stimuli with observed responses. These attempts are widely acknowledged to have failed, however. The purportedly neutral data turn out to be suffused with the categories that behavioral theories seek to exclude. To test a behaviorist theory, an observer would have to see whether a

supra note 11, at 1607.

144. Steven Weinberg, *Can Science Explain Everything? Anything?*, N.Y. REV. BOOKS, May 31, 2001, at 47.

145. See, e.g., SALMON, CAUSALITY, *supra* note 13, at 304.

146. See HEMPEL, *supra* note 139, at 58-59.

147. See SALMON, CAUSALITY, *supra* note 13, at 305.

148. This paragraph is drawn from the discussion in Brian Fay & J. Donald Moon, *What Would an Adequate Philosophy of Social Science Look Like?*, in READINGS IN THE PHILOSOPHY OF SOCIAL SCIENCE, *supra* note 13, at 21, 22-25.

person performs correctly under similar conditions—in other words, whether she votes on the next designated election day. But the selection of a particular day (“election day”) for observation and the criteria for evaluating correct performances (“person goes to a polling place”) make sense only because the observer already has in mind a structure of institutions, rules, and expectations. It becomes impossible to eliminate intentional or semiotic concepts from an explanation without rendering the explanation completely unintelligible.

To relate this somewhat abstract discussion to the problem of social norms, many critics from the rational-choice school have observed that interpretive explanations of behavioral regularities fall into circularity by including the observation to be explained (“there is a social norm against littering”) as a premise in the explanatory argument, which is supposed to explain why people do not litter. Certainly the explanation of the existence of or conformity to a social norm must be inductive, not deductive, in its logical form, because compliance with the norm is only probabilistic—there is always some number of people who do not follow particular norms. In addition, methodological purity in rational-choice theory demands that the explanation include only the minimal assumptions of economic rationality as utility maximization, and exogenously given preferences and utility functions. The assumptions of rational-choice theory are constructed to avoid importing interpretive categories into the model, to eliminate the value-laden categories that the theorist is concerned with explaining. The danger of restricting the explanation to these kinds of considerations, however, is the baroque complexity of Posner’s signaling account,¹⁴⁹ which is internally consistent, and avoids begging the question, but which seems jarringly discordant with other widely shared assumptions about human behavior.¹⁵⁰ The metatheoretical issue concerns the specification of explanatory considerations that are invoked by the theory. This brings us back to the criterion of explanatory relevance.

In a whimsical example from Wesley Salmon, suppose a man takes oral contraceptives regularly for a year and explains his failure to get pregnant on the basis of having taken the pills.¹⁵¹ His explanation is perfectly satisfactory as a matter of *form*—it invokes a general law regarding the efficacy of oral contraceptives—yet utterly fails as a genuine account of why he did not become pregnant. The explanation fails because the explanatory terms introduced are irrelevant to the nonpregnancy of a man—the facts do not make a difference to the occurrence of the fact to be explained. More serious examples may of course be offered: Salmon used as an illustration the disappearance of certain neurotic symptoms after a course of psychotherapy; in that example, the explanatory terms include premises about the effectiveness of psychotherapy, which are hotly contested in the relevant scientific and medical communities¹⁵² (for example, the spontaneous remission rate from neurosis may be quite high). Or, if we seek to explain why a teenager went on a shooting rampage, we may offer as relatively uncontroversial explanatory terms such

149. POSNER, *supra* note 19, at 18-27.

150. *See id.* at 39-46.

151. *See* WESLEY C. SALMON, *STATISTICAL EXPLANATION AND STATISTICAL RELEVANCE* 34 (1971).

152. *See* SALMON, *CAUSALITY*, *supra* note 13, at 310.

facts as the shooter's age, sex, upbringing, mental health, and grudges he may have borne against his classmates. On the other hand, whether additional terms may be adduced, such as his television-watching habits, or the nature of music he listened to or the video games he played, is a matter of intense debate. The point of these examples is that the nature of the observation we are seeking to explain puts some constraints on the kind of concepts we can introduce in the explanation. Psychotherapy may be offered as an explanation of the disappearance of neurosis, and video game playing may appropriately be included in an explanation of a school shooting, only if these factors are relevant to the observed event.

Elaborating what is meant by the term "relevance" turns out to be more difficult than it first appears. Salmon argues that explanatory relevance is essentially a pragmatic consideration.¹⁵³ For any event there is an "ideal explanatory text"—a hypothetical detailed account of everything that could have contributed to the event—which is objective in the sense of being correct in light of the known (or knowable, if one wishes to posit an ideal intelligence) facts about the causal structure of the universe. But we seldom have occasion to explain an event with reference to the complete ideal explanatory text. Rather, we ask "why?" questions that call upon aspects of the ideal explanation. The resulting partial explanation must reflect the ideal explanation,¹⁵⁴ but must additionally "honor the salience of the information it includes."¹⁵⁵

A considerable amount of debate has occurred in the philosophy of the social sciences, over what kind of information should be included in an explanation of human behavior. One principle that is frequently invoked is Ockham's razor, which states that a simpler theory should be preferred to a more complex one with the same explanatory power.¹⁵⁶ Another relevant pragmatic criterion is the probability that an explanatory factor will distinguish an observed event from its *contrast class*,¹⁵⁷ or the set of events that did not occur. This consideration is pragmatic because the contrast class is specified by the type of "why question" that is asked. For example, suppose we ask why a particular teenager named Albert stole a car.¹⁵⁸ The ideal explanatory

153. *See id.* at 315 (discussing Peter Railton).

154. By "reflect," Salmon means that it cannot appeal to considerations that are ruled out by our best scientific understanding of causal relationships; thus, an explanation of the death of John F. Kennedy that appealed to the position of stars and planets would not reflect the ideal explanation, but an explanation in terms of the CIA or the Mafia or the Cuban government might. *See id.* at 184. The last clause is my interpolation. Salmon never registers dissent from the Warren Commission's conclusions.

155. *Id.* at 315.

156. *See, e.g.,* GORDON, *supra* note 13, at 602-03; *William of Ockham*, in 8 *ENCYCLOPEDIA OF PHILOSOPHY* 306, 306 (Paul Edwards ed., 1967). As Gordon points out, very little is generally given in the way of argument for Ockham's razor. Perhaps simplicity is an aesthetic criterion. *Cf.* ISAIAH BERLIN, *THE ROOTS OF ROMANTICISM* 24-25 (Henry Hardy ed., 1999) (describing simplicity and conceptual unification as a matter of faith, or an ideal of the Enlightenment). Absent some more compelling reason to believe that it has some connection to the probability that a given theory will be true, Ockham's razor can be no more than a pragmatic reason for preferring simpler theories.

157. *See* SALMON, *CAUSALITY*, *supra* note 13, at 182-85 (responding to BAS VANFRAASSEN, *THE SCIENTIFIC IMAGE* (1980)).

158. *See id.* at 367-68 (discussing an example from James G. Greeno, *Explanation and*

text would include such factors as Albert's age, sex, upbringing, socioeconomic class, and so on, but notice that the answer given—the subset of the ideal explanatory text to which the explanation appeals—varies according to the information sought by the questioner:

Q: Why did *Albert* steal the car?

