

BEHAVIORAL LAW AND ECONOMICS: A CRITIQUE

by
Richard A. Posner



ECONOMIC EDUCATION BULLETIN

Published by

AMERICAN INSTITUTE *for* ECONOMIC RESEARCH
Great Barrington, Massachusetts

About A.I.E.R.

AMERICAN Institute for Economic Research, founded in 1933, is an independent scientific and educational organization. The Institute's research is planned to help individuals protect their personal interests and those of the Nation. The industrious and thrifty, those who pay most of the Nation's taxes, must be the principal guardians of American civilization. By publishing the results of scientific inquiry, carried on with diligence, independence, and integrity, American Institute for Economic Research hopes to help those citizens preserve the best of the Nation's heritage and choose wisely the policies that will determine the Nation's future.

The Institute represents no fund, concentration of wealth, or other special interests. Advertising is not accepted in its publications. Financial support for the Institute is provided primarily by the small annual fees from several thousand sustaining members, by receipts from sales of its publications, by tax-deductible contributions, and by the earnings of its wholly owned investment advisory organization, American Investment Services, Inc. Experience suggests that information and advice on economic subjects are most useful when they come from a source that is independent of special interests, either commercial or political.

The provisions of the charter and bylaws ensure that neither the Institute itself nor members of its staff may derive profit from organizations or businesses that happen to benefit from the results of Institute research. Institute financial accounts are available for public inspection during normal working hours of the Institute.

You can receive our twice monthly *Research Reports* and monthly *Economic Education Bulletin* by entering a **Sustaining Membership** for only \$16 quarterly or \$59 annually. If you wish to receive only the *Economic Education Bulletin*, you may enter an **Education Membership** for \$25 annually.

ECONOMIC EDUCATION BULLETIN

Vol. XLII No. 8 August 2002

Economic Education Bulletin (ISSN 0424-2769) (USPS 167-360) is published once a month at Great Barrington, Massachusetts, by American Institute for Economic Research, a scientific and educational organization with no stockholders, chartered under Chapter 180 of the General Laws of Massachusetts. Periodical postage paid at Great Barrington, Massachusetts. Printed in the United States of America. Subscription: \$25 per year. POSTMASTER: Send address changes to *Economic Education Bulletin*, American Institute for Economic Research, Great Barrington, Massachusetts 01230.

BEHAVIORAL LAW AND ECONOMICS: A CRITIQUE

By Richard A. Posner*

Abstract

There is increasing interest in the use of the branch of cognitive psychology usually called “behavioral decision theory” to modify neoclassical economic analysis of law, heretofore the dominant methodology of economic analysis of law. This paper, while acknowledging the fruitful insights generated by the behavioralist literature, emphasizes the weaknesses of the literature, which include a tendency to caricature the rational model of human behavior and by doing so to overlook the considerable extent to which the insights claimed by behavioralists are *already* a part of the rational-choice version of economic analysis of law; a puzzling and unexplained refusal to seek theoretical grounding for behavioralism in evolutionary biology; a neglect of methodological problems that afflict the empirical side of behavioralism; and an exaggerated claim to have contributed substantially as yet to understanding and improving the law.

Introduction: Rational Choice

Economists, lawyers, and psychologists are endeavoring to use the insights of cognitive psychology (more specifically of that branch of cognitive psychology usually referred to as “behavioral decision theory”) to refine, or in some versions to refute, economic analysis of law, which these analysts believe to be handicapped by its commitment to the assumption that people are rational.¹ This emerging field of “behavioral law and economics” (or “law and psychology,” as Rachlinski terms it²) has a number of useful insights, some of which I have employed in my own academic work.³ But the field as constituted at present has major weaknesses, and it is those that I shall emphasize in this paper. They include a tendency to caricature the rational model of human behavior and by doing so to overlook the considerable extent to which the theses of behavioral law and

* Judge, U.S. Court of Appeals for the Seventh Circuit; Senior Lecturer, University of Chicago Law School. An abridged version of this paper was given on July 19, 2002, at “Behavioral Economics and Neoclassical Economics: Continuity or Discontinuity?,” the Second Annual Symposium on the Foundation of the Behavioral Sciences sponsored by the Behavioral Research Council of the American Institute for Economic Research. The paper incorporates (but with many changes) portions of chapter 8 of my book *Frontiers of Legal Theory* (2001), which in turn is based on my article “Rational Choice, Behavioral Economics, and the Law,” 50 *Stanford Law Review* 551 (1998), and also portions of parts 8 and 9 of Gertrud M. Fremling and Richard A. Posner, “Status Signaling through Market Behavior” (Aug. 21, 2001, unpublished). I thank Brian Grill, David Kitchen, and Carl LeSeur for their helpful research assistance, and the participants in the Symposium for their helpful comments.