A: Because among all his fellow gang members, Albert was the most proficient at hot-wiring.

Q: Why did Albert steal the *car*?

A: Because it's harder to fence a television.

These explanations vary because the contrast classes vary. We are asking different questions, as indicated by the differences in emphasis above. The presuppositions standing behind the “why questions” make all the difference to the suitability of the offered explanation.

Philosophers who study social scientific explanations are attempting to construct a metatheory—that is, a theoretical account of the kind of theories we should expect from observers of human societies. These metatheories assume that an observer is trying to do something besides make predictions, although predicting is certainly a useful thing to do with a theory. Take a simple example from the natural sciences: the correlation between meteorological phenomena and the readings on instruments like barometers. If one's job is to evacuate coastal areas before a hurricane, one's conceptual toolkit need include only things that enable accurate predictions, such as radar data and barometric pressure readings. There would be a sense, however, in which a weather forecaster would not *understand* the hurricane, even though she could predict its occurrence with great accuracy. Similarly, rational-choice theorists emphasize the utility of their models in the predictive dimension, as if that were all a theory were supposed to do. Understanding is not the same thing as making empirically testable predictions, however, in the social as well as the natural sciences.¹⁵⁹

Understanding human behavior is usually taken to involve grappling with the reasons why people act in a given way. Reasons are a phenomenon internal to individuals—they generally cannot be simply “read off” an observation, like measuring the temperature of a chemical reaction. As Thomas Nagel puts it, even from the external point of view, we realize that evaluation, explanation, and understanding require engagement with the person's own subjective point of view:

What we see, unless we are artificially blind, is not just people being moved to act by their desires, but people acting and forming intentions and desires for reasons, good or bad. That is, we recognize their reasons *as reasons*—or perhaps we think they are bad reasons—but in any case we do not drop out of the evaluative mode as soon as we leave the

Information, in STATISTICAL EXPLANATION AND STATISTICAL RELEVANCE, *supra* note 151, at 89, 89-104).

159. *See id.* at 338-39.

subjective standpoint.¹⁶⁰

The obsession that many rational-choice theorists have with the natural sciences causes them to wish to eliminate the internal point of view from their theories. Again, however, we must be attentive to the reasons why we want to have a theory in the first place. As the example of Albert makes clear, an explanation of human behavior must be evaluated in the light of the presuppositions that underlie the question asked.

Because why questions in social science generally seek understanding, not merely prediction, many theorists of the social sciences argue that an explanation of human behavior must be interpretive, that is, concerned with questions of *meaning*.¹⁶¹ People are “up to something” when they act, and the “something” is the meaningful dimension of human life. Religious observances, gift-giving behavior, patriotic displays, and historic phenomena such as duels, lynchings, and tarring and feathering, can all be explained in terms of their symbolic significance. Clifford Geertz is brilliant at this kind of analysis, taking raw ethnographic data about, say, a sham sheep raid in Morocco and giving it meaning in terms of authority and status, offense and honor, ethnic conflict, and challenges to claims of political hegemony.¹⁶² In his story, a Jewish trader, Cohen, and his friends were attacked by neighboring Berber tribesmen; the friends were killed and Cohen’s sheep were taken.¹⁶³ Cohen complained to the local French colonial authorities and demanded his ‘*ar*—four or five times the value of his sheep, which he would be owed as recompense for the theft under the traditional trade-pact system. The French refused to comply, because they were attempting to displace the traditional rules and replace them with colonial government. Cohen, however, rallied the sheikh of the local tribe and rode off to steal sheep from the thieves. The thieving tribe, recognizing that the sheikh and Cohen were asserting their ‘*ar*, relented, and permitted Cohen to select five hundred sheep as payment. The French, seeing Cohen riding home with the sheep, locked him in jail and took his sheep.

These are the raw data of social science observations. What does it mean to provide a theoretical account or explanation of the event? Just as the physicist is trying to capture what is going on in his head when he thinks “Aha!” after hearing a theoretical explanation of some experimental data, the interpretive anthropologist or sociologist

160. THOMAS NAGEL, *THE VIEW FROM NOWHERE* 142 (1986) (emphasis in original).

161. See, e.g., JAMES BOHMAN, *NEW PHILOSOPHY OF SOCIAL SCIENCE: PROBLEMS OF INDETERMINACY* 102-45 (1991); GEERTZ, *supra* note 1; ROSENBERG, *supra* note 13, at 19; MICHAEL WALZER, *INTERPRETATION AND SOCIAL CRITICISM* (1987); Charles Taylor, *Interpretation and the Sciences of Man*, in *READINGS IN THE PHILOSOPHY OF SOCIAL SCIENCE*, *supra* note 13, at 181; Owen M. Fiss, *Objectivity and Interpretation*, 34 *STAN. L. REV.* 739, 739 (1982). There is obviously a dizzying variety of interpretive explanations of human behavior, each with its own specialized vocabulary, presuppositions, and methodology. It is impossible within the confines of a (relatively) brief article to do justice to phenomenological, structuralist, poststructuralist, semiotic, hermeneutic, postmodern, Marxist, psychoanalytic, feminist, and other interpretive perspectives. I concentrate on Geertz’s example because his essay is widely known, lively and engaging, and representative of arguments that are offered for the interpretive stance.

162. See GEERTZ, *supra* note 1, at 18-19.

163. See *id.* at 7-9.

is aimed at “figuring out what all that rigamarole with the sheep is about”¹⁶⁴ when given a description of an elaborately staged sheep-stealing expedition. Explaining the rigamarole with the sheep is not just a matter of identifying causal relationships. At the very least, the causal structure of the sheep raid is relatively unimportant, as compared with questions of meaning and interpretation. If someone asks what’s going on with the sheep raid, it’s unlikely to be out of idle curiosity. In Salmon’s pragmatic terms, the properly formulated “why?” question, in light of the relevant contrast class, is not “Why did Cohen go with the sheikh and raid the neighboring tribe’s sheep herds, rather than staying home to watch television?” Rather, it is something like, “Why, in light of the abolition of the traditional trade-pact structure by the French, did Cohen resort to self-help, instead of accepting that there was no remedy under the French authorities?” Answering the properly formulated why-question requires an understanding of the reasons that make sense to the trader, the Berbers, and the French—considerations like claims owing as a matter of honor (the concept of *‘ar*), traditional norms for resolving conflicts, the respect (or lack thereof) owed to the colonial authorities and the sheikh, and the challenge posed to the French by the traditional trade-pact system. Without these terms, the explanation fails to make sense of the reasons that actually motivated Cohen, the sheikh, and the French. If, for example, a theorist tried to explain the sheep raid in terms of Cohen’s desire to signal that he had a low discount rate, and would therefore be a good transactional partner, the explanation would miss the mark because of its failure to take into account the concept of *‘ar*. Because it is *‘ar*, not signaling, which actually motivated Cohen, the signaling explanation would have selected the wrong portion of the ideal explanatory text.

This is not to deny that Cohen’s actions would have the effect of signaling his willingness to bear short-term costs (the possibility of getting killed in the sheep raid) in order to realize long-term gains (the stability provided by relationships protected by the trade pact). Pragmatically speaking, however, the best account of the sheep raid, and the relevant portion of the ideal explanatory text, is that which makes use of the concepts that make sense to the participants in the story, from the first-person point of view. It is only by taking that point of view that we can actually understand what Cohen was up to when he went riding out with the sheikh to enforce his rights under the trade pact. Ironically, adopting this point of view will enhance the ability of the theory to generate accurate predictions about future behavior, since it is more finely calibrated to the factors that are likely to motivate people to act.¹⁶⁵

Another way to formulate why-questions concerning the sheep raid is in a normative or interventionist spirit. We could ask what went on with the sheep raid in order to understand political and ethnic conflicts in Moroccan society so that we can do something with that knowledge, such as reduce violence or intergroup tensions. It may be a bit grandiose to say that the social sciences are aimed at “improving the human condition,”¹⁶⁶ but all theories of social norms that are offered as part of legal discourse are offered as premises in an argument, the conclusion of which generally takes the form: “Therefore some legal actor should (or should not) intervene in the

164. *Id.* at 18.

165. I am grateful to David Millon for this point.

166. ROSENBERG, *supra* note 13, at 21.

world through the means of a legal rule that provides such and such.” In order for these legal rules to be sensitive to the way they will affect behavior in the real world, I contend, the theory that one uses to describe social behavior must take account of the reasons for action that are actually employed by the people in question. For example, in interpersonal relationships we don’t think about “how to treat people” by strategizing signals that will communicate one’s suitability as a transacting partner. We think about how to treat people in terms of tolerance, respect, dignity, rights, fairness, loyalties, duties, and emotions such as affection, kindness, or loathing. People do not refrain from cheating on their spouses so they can signal a low discount rate and thereby preserve marital surplus; they do so because they are motivated by love, loyalty, or their commitment to their expressed vow of fidelity. These terms are required to understand action, as opposed to simply making predictions about when something will occur.

Similarly, to the extent that social science theories are deployed as premises in policy arguments, they must make use of explanatory concepts that go beyond concepts that are useful in making predictions into terms that have normative significance. This is the difference between using quadratic equations to predict the path of a baseball, and using methodologically thin economic theories to predict human behavior. Both theories may have some predictive value, but in the case of a theory of human behavior, the theorist also seeks to use the model to buttress an institutional-design argument. Of course, a theorist may care about nothing other than prediction, in which case many of the theory-acceptance principles in this section have no bite. In that case, however, the argument will be of little interest to policymakers and legal actors. In a discourse community that is concerned about legal responses to social problems, purely predictive theories can be set to one side, in favor of theories that have evaluative conclusions.

B. The Appeal of Economic Models

We have already observed the tendency of some rational-choice theorists, notably Eric Posner, to start off with relatively uncontroversial models of social norms, that could appeal to both law and economics and law and society sympathizers. Posner initially admitted a variety of motivational factors into his account, and conceded that interests other than utility maximization could account for the origin of norms. As his thinking developed, however, Posner has moved increasingly in the direction of methodological purity, at the cost, I think, of the plausibility of his theory. One might wonder, therefore, about the attraction of the methods of rational-choice theory for legal scholars. There are two principal categories of reasons why economic methodology is appealing to social theorists, although I think neither is sufficiently attractive, normatively speaking, to justify wholesale acceptance of rational-choice theories of social norms.

1. The Ideal of Freedom and the Rejection of Authority

One of the most enduring and powerful ideological underpinnings of rational-choice theory is political liberalism—the notion that the modern subject is free to determine for herself what she values, “without interference from external authority”:

"Each person is the best judge of his own happiness."¹⁶⁷ As David Gauthier puts it, each person has the liberty of a Robinson Crusoe.¹⁶⁸ If a person likes to eat Taco Bell and watch Jerry Springer, that is his prerogative, and the state should not interfere with these choices, unless they cause some kind of harm to others. This position is sometimes stated more strongly, as the claim that there is no source of value beyond a person having chosen something as valuable.¹⁶⁹ In contrast with classical and medieval thinkers, many contemporary ethical theorists hold that there is nothing like God's will, the Form of the Good, human nature, or natural law, that makes choices good or bad. All that is important is securing (often hypothetical) agreement, by settling on procedural principles that can be justified to one another or by analyzing the structure of discourse itself to construct a minimal set of regulative criteria that must be complied with in order for an agreement to count as a principle of ethics.¹⁷⁰ In all of these systems of ethics that can be called proceduralist, the emphasis is not on "getting it right" as it was for Plato; rather, excellence in ethical reasoning is defined in terms of following a sound method or style of thought.¹⁷¹

Rational-choice theory is in some ways the ultimate procedural ethic, but it does contain (sometimes covert) substantive moral principles. Applied to law, economic reasoning tells state actors to maximize wealth or efficiency (often defined in terms of wealth). Because the ideal of autonomy is so pervasive as a leitmotiv of modern thinking, however, some economic theorists have tried to make the concept of autonomy do work that it is plainly not suited to do: for example, notwithstanding the obvious affinity between law and economics and utilitarianism, Richard Posner has tried to analogize the Paretian criterion of efficiency to Kantian ethics which, in Posner's reading, locates the highest value in individual autonomy.¹⁷² If even so committed a consequentialist as Posner is trying to capture the Kantian flag, it is a clear indication that autonomy is a powerful cultural value. But one must be careful with the concept of autonomy. For example, Posner misreads Kant in this respect, because autonomy alone has no intrinsic value for Kant; what is essential to his ethics is respect for the moral law.¹⁷³ This is not exegetical nitpicking—it is absolutely central to Kant's position. For Kant, it is possible to be right or wrong about a

167. CHARLES TAYLOR, *SOURCES OF THE SELF: THE MAKING OF THE MODERN IDENTITY* 82 (1989).

168. GAUTHIER, *supra* note 70, at 91.

169. See, e.g., MIDGLEY, *supra* note 102, at 177-84 (criticizing existentialism); ROBERTO MANGABEIRA UNGER, *KNOWLEDGE AND POLITICS* 67-69 (1975) (restating a premise of "liberal political theory" with which he disagrees).

170. See, e.g., GAUTHIER, *supra* note 70; 1 & 2 JÜRGEN HABERMAS, *THE THEORY OF COMMUNICATIVE ACTION* (1984); STUART HAMPSHIRE, *INNOCENCE AND EXPERIENCE* (1989); RAWLS, *supra* note 71; SCANLON, *supra* note 43. Although these writers differ greatly in the details of their theories, they are all broadly *contractarian* in their approach to ethics. Antecedents of contractarianism in modern moral and political philosophy include Hobbes and Locke.

171. TAYLOR, *supra* note 167, at 85-86.

172. See Richard A. Posner, *The Ethical and Political Basis of the Efficiency Norm in Common Law Adjudication*, 8 *HOFSTRA L. REV.* 487 (1980).

173. See CHRISTINE M. KORSGAARD, *THE SOURCES OF NORMATIVITY* 97-98 (Onora O'Neill ed., 1996).

principle of ethics. Values are not merely the projection of our subjectivity onto a value neutral world. Rather, principles of pure reason determine the moral law, and one acts autonomously by acting pursuant to that law.