economics are already a part of the rational-choice version of economic analysis of law; a puzzling and unexplained refusal to seek theoretical grounding for behavioralism in evolutionary biology; a neglect of methodological problems that afflict the empirical side of behavioralism;⁴ and an exaggerated claim for the implications behavioralism for understanding and reforming law. My title may seem a bit misleading, since the focus of the paper is less on the application of economics to law than on the internal weaknesses, as it were, of behavioralism; but this is the right focus because as yet the applications of behavioralism to legal issues have been limited and often rather vague; specific reform proposals have been few, and have primarily to do with how issues are best presented to juries, how risks to health or safety are most informatively disclosed and compared, and how workers are likely to respond to regulations of the labor market. I shall touch on a few of these proposals, but the touch will be light.

To take the first point of criticism first: economic analysis of law, without abandoning its commitment to the rational model of human behavior, has abandoned the model of hyperrational, emotionless, unsocial, supremely egoistic, omniscient, utterly selfish, nonstrategic man (or woman), operating in conditions of costless information acquisition and processing, that cognitive psychologists rightly deride as unrealistic, and, more important, that is deficient in explanatory and predictive power with regard to a number of the phenomena in which economists and economically minded lawyers are interested. The availability heuristic is one of the principal discoveries of the behavioralists, yet it is consistent with rationality once one acknowledges that imaginative reconstruction requires more “effort” (that is, cost) than immediate perception; in other words, once thinking is understood to be a costly activity. For example, the argument for allowing the prosecution in a capital case to place “victim impact” statements before the jury is that without them jurors would have to exert extra effort to imagine the victim’s suffering in order to counterbalance the impact of the immediate perception of the suffering defendant, pleading for his life. The assumption that thinking is costly will play a key role in the analysis that follows, in which I show that much of the behavioral critique of rational choice rests on a too-narrow conception of the rational-choice model. This in fact is my most fundamental criticism of behavioralism.

It may be objected that the question whether to allow victim-impact statements is really no different from the question whether a restaurant should be allowed to serve lobster without forcing customers to view the living lobster, the “victim.” Analytically they are quite similar questions; they are both about the stopping point in providing people with the history of the “product” that they are being asked to buy. But there is no reason why they have to be answered the same way. Regard for human and for

crustacean life differs; yet one can imagine animal rights insisting, as a way station to compulsory vegetarianism, that people who eat meat be forced to watch the animal killed.

Behavioral Law and Economics: A Taxonomy

Following the taxonomy in an article by Jolls, Sunstein, and Thaler,⁵ I shall discuss the fundamental theses of behavioral law and economics under the following rubrics: “bounded rationality,” “bounded willpower,” and “bounded self-interest.”

Bounded rationality asserts that people have cognitive quirks that prevent them from processing information rationally. These include the already-mentioned availability heuristic, overoptimism, the sunk-cost fallacy, loss aversion, and framing effects. There is much evidence for the existence of these quirks, but because most of it consists of experiments with students or of responses to surveys, it is unclear to what extent the quirks are due to serious, stubborn obstacles to rational action; to what extent they are due merely to the operation of mental shortcuts in settings in which the optimal investment in thought is low because so little turns on thinking through the problem posed by the experiment or survey; and to what extent—as seems to me the most plausible explanation—they are due to the artificiality of the experimental settings that have furnished the bulk of the evidence for the cognitive quirks, a point to which I return later in this paper.

One important distinction that behavioralists typically fail to make is between impediments to clear instrumental reasoning and preferences that enlightened observers may think silly. A person might eat a lobster contentedly if he doesn't see it when it's alive, but if asked to pick it out of a lobster tank he might well lose his appetite for it. Has his mind been fogged by the availability heuristic? That is one possibility but another is that he has different preferences for two different goods: one a lobster seen only after being cooked and the other a lobster seen before, in his living state, as well as after. These are different goods in the same way that a good that comes in pretty wrapping paper is different from one that comes in a brown paper bag. There is no basis for pronouncing a difference in preferences with regard to such pairs of goods irrational (although ethical criticism may be possible, as animal-rights advocates remind us) or for dismissing the difference as a product of “emotion.” When people react with fright to a horror film, we might be tempted to say that they are being irrational, because the movie is make-believe. But preferences cannot be divorced from emotion, or emotion from their stimuli, and so instrumental reasoning cannot be thought pervaded by irrationality merely because the goal to which such reasoning is instrumental may be a preference that we would

not have if we were not emotional beings. The way to distinguish a real cognitive quirk from an emotion-driven preference is to ask whether if you point out to a person the “irrationality” of his action he either will change it or at least will admit that he’s being irrational. Obviously it would do nothing to a liking for horror movies to point out that they are make-believe; the people who watch them know that already.