By contrast, the rational-choice conception of value holds that values are nothing more than a name for preferences, which an individual adopts arbitrarily. This insulation from the grounds for adopting preferences means that economic theories of anything, including social norms, cannot take into account these reasons. Economic theories of social behavior suffer from the same problem that afflicts emotive theories of ethics. It is of little help to say that a negative moral evaluation reflects an observer's judgment of disapproval. We are interested in *why* the observer expresses disapproval—the attitude standing alone does not suffice to provide an explanation of this attitude.¹⁷⁴ This explanation requires deeper digging into the grounds of the judgment of disapproval, that is, into reasons and values. Similarly, to the extent that the basis for expressed preferences is essential to an explanation of some aspect of human behavior, rational-choice theory will be unable to use them as explanatory concepts.

2. Science Envy and Value Neutrality

The humanities and social science disciplines have long labored with something of an inferiority complex, vis-à-vis the natural sciences.¹⁷⁵ Moral philosophy, for example, is still struggling to liberate itself from the legacy of logical positivism, which held that only empirically verifiable statements could be regarded as truth-functional, leaving evaluative judgments to be explained as simply expressions of the observer's attitude.¹⁷⁶ Going in the opposite direction, the discipline of political *science* has tried to purge itself of the taint of political *philosophy*, that is, of any form of knowledge that cannot be verified by the study of observable facts.¹⁷⁷ Similarly, value neutrality has at times seemed to be the Holy Grail after which legal theorists are questing.¹⁷⁸ Positivism in legal philosophy is the claim that legal rules are at least analytically separable from moral principles, although these extralegal principles may influence the interpretation of legal texts in some way.¹⁷⁹ Some, though by no means

174. See Stephen Darwall et al., *Toward Fin de siècle Ethics: Some Trends*, 101 PHIL. REV. 115, 149 (1992).

175. See generally TAYLOR, *supra* note 167, at 79 (stating that epistemological assumptions "tend to allow the natural sciences a paradigm status for all forms of knowledge, including that of human affairs . . ."); Brian Fay & J. Donald Moon, *What Would an Adequate Philosophy of Social Science Look Like?*, in READINGS IN THE PHILOSOPHY OF SOCIAL SCIENCE, *supra* note 13, at 21, 22; Benjamin Zipursky, *Legal Coherentism*, 50 SMU L. REV. 1679, 1679-88 (1997) (book review of Dennis M. Patterson's book, LAW AND TRUTH (1996)).

176. See, e.g., A. J. AYER, LANGUAGE, TRUTH AND LOGIC (1936).

177. See, e.g., Charles Taylor, *Neutrality in Political Science*, in READINGS IN THE PHILOSOPHY OF SOCIAL SCIENCE, *supra* note 13, at 547.

178. See, e.g., Robert Bork, *Neutral Principles and Some First Amendment Problems*, 47 IND. L.J. 1 (1971); Antonin Scalia, *The Rule of Law as a Law of Rules*, 56 U. CHI. L. REV. 1175 (1989); Herbert Wechsler, *Toward Neutral Principles of Constitutional Law*, 73 HARV. L. REV. 1 (1959).

179. See, e.g., Owen M. Fiss, *Objectivity and Interpretation*, 34 STAN. L. REV. 739 (1982).

all, legal theorists who call themselves positivists are attracted to positivism because it supposedly keeps the enterprise of deciding cases separate from resolution of the moral disagreement that is endemic in our society. As Owen Fiss points out, however, legal texts are inherently prescriptive.¹⁸⁰ They embody values and direct that a certain state of affairs consistent with those values be brought about. As a result, the interpretation of legal texts must be an evaluative enterprise, otherwise it cannot account for the sources of the law's authority. The implication for a theory of social norms, as I have previously noted, is that the theory must make use of explanatory concepts that can function in a normative argument about how a legal text ought to be interpreted, if the theory is to have any relevance for law. An austere rational-choice theory should be rejected on these grounds.

One of the most significant discontinuities between sociological accounts of social norms and strict varieties of rational-choice explanations is the inability of economic analysis to account for rational deliberation about preferences,¹⁸¹ or to support any kind of critique of preferences, as in, for example, the case of objectionable preferences like racism or sadism.¹⁸² Classic rational-choice theory assumes preferences as given; in operational terms it constructs utility curves based on consumers' revealed preferences, using data from market transactions. This is fine, for certain tasks, but singularly unhelpful for the project of understanding social norms. For when we ask why people do *X*, the answer, "Those actions reveal a preference for *X*" is question-begging. The issue is not whether people prefer *X*, but *why*, and the answer to this question requires an inquiry into the reasons behind the preference.

As Mary Midgley points out, in her critique of existentialism, one cannot understand a choice between conflicting desires as a function of the will, or arbitrary choice, because one cannot will something prior to deciding what to will.¹⁸³ The initial choice, which can later be backed up by the will (in the sense of being resolute or determined), must be made on the basis of *reasons*, unless a person is to surrender her agency entirely and be simply tossed about on the winds of random passions. I suppose such persons do exist, but they are generally committed to asylums. The philosopher Harry Frankfurt has famously argued that the distinguishing feature of persons is that they have "second-order volitions"—they can decide what to will.¹⁸⁴ A person who lacked second-order desires altogether would be a wanton, more like an animal than a person.

Oddly, that seems to be the picture of human nature that is attractive to rational-choice theorists, perhaps because it avoids any messy entanglement with the reasons for deciding what to will. But it puts them in the awkward position of subscribing to a kind of behavioralist conception of human life, where the only thing that matters to

180. *Id.* at 751-53.

181. Some law and economics scholars concede this point. *See, e.g.*, ELLICKSON, ORDER WITHOUT LAW, *supra* note 3, at 156 n.3.

182. Kaplow & Shavell, *supra* note 22, at 1339-50 (arguing that welfare economics can be saved from the malevolent-preferences objection, because our grounds for objection to certain preferences are ultimately welfarist in nature).

183. *See* MIDGLEY, *supra* note 102, at 184.

184. HARRY G. FRANKFURT, *Freedom of the Will and the Concept of a Person*, in THE IMPORTANCE OF WHAT WE CARE ABOUT 11, 16 (1988).

the study of some activity is the outward manifestation of human action. The social sciences were once more accepting of behavioralism, for all the reasons that rational-choice theories are appealing—it seems somehow “harder,” that is, more authentically scientific as compared with conceptually pluralistic accounts that make reference to beliefs, values, attitudes, and other terms that are relevant to an interpretation of human action, rather than a mere description. Although some economists deny that outward manifestations of preferences are equivalent to “rational, fully-informed preferences,” thus avoiding the association with behavioralism,¹⁸⁵ it nevertheless remains the case that at some point, the individual’s expressed preference (whether manifested through behavior or some kind of verbal affirmation of desire) is taken as given, and placed beyond the reach of normative criticism.