Even the fact that human beings are not always rational need not seriously undermine rational-choice economics. Many people have an irrational fear of flying. We know that it is an irrational fear rather than just an aversion that we may not share because the people who harbor it acknowledge that it is irrational and often (if they have the time and money) will seek therapy to “cure” them of their fear. For they know that the surface-transportation alternatives are more dangerous and they want above all to avoid being killed, yet they choose the more dangerous mode (and not because it’s cheaper). Their regret, embarrassment, and annoyance with themselves confirm their case as one of genuine irrationality and thus distinguishes it from the case of the people who like horror movies. But their irrationality does not cast a cloud over the economic analysis of transportation, although it may show why pecuniary and time costs, and accident rates, may not explain the entire difference between the demand for air transportation and the demand for its substitutes. Most of the usual predictions that economists make about behavior will go through: a fall in the price of air transportation will result in an increase in demand, as will a rise in the price of a substitute or a fall in the price of a complement. A preference can be taken as a given, and economic analysis proceed as usual, even if the preference is irrational.

Another example is the market price of residential real estate adjacent to a funeral parlor. Because people don’t like to be reminded of death, the proximity of the funeral parlor is likely to depress that price. It is an example of the availability heuristic at work. Yet the usual predictions of economics will go through, such as that if the local public schools improve, the market price of the residential real estate will increase notwithstanding its proximity to the funeral parlor.

Voting, long used by critics of rational-choice economics as a prime example of irrational behavior, can be analyzed similarly. It is indeed true that when viewed as an instrumental act voting in a political election is irrational, because it costs something (chiefly time) to vote yet there is no offsetting benefit to the individual voter because no such elections are ever decided by one vote. But, treating the desire to vote as a given, in much the same way that other expressive behavior (for example, applauding at a concert or other public performance) is normally treated in economics as a

given rather than something the economist is obliged to explain, the economist can answer important questions about voting behavior. These questions include why the old vote more than the young, why retired people vote more than unemployed people (even though both groups might seem to have low costs of time), and why turnout is greater in a close election.⁶ Turnout is greater in a close election not because one's vote is more likely to make a difference—even close elections are not decided by one vote—but because the costs of information are lower the more publicity an electoral contest generates, and close elections generate more publicity than one-sided ones, not only because a close race is more exciting but also because candidates will tend to spend more the closer the race; the marginal benefit of campaign expenditures is greater. The effect of increased campaign expenditures on turnout is similar to that of increased advertising expenditures on the sales of a product. Although increased advertising may to a great extent merely alter the market shares of the different brands of the product, there is likely to be at least some effect on the overall demand for the product.

Bounded willpower, the second category of alleged departures from rationality, is essentially just a relabeling of weakness of will. Most of us have experienced the sensation of being torn between two selves—a “good” self that has our long-run welfare in mind and a “bad,” short-sighted self—and of the “bad” self winning unless strenuous efforts are made to thwart it. Hyperbolic discounting is said to illustrate the operation of weakness of will, although it can equally well be understood in terms of information costs. A hyperbolic discounter increases his discount rate as the costs or benefits that he is discounting become more imminent. For example, if you asked me whether I would rather have \$1,000 in the year 2010 or \$800 in the year 2009, I would almost certainly say \$1,000 in 2010. But if you asked me whether I would rather have \$800 today or \$1,000 a year from now, I might very well say \$800 now. And this would mark me as a hyperbolic discounter. But the reason for the different reactions may simply be that I lack a clear conception of my consumption needs a decade and more hence; the reason may, in other words, be the imagination cost that I mentioned earlier. I cannot imagine what might make me pay in effect a huge interest rate to reallocate consumption from 2010 to 2009. The fact that knowledge and imagination are “bounded” just shows, what no rational-choice economist doubts, that information costs are positive. Similarly, when Adam Hirsch in an interesting recent paper argues that the judicial policy of adhering to precedent is a response to bounded rationality,⁷ one is inclined to respond that it is equivalently an argument about information costs.

The “imagination cost” interpretation of hyperbolic discounting is

strengthened by taking a closer look at the difference between hyperbolic and the normal, “rational” exponential discounting. When the period over which some future cost or value is being discounted is short, hyperbolic discounting produces a higher discount rate than exponential. But when it is long, the reverse is true because of the explosive effect of exponential expansions. This makes good sense in terms of imagination cost. It will usually be rational in the narrow sense to prefer \$1.10 in a year to \$1 today; but it does not follow that the same person asked to choose between \$1 in 10 years and \$2 in 20 years would choose the \$1, even though with exponential discounting at 10 percent \$1 after 10 years is worth a great deal more than \$2 after 20 years. But it would take a great exertion of the imagination to imagine what difference it would make to the individual’s welfare to have an extra dollar in 10 years versus an extra two dollars in 20 years.

Yet no one can doubt that there is such a thing as weakness of will, even if hyperbolic discounting is not a good example of it. But unlike many of the cognitive quirks that go to make up “bounded rationality,” it can be analyzed within the framework of rational-choice theory—easily when we are torn between alternative courses of action because of uncertainty, less easily when there is no uncertainty, as in the case of refusing to keep chocolate in the house because of doubt of one’s being able to overcome temptation. In this example there is no deficiency of information or of reasoning; everything is understood but the individual is incapable of making the “correct” choice. Explaining this second type of behavior in rational-choice terms is still possible, but may require abandoning a tacit assumption of most economic analysis—that the self is a unity—in favor of a conception of the person as a locus of different selves. All the selves are rational but they have—rationally—inconsistent preferences. Examples are a young self versus an old self, with the former unwilling to save money so that the latter can enjoy a high level of consumption; a pre-accident self unwilling to spend heavily on accident insurance versus a post-accident self that would have liked the pre-accident self to buy a lot of accident insurance; and in the case of the chocolate, a present-oriented self that lives for the moment and a future-oriented self.⁸ (The last example is related to the first.) The assumption of a unitary self is not inherent in the concept of rationality used in economics; it is merely a convenient assumption in most situations that economists analyze. When a person takes an action to prevent himself from yielding to temptation, the concept of the individual as a locus of sometimes warring factions becomes the most natural way of modeling the behavior. The movement from analyzing the individual as one self to analyzing him as a locus of multiple selves can be analogized to the movement from treating the business firm as a “black

box” to analyzing the relations among the people involved in the firm, such as managers and shareholders, who may have conflicting interests.

As in this example, behavioral economists tend to give up on rational-choice economics too soon.⁹ Another example is hindsight bias: “when someone is asked to estimate the probability of some event before it occurs and then asked later what the prior estimate was,” he is likely to indicate that his *ex ante* estimate was higher than it actually was.¹⁰ Paul Rubin, whom I’m quoting, does not regard this phenomenon as the result of some inexplicable quirk, but instead as explicable in terms of the costs of processing information: “When asked for the *ex ante* estimate, most people cannot recall it, and so go through the same calculation process that was used to provide the original estimate. However, because we routinely update information as new data become available, when we perform the calculation, we get a bias because we automatically use the new data in the analysis. The costs of any hindsight bias are less than would be the costs of increased memory capacity to avoid this problem.”¹¹ Hindsight bias is a well-attested phenomenon, one that we should worry about when for example a jury is asked to decide whether the injury giving rise to a negligence case was foreseeable (probable) before it happened.¹² But as Rubin shows, we need not give up on rational-choice economics in order to acknowledge its existence and importance.

Here is another example of giving up too soon on rational-choice economics: Matthew Rabin writes that “a nominal wage increase of 5 percent in a period of 12 percent inflation offends people’s sense of fairness less than a 7 percent decrease in a time of no inflation.”¹³ People know that not all wages will increase by the rate of inflation—inflation spells economic trouble; only with perfect indexing would real wages remain unaffected by it. So the failure of one’s wage to rise by the rate of inflation need not imply that one’s employer is critical of one’s work. But a sharp wage cut out of the blue is often a signal of dissatisfaction with an employee’s work, and so engenders anxiety or resentment.

There is an even simpler and more fundamental explanation for the different responses to the two methods of reducing a worker’s real wage, and again it has nothing to do with “fairness” or even with the endowment effect (preferring what you have just because it is yours). Even when a worker does not have an employment contract that fixes his wage at a definite amount, it will generally be the implicit understanding between him and his employer that if his work is satisfactory and the demand for the employer’s product stable or growing, his wages will not be reduced. There is no similar expectation that his wages will grow in lockstep with any increase in inflation. That just is not part of the implicit bargain, and so

there is no breach of an implied understanding if, in a period of inflation, wages do not rise commensurately. If a worker wants his wages indexed to inflation, he will have to negotiate specifically for that. A similar explanation is available for a related reaction that behavioralists in my view unnecessarily attribute to the endowment effect—the greater disappointment at having one’s wage cut than at not receiving as large a bonus as one did last year, even if the two cuts are identical in magnitude. An employer awards a bonus rather than an increase in wages because he wants the employee to understand that he is not committing himself, even tentatively, to continuing to pay the higher income in future years. Stated differently, the employer increases the wage rather than the bonus when he wants to signal a high probability that the employee’s income will continue at that level, and he raises the bonus instead when he wants to signal a lower probability of such a continuation. The difference in signals is helpful to the employee in planning his expenditures, since we know from studies of the consumption function that people tend to save a higher fraction of their transient than their permanent income.