The emphasis on neutrality in the social sciences has been motivated, in part, by laudable reasons. One of these reasons is the elimination of parochial concerns from political principles, so that the good for humans is defined in terms that are not bound up with a particular culture or contestable conception of the good life.¹⁸⁶ Proceduralist philosophers like Rawls and Habermas maintain a distinction between universally acceptable procedural principles (often referred to as “rights”) and substantive principles of ethics (often referred to as “the good”) in order to defend against “chauvinistic and ethnocentric aggression in the name of one’s way of life, or tradition, or culture.”¹⁸⁷ It is no accident that contractarianism in political philosophy owes a great deal to John Locke, who was motivated by his experience of bitter religious warfare to seek a basis for social ordering in something other than the revealed truth to which religious traditions appealed.¹⁸⁸ Although most Western nations no longer experience anything like the religious strife of Locke’s time, we nevertheless do perceive profound, sometimes violent conflict over moral questions such as abortion, animal rights, environmental protection, and nuclear disarmament. Substantive principles of ethics seem to be part of the problem, not part of the solution, since the disputing sides subscribe to different substantive moral positions. For this reason, the search for neutrality through proceduralism and the Enlightenment liberal project offer a way out—it is a response to the pluralism or fragmentation of value in modern cultures that is the basis for so much interminable disagreement.¹⁸⁹

There are two reasons why the supposed advantages of natural science should not

185. Kaplow & Shavell, *supra* note 22, at 1331 n.901.

186. See TAYLOR, *supra* note 167, at 80-88.

187. *Id.* at 88.

188. For a recent history of moral philosophy that reveals the connection between religious disputes and the emergence of Enlightenment liberalism, and ultimately the views of Kant, see generally J. B. SCHNEEWIND, *THE INVENTION OF AUTONOMY: A HISTORY OF MODERN MORAL PHILOSOPHY* (1998).

189. For accounts of this fragmentation, see, for example, BERLIN, *supra* note 17, at 10-11; BARBARA HERMAN, *THE PRACTICE OF MORAL JUDGMENT* (1993); CHARLES LARMORE, *PATTERNS OF MORAL COMPLEXITY* (1987); ALASDAIR MACINTYRE, *AFTER VIRTUE* (2d ed. 1984); THOMAS NAGEL, *The Fragmentation of Value*, in *MORTAL QUESTIONS* 128 (1979); TAYLOR, *supra* note 167, at 496. None of these writers, except perhaps Berlin, can be described as a proceduralist—their responses to pluralism are varied, but do not take neutral principles as a foundation.

push us in the direction of a pure, methodologically thin rational-choice theory of social norms. The first is merely a quibble: If rational-choice theory is enamored of the methodology of natural science, it does seem fair to expect that it exhibit some of the advantages of the scientific method. Foremost among them is empirical testability (often couched in terms of falsifiability of a hypothesis) and predictive power.¹⁹⁰ But as many critics, even those who are generally sympathetic with economic analysis, have observed, claims of empirical verifiability have largely not been borne out.¹⁹¹ Instead, some rational-choice theorists have been content to offer only the kind of armchair empiricism they deride in their opponents. Posner's book, for instance, offers commonsense observations (many of which are probably apt) in the place of data that would satisfy a scientist.¹⁹² There is no reason we should take on faith his assumptions about the amount of signaling that would occur under certain conditions, in the absence of an empirical demonstration that the theory's prediction has been borne out.

The second objection goes deeper. It challenges the assumption that empirical observation can be detached from normative judgments. The construction of scientific theories cannot proceed merely by agglomerating together a bunch of value-neutral observations. For one thing, facts do not come prepackaged as such. An observed event is a "fact," for the purpose of empirical evidence for or against a theory, only by virtue of a correlation with other similar events. Significantly, this correlation depends on categories that are constructed by a theory and exist prior to the observation.¹⁹³ In the physical sciences, a chemist can observe the behavior of "benzene rings" and a particle physicist can describe tracks left by "quarks" only because these scientists are working with concepts that organize raw sense-data into some intelligible form. For what is a benzene ring other than a convenient schematic shorthand to represent and explain the observed inputs and outputs of a particular chemical reaction? In the social sciences these prior categories are even more overt. "Crime," "gang," "preference," "utility," "market," "surplus," "cartel," "political party," "interest group," "faction," "lobbying," and an almost infinite variety of other terms familiar from the social sciences are not natural kinds (think of the categories from any ethnographic account, such as Clifford Geertz's story of the sheep raid, which sorts people into "Jews," "Berbers," and "French," labels having considerable significance for the correct interpretation of the story). When applied to law, the normativity is even clearer. As the British historian E.P. Thompson observed, categories such as "serf," "cottager with common rights," and "landless labourer who

190. See, e.g., ROSENBERG, *supra* note 13, at 24-25.

191. See, e.g., DONALD P. GREEN & IAN SHAPIRO, *PATHOLOGIES OF RATIONAL CHOICE THEORY* (1994); Mark Cooney, *Why Is Economic Analysis So Appealing to Law Professors?*, 45 STAN. L. REV. 2211, 2222-27 (1993) (reviewing ELLICKSON, *ORDER WITHOUT LAW*, *supra* note 3); Barbara Yngvesson, *Beastly Neighbors: Continuing Relations in Cattle Country*, 102 YALE L.J. 1787 (1993) (reviewing ELLICKSON, *ORDER WITHOUT LAW*, *supra* note 3); McAdams, *Signaling*, *supra* note 11.

192. See, POSNER, *supra* note 19, at 49-166.

193. This is a point made again and again in the philosophy of science literature, to the point where it is something of a cliché. See, e.g., GORDON, *supra* note 13, at 604-08; ROSENBERG, *supra* note 13, at 201-06; TAYLOR, *supra* note 167, at 25-32. Like many clichés, though, it contains a considerable amount of truth.

may sue . . . his employer for assault," are essential ideas that one must use for making sense of eighteenth century history, yet they cannot be disentangled from theories of justice, equity, and good; they do not simply exist to be "read off" the text of history.¹⁹⁴ In short, the descriptive concepts used in observing, describing, and memorializing human behavior name intellectual constructs that give order to what would otherwise be a chaotic series of sense impressions of sound and fury, signifying nothing.

Not only are these categories socially constructed, they are pervaded with normativity. Consider the example mentioned earlier, of constructing a theory of voting (this is a favorite puzzle for rational-choice theorists to take a crack at, since it is so mystifying on their own terms. There seems to be no reason for a rational utility maximizer to incur the costs associated with voting, given the infinitesimal payoff that voters receive, measured by the effect they have on the outcome of elections). If the person's behavior, or the behavior of an aggregate of individuals, were described in truly value-neutral terms, the story would be incomprehensible. The account begins to make sense only when normative concepts are admitted, such as moral obligations of gratitude or fair play owed to one's country or fellow citizens. Why did someone go to a particular place, stand in a booth, and pull a lever (or dimple a chad)? An explanation that referred only to behavioral regularities ("people do this sort of thing on a designated date") obviously begs the question. Significantly, even methodologically thin rational-choice explanations which have a hard time accounting for why people incur the costs associated with voting describe the practice in normative terms: People seem to *feel compelled* to turn out and vote—they are moved by reasons other than utility maximization. This may be a puzzle for rational-choice theorists, but notice that already the description has departed from the purported ideal of value neutrality. There is nothing wrong with using evaluative terms in one's description, provided one is aware of this tendency and sufficiently critical about the content of the evaluative concepts used. But it is vain to hope for a "science" of human behavior that dispenses entirely with normative explanatory terminology.