And here is a simpler example of giving up to soon on rationality. Behavioral experiments reveal that most people would rather a rare disease be eliminated entirely than that the incidence of a more common disease be reduced even if the reduction would provide a somewhat greater benefit in overall health. Thus, if there were two equally serious diseases, one with a 1 percent incidence and the other with a 2 percent incidence, and the choice was between wiping out the first disease altogether and reducing the incidence of the other to .5 percent, apparently most people would prefer the former. But this would probably make economic sense, because by eliminating a disease entirely (such as smallpox or polio), the medical profession and the general population would completely eliminate the costs of prevention, treatment, and training. To the extent that these are fixed costs, rather than costs that vary with the incidence of the disease, the cost savings might well exceed those that would be generated by a greater reduction, but not to zero, in the incidence of the more common disease.

But I have strayed from “bounded willpower” and let me come back to it. Self-destructive behavior, such as drug addiction and unsafe sex, may seem dramatically to illustrate weakness of will. Not necessarily. It is important to distinguish between *ex ante* and *ex post* choices. *Ex post*, the addict may be incapable of resisting a craving for the addictive substance even if he “knows” that it is bad for him to give in to the craving. *Ex ante*, however, his choice to become addicted (or, more realistically, to risk becoming addicted) need not bespeak any weakness of will; it is just the opposite of self-binding behavior, such as not keeping chocolate in the house, knowing that one would could not resist eating it. The rational

would-be addict merely figures in the downside of addiction as one of the costs of becoming addicted; the costs may nevertheless be outweighed by the perceived benefits. If addicts are rational, the long-run price elasticity of addictive goods should be high, rather than low as conventionally believed, because the rational addict expects his consumption of the addictive good to increase over time and so an increase in the price of such a good (if the increase is expected to be long term) will have a greater long-term effect on his expenses than the same increase in the price of a nonaddictive good would have.¹⁴ Or consider the hypothesis that the AIDS epidemic will cause an increase in the rate of unwanted pregnancies by inducing a rational substitution of condoms—which are good prophylactics against disease but mediocre contraceptives—for the pill, which is an excellent contraceptive but no prophylactic. To the extent this hypothesis is supported,¹⁵ it suggests that highly emotional, frequently impulsive behavior, such as sexual intercourse, may nevertheless be quite rational.

Behavioral economists like to cite Gary Becker's demonstration that random choice in a situation of scarcity will generate a downward-sloping demand curve.¹⁶ From this it might seem to follow that the Law of Demand, probably the most robust hypothesis in economics, cannot be used to support the rational model of behavior. Becker's argument is that consumers have limited budgets, and so on average, even if their purchasing decisions are made randomly, they will purchase less of a pricier good because a fixed amount of money will not buy as much of it. He did not suggest that most consumers *are* irrational, however, or that well-attested economic phenomena other than the downward-sloping market demand curve, such as the tendency of the prices of the same good to be equalized, could be explained without assuming rationality. Buyers do not in fact choose randomly. Rationality is the only reasonable explanation for their reactions to changes in relative prices.

Randomness actually helps explain *why* rational-choice economics can explain so much behavior. Most questions that economists ask concern aggregate rather than individual behavior, for example the effect on the quantity purchased of cigarettes of an increase in the cigarette excise tax, not the effect of the tax increase on Mr. Cigarette Smoker A or Ms. Cigarette Smoker B. Suppose the cigarette tax is increased by 2 percent and rational smokers respond by reducing their purchases of cigarettes by an average of 1 percent, while the irrational ones respond randomly—some reduce their purchases by 50 percent, some actually increase their purchases, and so on. If the distribution of these random behaviors has the same mean as the rational smokers' reaction to the tax, the effect of the tax on the quantity demanded of cigarettes will be identical to what it would be if all cigarette consumers were rational. This is true no matter what fraction

of cigarette consumers is irrational, so long as not all are.

“So long as not all are....” This qualification brings to the fore an important point made by Mitchell in his criticism of the empirical methodology of the behavioralists.¹⁷ He points out that most tests of behavioral hypotheses are “between subject” rather than “within subject” in design. That is, rather than inquiring whether a given individual is rational, the experimenter seeks to determine what percentage of individuals in the study group failed to give the rationally correct answers to the test questions. The fact that that percentage is greater than zero does not show that all of us are afflicted with serious cognitive biases, and is consistent with a socio-economic system in which people are sorted into different activities in such a way as to minimize the bad effects of cognitive biases.

Sorting casts a shadow on the experiments that provide the principal basis for the behavioralists’ thesis that rational-choice economics fails to explain as much human behavior as rational-choice economists think it does. The experimental subjects are normally a random draw from a relatively heterogeneous population; but people are not randomly sorted to jobs and other activities.¹⁸ Experiments in how well people handle probabilities are not conducted on future actuaries, statisticians, or econometricians, but on student or other groups containing a fair share of individuals who have “math block” or who in any event are not very adept at handling mathematical concepts. Maybe only 1 percent of the test group is good at math, in which event the results of the test will be taken as strong evidence against the rational-choice model—yet it is from that 1 percent that the future actuaries, etc. will be drawn, and there is no basis for thinking that *their* behavior will be deformed by cognitive quirks and as a result that insurance markets will behave irrationally. Similarly, people who are unusually “fair” will avoid (or, again, be forced out of) rough-house activities—including highly competitive businesses, trial lawyering, and the academic rat race. Hyperbolic discounters will tend to avoid the financial-services industry.

This is the glass-half-empty/glass-half-full phenomenon (itself a “framing” effect that should interest behavioralists). An experimental study that finds that a third of the subjects achieve results close to the economic optimum, another third do pretty well, and the remaining third are confused¹⁹ is really quite reassuring; for a third of the population is probably ample to furnish society’s needs for people able to calculate rationally.

Selection²⁰ will not work perfectly, but it will drive a wedge between experimental and real-world consequences of irrationality. It would be interesting, therefore, to compare the subsequent career paths, and earnings, of students who score high in rationality in experiments conducted by

behavioral economists with those who score low.

Bounded self-interest refers to the fact that people sometimes act out of motives (termed in the behavioral literature “fairness”) that do not seem explicable by self-interest even in the sense, which is now conventional in rational-choice economics, in which an altruistic act is self-interested. If an increase in *A*’s utility will increase *B*’s utility, this means that *B* is altruistic toward *A* and it may therefore be in *B*’s self-interest to transfer resources to *A*. This is different from a situation in which a person will do something for other people—or against other people—because he thinks it the fair thing to do. Acting out of duty (or out of gratitude²¹) is not the same thing as acting out of altruism, because it is not motivated by sympathy or antipathy for the object of the act.

As an aside, I note that lumping fairness in with cognitive quirks and weakness of will suggests that behavioral economics is merely the negative of rational-choice economics—the residuum of social phenomena unexplained by it. The fact, if it is a fact, that “fairness” is important to some people some of the time has nothing to do with the fact that people have difficulty processing some types of information (bounded rationality) and subordinating short-run to long-run interests (weakness of will). Those are disabilities or insufficiencies; acting in accordance with notions of fairness is a strength. The cognitive quirks belong to cognitive psychology, weakness of will to the psychology of neurosis and other abnormalities, and fairness to moral philosophy or moral psychology.

Behavioral Man: An Undertheorization

The picture that emerges from the three assumptions that I am taking to define the behavioralist approach is of a person who has trouble thinking straight or taking care for the future but who at the same time is actuated by a concern with being fair to other people, including complete strangers. This may be a psychologically realistic picture of the average person,²² and it responds to the familiar complaint that “economic man” is unrecognizable in real life. But quite apart from the fallacy of supposing that everybody is average and thus neglecting selection effects, the behavioralist preoccupation with descriptive accuracy gives too short shrift to the point that in theory-making descriptive accuracy is purchased at the price of sacrificing predictive power. The rational-choice economist asks what “rational man” would do in a given situation.²³ Usually the answer is pretty clear and can be compared with actual behavior to see whether the prediction is confirmed. Sometimes it is not confirmed—and so we have behavioral economics. But it is profoundly unclear what “behavioral man” would do in any given situation. A compound of rational and nonrational capacities and impulses, he might do anything. Behavioralists have neither a

causal account of behavioral man nor a model of his decisional structure. The result is a host of unanswered questions. Do cognitive quirks diminish as the costs of yielding to them rise? If so, why? Does weakness of will vary across people, and, again, if so why? Do behavioralists believe that their own analysis is marred by cognitive quirks or weakness of will, or actuated by a sense of fairness, or of resentment at being treated unfairly? If not, why not? And are the quirks curable? Is weakness of will curable?

These questions are made both urgent and mysterious by the undertheorization of behavioral economics. It is undertheorized because of its residual, and in consequence purely empirical, character. Behavioral economics is defined by its subject rather than by its method and its subject is merely the set of phenomena that the most elementary, stripped-down rational-choice models do not explain. It would not be surprising if many of these phenomena turned out to be unrelated to each other, just as the set of things that are not edible by man include stones, toadstools, thunderclaps, and the Pythagorean theorem. Describing, specifying, and classifying the empirical failures of a theory is a valid and important scholarly activity. But it is not an alternative theory.