Predictability and falsifiability are certainly criteria of theory acceptance, but they are not the only things that matter to an adequate theory of social behavior. It may be important for some purposes to "operationalize" theories—that is, by defining the theory in terms of empirical testing procedures. For other purposes, however, experimental verifiability may be a subordinate consideration. Theories that are deployed for explanatory purposes must surely be testable, but they also must go beyond just appearances or, in the case of social science explanations, the outward manifestations of behavior.

[T]hat which is without any observable manifestations whatsoever, which, so to speak, casts no shadow onto the plane of experience, would never have been considered as being of any use to science. Nevertheless, it is not unimportant whether we regard our operations as capturing at most the shadows of the furniture of the universe or as dealing with the furniture itself. Objects totally dissimilar in substance and even in size

194. E. P. THOMPSON, *WHIGS AND HUNTERS: THE ORIGIN OF THE BLACK ACT 266-67* (1975).

and shape may under particular circumstances cast identical shadows.¹⁹⁵

In the domain of social norms, a theory that utterly fails to connect with the meaningful, symbolic, interpretive, or expressive dimension of human action is no more favored than a theory whose morass of ad hoc explanatory postulates renders it incapable of empirical testing. Although the shadows cast by signaling, or some other manifestation of economic rationality, may resemble the shadows cast by other actions, it is important to discern whether the two objects are in fact the same.¹⁹⁶ It may be necessary, that is, to push back the level of theoretical inquiry to the point where we are observing the furniture itself, not merely the shadows cast on the plane of observable experience. The importance of this distinction, as I will argue in the last Part, inheres in the connection between a theory of social norms and the prescriptive conclusions that a theorist draws with respect to the legal system.

VI. "SHADOWS OR THE FURNITURE ITSELF": ARBITRATING THE DISPUTE

Law professors influenced by disciplines such as philosophy, sociology, anthropology, and psychology have been criticizing the law and economics movement practically from the moment it took root in the legal academy. Collectively, we need another voice in this debate like I need a hole in my head. What would be useful, however, is a way out of the conundrum, a roadmap toward a more productive line of discussion. I think there is a way to reframe the dispute, which depends on the practical use for which law and economics scholars propose their models be employed. This is an argument that relies upon what philosopher of science Wesley Salmon calls the pragmatics of explanation.¹⁹⁷ Salmon argues that the kinds of considerations that are relevant to an explanation vary according to the purpose for which an explanation is sought. It is my contention that it is possible to develop criteria for an adequate theory of social norms in the domain of legal scholarship by considering the way in which legal scholarship aims to be prescriptive or normative.¹⁹⁸

195. G. Schlesinger, *Operationalism*, in 5-6 *ENCYCLOPEDIA OF PHILOSOPHY* 543, 546 (Paul Edwards ed., MacMillan Reference 1996) (1967).

196. Compare this with the admittedly very different debate over whether the rule of law is supervenient (or superstructural, as the Marxists say) on class hierarchies, or has independent significance for both the ruling class and the underclass. See THOMPSON, *supra* note 194, at 258-69. The point of this seemingly obscure comparison is that Marxist explanations do make a certain amount of sense. The law can mystify and obscure abuses of power by the dominant classes. Even within the Marxist tradition, however, it is important to get it right. If the rhetoric of law also enabled the oppressed to obtain justice occasionally—to speak truth to power, as the saying goes—then a more nuanced conception of the rule of law is necessary. The debate about the rule of law described in this section of Thompson's book is an excellent example of a theorist seeking to understand and explain social phenomena, not merely making predictions. See *id.*

197. See *supra* text accompanying notes 146-48.

198. Cf. KORSGAARD, *supra* note 173, at 46 ("[E]xplanation and description of . . . phenomena is [not] the sole or primary function of human concepts. . . . Normative concepts exist because human beings have normative problems.").

Pick up an article on the economic analysis of social norms and it is a sure bet that it will end with a series of policy prescriptions, generally arguing that the law should take account of norms in one way or another, either by abstaining from regulation because prevailing norms do a sufficient job, or by entering the regulatory field because informal norms are either absent or unsatisfactory in some way. Robert Cooter, for instance, recommends incorporating social norms by reference into the formal regulatory structure of an industry.¹⁹⁹ Others, like Lisa Bernstein, recommend a “hands-off” approach to informal regulation within an industry or trade association.²⁰⁰ Looking at the reverse side of the same coin, Lawrence Lessig, Cass Sunstein, and Richard McAdams argue that the state should interfere with the development of pernicious group norms, such as dueling, racial discrimination, and the social approval granted to smoking.²⁰¹ Eric Posner offers a welter of suggested modifications to the law in a variety of substantive areas, to counteract the evolution of inefficient norms or to maintain efficient ones.²⁰² For example, he hypothesizes that the law disallows marital arrangements that differ from the standard one-man one-woman model contemplated by Congress in the Defense of Marriage Act, because community punishment of spouses who defect from a cooperative arrangement requires clear information about whether a person’s sexual activity is deviant or not.²⁰³ He is noncommittal as to whether involvement of the crowd in the enforcement of marital norms is a good thing, compared with state intervention, but his approach to these policy problems should be clear—identify the relevant costs, benefits, surpluses, inefficiencies, barriers to entry, and whatnot, and argue from some conception of efficiency to a normative conclusion. Significantly, Posner does not intend his model to be solely used for making predictions—it is avowedly offered as the basis for normative judgments about legal intervention in the marketplace: “The goals of [Part 3 of the book] are . . . to convince you that common normative judgments in legal analysis should take account of complexities of nonlegal regulation more often than they do.”²⁰⁴

Like Posner’s judgments, the policy proposals offered by other rational-choice theorists are generally underwritten by second-order normative arguments, of the following form: “*X* is a desideratum of the legal system. This social norm enhances (or detracts from) *X*. Thus, the social norm should be encouraged or left alone by the law (or the law should intervene and attempt to change the norm).” Unsurprisingly, *X* is often identified with economic efficiency by rational-choice scholars.²⁰⁵ Even within the law and economics circle, however, social norms are sometimes praised for furthering other values such as neighborliness and stability.²⁰⁶ Whatever the nature of *X*, the one constant is that it functions as a term in an evaluative argument and, as

199. See Cooter, *Decentralized Law*, *supra* note 98; Cooter, *Structural Adjudication*, *supra* note 98.

200. See Bernstein, *supra* note 3.

201. See Lessig, *supra* note 3; McAdams, *Discrimination*, *supra* note 9; Sunstein, *supra* note 3.

202. See Posner, *Inefficient Norms*, *supra* note 98, at 1725–43.

203. See POSNER, *supra* note 19, at 76–87.

204. *Id.* at 169.

205. Kaplow & Shavell, *supra* note 22; Posner, *Inefficient Norms*, *supra* note 98, at 1701.

206. See, e.g., ELLICKSON, *ORDER WITHOUT LAW*, *supra* note 3.

a result, it can be used to generate criteria of explanatory relevance for our metatheory of social norms. To emphasize, these criteria go beyond predictive value as soon as law and economics scholarship takes a normative turn. Although rational-choice proponents sometimes disavow any intention to go beyond constructing models that enable prediction,²⁰⁷ they seem unable as a group to eschew prescriptive conclusions.