Granted that it is easy to formulate a theory that will explain, in the sense of subsume, all observations within its domain, however anomalous they are from another theoretical standpoint; the trick is simply to relax whatever assumptions in the other theory made some of the observations anomalous. The rotation of the moons of Jupiter was anomalous in medieval cosmology because each planet (other than the earth, which was not considered a planet, but instead the center around which the planets revolved) was thought to be fastened to a crystalline sphere, with which the moons would have collided in their rotation. The anomaly could be dispelled by assuming either that the sphere was permeable or that the telescopic observations that had disclosed the rotation of Jupiter's moons were a deceit by the devil. Whichever route was taken, the amended theory would not generate any predictions about planetary satellites; all it would predict was that whatever would be, would be.

Similarly, if rational-choice theory bumps up against some example of irrational behavior, the example can be accommodated by changing the theory to allow for irrational behavior, in other words, by substituting behavioralism for rationality. But there is no greater gain in predictive power than in the cosmology example; in both cases, in fact, there is a loss. If a theory is so vague or elastic that it cannot be falsified, neither it nor its predictions can be validated; everything that happens is by definition consistent with the theory. When people act rationally, the behavioralists do not treat this as contradicting the assumption of bounded willpower. When

people resist temptations, thus demonstrating strength of will, this is not treated as contradicting the assumption of bounded willpower. And when they act selfishly, this is not thought to contradict the assumption of bounded self-interest. If people became more rational, this would be attributed to their having learned the lessons of behavioral economics, and so would operate to confirm rather than refute it. So the question arises, what if any observation would falsify the theory? If none, there is no theory but merely a set of challenges to the theory-builders, who in the relevant instances are rational-choice economists and—as I am about to argue—evolutionary biologists.

Behavioral Phenomena Theorized; Herein of Evolutionary Theory

“Fairness” is the vaguest word in the language. Most of morality, including all of welfare economics, could be subsumed under it.²⁴ What behavioralists are particularly interested in is the *sense* of fairness that, they argue, frequently causes people to act against their self-interest as ordinarily conceived. I believe that to a great extent this sense of fairness can be made precise, and explained, and subsumed under a broad conception of rationality, with the aid of the evolutionary biology of positive and negative altruism. The behavioralists’ lack of interest in, and indeed hostility to, evolutionary theory²⁵ is an example of their lack of theoretical ambition. But it is more. Most though by no means all behavioralists are political liberals. The use of evolutionary theory to explain human social rather than merely physical traits, the use that goes by the name “sociobiology” (recently renamed by its proponents “evolutionary ecology” because of the negative connotations that “sociobiology” had acquired among the politically correct), is anathema to liberals²⁶—as, indeed, is economics; and much of “behavioral economics” is really anti-economics. Political bias is especially conspicuous in the neglect by behavioralists of vengeance, though it is the best attested example of the “fairness” instinct. Liberals do not like vengeance and prefer to think that our instinct for fairness is dominated by altruistic concerns that might provide a foundation for organizing society along socialist or collectivist rather than free-market lines.

Evolutionary biology sees altruism as a trait that promotes inclusive fitness, defined as maximizing the number of copies of one’s genes by maximizing the number of creatures carrying them weighted by the closeness of the relation.²⁷ The inclusive fitness of a social animal like man is greatly enhanced by his having a proclivity to help his relatives, and so it is plausible to suppose that this proclivity evolved as an adaptive mechanism.²⁸ In the prehistoric epoch in which our instinctual preferences were formed, people lived in tiny, isolated bands. Most members of one’s com-

munity would have been either one's relatives, or nonrelatives having very close affective ties to one (such as one's mate and his or her family), or at least having very frequent—indeed virtually continuous—face-to-face dealings with one. In these circumstances it would not have been essential to have an innate capacity to discriminate between relatives and other intimates, on the one hand, and, on the other hand, those people—call them “strangers”—with whom one did not have repeated face-to-face interactions.²⁹