When rational-choice theorists, or any legal theorists for that matter, make prescriptive arguments based on a model of human behavior, they are selecting portions of an ideal explanatory text on persuasive or rhetorical grounds.²⁰⁸ They are trying to persuade legal actors—judges, legislators, and law students—that because of Theory *T*, they should do such and such. The theoretical explanation is supposed to supply a *reason* for the legal actor; thus, we can criticize the explanation in the same way we would criticize any other proffered reason that a legal actor should do something—in jurisprudential terms. If Theory *T* predicts that a given rule will lead to inefficiencies (say, heightened transaction costs), the conclusion that the legal actor should not implement the rule depends on an implicit jurisprudential premise that legal rules ought to be structured to maximize efficiency. This premise may be true, or it may not, but it is essential to move from the theoretical structure to the prescriptive conclusion. Once we have identified this logical structure, one can readily see how theories of social norms may be criticized. The implicit jurisprudential premises that enable the move from theory to prescriptive conclusion supply the pragmatic criteria for selecting from the ideal explanatory text.

I will conclude this discussion with a hypothetical illustration of the linkage between practical issues in legal scholarship and theories of social norms. It is deliberately impressionistic and merely suggestive in nature. Also, I am not necessarily arguing for any particular theory of social norms, merely pointing out how one might criticize these explanatory models, in light of their purpose of buttressing normative arguments about the law. The topic is also essentially arbitrary—a great many disputed social issues could have provided an effective demonstration of the connection between theories of social norms and the law. I invite the reader to imagine other topics of interest, and apply them to theories of social norms.

Suppose a legal scholar is concerned about the sexual harassment of working women,²⁰⁹ and wishes to consider modifying the legal landscape in order to provide protection to these women. Our legal theorist recognizes that workplace relationships are structured not only by formal legal rules (existing antidiscrimination statutes, collective bargaining agreements, the law of contracts, and so on) but by informal social norms. As sociologist Robert Jackall has demonstrated, life in a large, bureaucratic organization is largely shaped by unwritten rules of conduct, such as: personal relationships come before abstract principles and moral values; never embarrass your boss publicly; don't go over someone's head in the chain of command; if someone superior to you in the organization tells you to drop a matter,

207. See, e.g., *id.* at 158 (“I am a positivist and am therefore interested in making and testing predictions.”).

208. Cf. GORDON, *supra* note 13, at 610-11.

209. An echo, naturally, of the title of Professor MacKinnon's groundbreaking book. See generally CATHARINE A. MACKINNON, *SEXUAL HARASSMENT OF WORKING WOMEN: A CASE OF SEX DISCRIMINATION* (1979).

do so without question; and, above all, don't rock the boat.²¹⁰ Daily life in the organization is also structured by social norms that are not specific to the workplace, such as the nature of joking and sexual innuendo that is widely considered acceptable among men and women,²¹¹ as well as (perhaps sexist) assumptions about the appropriate behavior of working women. The legal scholar would like to harness these norms, if possible, in order to make changes in the law more effective; at the very least, she is wary of modifying the law without paying attention to underlying social norms. Going too dramatically against the current in making legal changes could result in a backlash against the legal rules. Finally, our theorist has read and found persuasive the work of scholars who argue that legal rules have an expressive dimension, so that changes in the law will be understood as the state taking sides on contested moral issues and expressing respect for particular values.²¹²

Now the problem: Suppose we perceive a great deal of sexualized banter in a particular work setting, which is unwelcome and intimidating to women employees. The banter is essentially mandatory—a new employee must engage in it, in order to be considered “one of the guys.” (Because of the informal compulsion that attends the lockerroom environment, we can say that there is a social norm requiring employees to participate in the sexualized talk.) In light of the social norm to engage in harassing speech, the legal scholar seeks to design a new legal rule, or modify an existing one, to accomplish the goal of equalizing opportunities for equal participation in the workplace for women. Regarding the role of a *theory* of social norms in this analysis, the legal theorist must get her explanatory account of social norms correct so that she may properly calibrate the legal rules to take account of the reasons for the observed behavior in the workplace.

Let us consider a variety of alternative explanations for this phenomenon:

- *Interpretive Theory—Feminist*: Men regard women as sexual objects, not as equal participants in the workplace. Sexual harassment is simply a social ritual that reinforces this message and contributes to the exclusion of women from positions of equality at work.²¹³

210. See ROBERT JACKALL, *MORAL MAZES: THE WORLD OF CORPORATE MANAGERS* (1988).

211. For example, there was a tremendous public outcry (and a jury verdict in excess of \$20 million for wrongful termination) following an employer's decision to dismiss an employee for discussing the “Delores!” episode of *Seinfeld* in a mixed-sex company of coworkers. See *Mackenzie v. Miller Brewing Co.*, 2000 Wis. Ct. App. 48, ¶ 95-96, 608 N.W.2d 331, 359 (reversing jury verdict).

212. See, e.g., Matthew D. Adler, *Expressive Theories of Law: A Skeptical Overview*, 148 U. PA. L. REV. 1363 (2000); Elizabeth S. Anderson & Richard H. Pildes, *Expressive Theories of Law: A General Restatement*, 148 U. PA. L. REV. 1503 (2000); Dan M. Kahan, *What Do Alternative Sanctions Mean?*, 63 U. CHI. L. REV. 591 (1996); Richard H. Pildes & Elizabeth S. Anderson, *Slinging Arrows at Democracy: Social Choice Theory, Value Pluralism, and Democratic Politics*, 90 COLUM. L. REV. 2121 (1990); Richard H. Pildes & Richard G. Niemi, *Expressive Harms, “Bizarre Districts,” and Voting Rights: Evaluating Election-District Appearances After Shaw v. Reno*, 92 MICH. L. REV. 483 (1993); Cass R. Sunstein, *On the Expressive Function of Law*, 144 U. PA. L. REV. 2021 (1996).

213. This is one explanation offered by feminist theorists for sexual harassment at work. See, e.g., Kathryn Abrams, *Gender Discrimination and the Transformation of Workplace Norms*,

- *Interpretive Theory—Conservative*: Men intend to flatter the women with whom they work, by calling attention to their attractive physical features.²¹⁴ Only a minority of strident feminists object to sexual talk at work.²¹⁵
- *Interpretive Theory—Secret Handshakes*: There is no underlying sexual significance to the banter, but the clever use of sexualized insults and epithets is a ritual that is required to gain admission to a particular subculture in society—perhaps like the exchange of stylized insults and boasts known as “playing the dozens” or “signifying” in hip-hop culture.²¹⁶
- *Psychological Theory—Conformism*: It really means nothing at all; it is empty talk, like commenting on baseball games, traffic, and the weather. People feel compelled to participate because they do not want to appear different to their coworkers.²¹⁷
- *Psychological Theory—Gender Roles*: In a society in which women are striving to be viewed as equals, men sexually harass women because they feel profound anxiety about their own role in the workplace and in the family.²¹⁸
- *Rational-Choice Theory—Esteem*: Directing sexual banter at particular employees is a way of informally rewarding or punishing behavior.²¹⁹ The reasons why particular actions are praised or discouraged are exogenous to the theory.
- *Rational-Choice Theory—Signaling*: By alienating women through their obnoxious speech, men drive up the costs associated with being shunned by their male peers if they should behave disloyally toward men. Thus, they show that they are willing to endure expected disutility in the short run—that is, that they have a low discount rate.²²⁰

42 VAND. L. REV. 1183 (1989); Deborah Epstein, *Can a “Dumb Ass Woman” Achieve Equality in the Workplace? Running the Gauntlet of Hostile Environment Harassing Speech*, 84 GEO. L.J. 399 (1996); Vicki Schultz, *Talking About Harassment*, 9 J.L. & POL’Y 417 (2001).

214. This seems to be one of the implicit explanations of some conservative opponents of sexual-harassment law. See, e.g., Kingsley R. Browne, *Title VII as Censorship: Hostile-Environment Harassment and the First Amendment*, 52 OHIO ST. L.J. 481, 493-98 (1991).

215. See *Rabidue v. Osceola Refining Co.*, 805 F.2d 611, 626-28 (6th Cir. 1986) (Keith, J., dissenting).

216. I am not claiming to be an expert on urban black culture, but this phenomenon has been widely discussed in sociology literature. See, e.g., KATHERINE S. NEWMAN, *NO SHAME IN MY GAME: THE WORKING POOR IN THE INNER CITY* 115-16 (1999).

217. The macho, hypersexual environment of restaurant kitchens can be seen, for example, in ANTHONY BOURDAIN, *KITCHEN CONFIDENTIAL* 219-26 (2000), although it has the obvious effect of excluding women (as well as many men) who find the atmosphere threatening.

218. See, e.g., SUSAN FALUDI, *STIFFED: THE BETRAYAL OF THE AMERICAN MAN* (1999).

219. See, for example, the highly sexualized terms of approbation (for example, “big swinging dick”) used by securities traders, as described in MICHAEL LEWIS, *LIAR’S POKER: RISING THROUGH THE WRECKAGE ON WALL STREET* 153, 180 (1989).

220. Compare Posner’s account of hostile subcommunities, which develop norms at odds

Again, my claim is that all of these theories may be plausible fits with the observed behavior, and some may be empirically testable, but not all of them are appropriate theories—with “appropriateness” being understood on pragmatic grounds—to underwrite an argument about how the law of workplace sexual harassment ought to be structured.

Recall that the theory chosen by the legal scholar is a premise in a prescriptive argument. The conclusion of an argument of this form, generally directed to a court, legislature, or administrative agency, is that the law should be changed to better accomplish some end. A theory of human behavior is relevant to the argument if the legal theorist is sensitive to the considerations described above, such as the recalcitrance of some practices, even in the face of explicit legal sanctions, or the opportunity provided by synergistically linking legal rules with informal norms. In the case of sexual harassment, suppose the psychological conformism theory were true. Changing the law to prohibit sexual banter at work would probably be effective, because employees would simply move on to some other subject for chit-chat. If, on the other hand, the conservative interpretive theory were true, we might anticipate a great deal of resistance to legal change, because the regulated employees might perceive that only a few “politically correct” fanatics object to their behavior, which should properly be regarded as harmless, and because the institutions charged with enforcing the rule may share this attitude and refuse to take seriously a hostile-environment harassment claim that can be dismissed as just “boys being boys.” Perhaps an effort at changing the hearts and minds of men would be required along with the legal change²²¹ (alternatively, one might hopefully argue that the legal change would itself have this educative effect²²²). Similarly, if the signaling model were the correct explanation, changing the law of sexual harassment might stop the harassing speech, but men would figure out another way to antagonize women, in their continuing effort to show one another that they have low discount rates.

Remember also that it is necessary to reason backwards, from jurisprudential considerations to pragmatic criteria for theory construction. It matters what the law is about when we are constructing social theories with an eye towards being relevant to law. If law is about values like equality and human dignity, among other things, then a theory of social norms ought not to eliminate the semantic dimension of an explanation of sexual harassment at work. Even if sexual harassment does have a signaling function, and even if an adequate, empirically testable theory of social norms can be elaborated in terms of signaling (although I do not believe this to be the case), a theorist is not justified in treating meaning, intention, and the first-person perspective as exogenous to theory if those concepts have a role in the justification of legal rules. The difficulty with this approach, of course, is getting the interpretation right. Some recent feminist writing on sexual harassment has tended to deemphasize the sexual dimension and focus instead on the attempt by men to exclude women

with those of the dominant group. POSNER, *supra* note 19, at 97-103.

221. “Litigation also can be too crude a tool for achieving the often subtle changes in understanding that produce equal treatment or regard for women.” Abrams, *supra* note 213, at 1196.

222. See, e.g., Lessig, *supra* note 3.

from the workplace, by sexual harassment as well as by other means.²²³ I do not wish to enter into this intramural debate—I wish only to suggest that a theory of social norms that takes questions of meaning seriously is bound to become bogged down in these sorts of debates.

The alternative, though, is to ignore the semantic dimension of behavior altogether. Rational-choice theories that either treat meaning and intention as exogenous (as McAdams's does) or attempt to shoehorn all behavior into a one-dimensional signaling game (as Posner's does) risk misguiding a legal theorist who seeks to use legal processes to intervene in the world. If Vicki Schultz is right, and sexual harassment is an attempt by men to exclude women from equality in the workplace,²²⁴ then a rule that focuses exclusively on the sexual dimension of harassment would be unresponsive to the real underlying problem of power and status differentials between men and women in occupational settings. If sexual harassment is just using sex to exclude women from work, then it is the exclusion, not the sexuality, that lies at the heart of the social problem to be addressed. The central interpretive question for a theorist of workplace norms is, therefore, whether sexual harassment is misconduct that is embedded in a pervasive structure of inequality, or whether the social meaning of sexuality itself constructs the inequality of women. The answer to this question can make a great deal of practical difference to the question of whether harassment, for example, should be seen as integrated into a programmatic response to other instances of gender inequality, such as "mommy tracking" of women employees or the existence of glass ceilings.

The rational-choice scholars, for the most part, ignore questions like this, treating them as beyond the scope of their theories of social norms. By doing so, however, they render their theories irrelevant to the legal problems that need solving. If the ambition of economic theorists is purely positivistic—if they are concerned only with making and testing predictions—then all that really matters is behavior in the aggregate. Selecting only a small portion of the ideal explanatory text would be justified in that event. But the rhetoric of the economics crowd belies their stated modest positivistic ambitions. The prescriptive conclusions always seem to follow from their theories. Of course, law and economics scholars may argue that interpretive questions, such as those that pervade the debate over sexual harassment, are irrelevant, since the law is only "about" maximizing overall social welfare. This is the tack taken by Kaplow and Shavell, for instance, in their recent article.²²⁵ Making this move, however, commits rational-choice theorists to defending a strong jurisprudential position that is likely to be unpersuasive in the courts and the legal academy. Time and space prohibit launching on an extended critique of the welfarist theory of jurisprudence and, again, my ambition here has been to suggest lines of criticism, not to definitively settle disputes. This Article has been an exploration of the inference-to-the-best-explanation pattern of argument, and as long as the structure of this rhetorical device has been laid bare, it has accomplished its modest ambition.

223. See, e.g., Vicki Schultz, *Reconceptualizing Sexual Harassment*, 107 YALE L.J. 1683, 1710-69 (1998); Schultz, *supra* note 212.

224. See *supra* text accompanying note 212.

225. Kaplow & Shavell, *supra* note 21.