Conditions today are different. We interact a great deal with strangers. But our instincts are easily fooled when confronted with conditions to which human beings never had a chance to adapt biologically. That is why a pornographic movie can arouse a person sexually or a violent one frighten an audience, why people can love an adopted infant as much as they would their own biological child, why people are more frightened of spiders than of cars and of airplanes than of far more dangerous terrestrial forms of transportation, and why men do not clamor to be allowed to donate to sperm banks. It may explain “loss aversion,” which “means that we dislike losses much more than we like gains—about twice as much.” Ruben, whom I am quoting again, explains that in the ancestral environment people lived at the subsistence level, so that even modest fair gambles would have asymmetric effects: starvation if one lost, a moderate increase in fitness if one won.³⁰ Voting, giving to charities, and refraining from littering, in circumstances in which there is neither visible reward for these cooperative behaviors nor visible sanctions for defection, may illustrate an instinctual, and as it were biologically mistaken, generalization of cooperation from small-group interactions, in which altruism is rewarded (hence reciprocal) and failures to reciprocate punished, to large-group interactions in which the prospects of reward and punishment are so slight that cooperation ceases to be rational.³¹

Negative altruism is illustrated by the indignation that we feel when someone infringes our rights. The extreme expression of this emotion is the passion for revenge. This may seem the antithesis of rational thinking, because it flouts the economist's commandment to ignore sunk costs, to let bygones be bygones. Not that it is irrational to threaten retaliation in order to deter aggression; but if the threat fails to deter, carrying out the threat will often be irrational. No matter how much harm you do to the aggressor in return for what he has done to you, the harm that you have suffered will not be annulled. Whatever dangers or other burdens you assume in order to retaliate will merely increase the cost to you of the initial aggression. But if retaliation is futile for rational man, this will make the aggressor all the more likely to attack not the average man but—rational man. The aggressor will know that rational man treats bygones as bygones (or as econo-

mists say, ignore sunk costs) and is therefore less likely to retaliate than a less rational person. This calculation will lower the anticipated costs of committing aggression.

What was needed for deterrence and hence survival in a human society before there were any formal legal or political institutions, and thus before it was possible to make a legally enforceable commitment to retaliate against an aggressor, was an *instinctual* commitment to retaliate. People who were endowed with an instinct to retaliate would have tended to be more successful in the struggle for survival than others. Sometimes retaliation ends in disaster; but inability to make a credible threat to retaliate renders a person virtually defenseless in a prelegal, prepolitical society. The desire to take revenge for real or imagined injuries—without calculating the net benefits of revenge at the time it is taken, because such calculation would, as I have suggested, reduce the credibility of the threat to retaliate and so would invite aggression that would in turn reduce a person's inclusive fitness—may therefore have become a part of the human genetic makeup.

I have contrasted rational man with vengeful man, but the contrast is superficial. The real contrast is between *ex post* and *ex ante* rationality. Having an unshakable commitment to retaliate may be *ex ante* rational by lowering the risk of being a victim of aggression, even though, if the risk materializes, acting on the commitment will then (that is, *ex post*) become irrational. Put differently, a certain emotionality may be a component of rationality, which I defined at the outset as suiting means to end rather than as a particular form of ratiocination.

But we must consider *why* a person may become indignant not only when his own rights are infringed but also when another person's rights are infringed. The key is altruism (so positive altruism lies at the base of negative altruism). This is easy to understand in the case in which the person whose rights have been infringed is a relative or close friend. But it is operative even when he is a stranger. For in that case the phenomenon of "fooling the instincts" is in play and the attenuated but nonetheless positive altruistic feelings that we have even for complete strangers engenders a corresponding indignation if the stranger's rights are infringed.

This analysis may explain what has long puzzled moral philosophers—why we are more indignant at the driver who runs down a child carelessly than at the more careless driver who through sheer luck misses the child.³² The altruistic instinct is triggered in the first case but not the second. We are hurt in the first case by the loss of the child even though it is not our own child. There is no loss in the second case.

The behavioralists' favorite example of how fairness can trump rationality is drawn from the "ultimatum game," which is played as follows. *A* is given an amount of money. He can offer as little or as much of it as he pleases to *B*. If *B* accepts the offer, *A* gets to keep the rest; if *B* refuses, neither gets anything. One might suppose, therefore, that however large the stakes, *A* would offer only a penny, since he would expect *B* to accept the offer rather than go away empty-handed. Yet in fact the *A*'s in the ultimatum game invariably offer the *B*'s a substantial fraction of the stakes. The explanation that behavioralists give for this result is that the proposer (*A*) and the respondent (*B*) share a concept of fairness. But this is just a labeling of the result of the game; the process that generates it remains mysterious.

We can make progress by viewing the game through the lens of negative altruism. To gain anything from playing the game, the proposer has to make an offer generous enough to induce the respondent to accept. As this necessity exists whether or not the proposer has any sense of fairness, there is nothing even remotely irrational—hence nothing that requires a concept of fairness to explain—about his offering more than a penny, provided however that he has a good reason to think that the respondent will not accept a penny offer. So we can forget about the proposer and concentrate on the respondent, and ask, Why won't he take a penny? (If he won't, and the proposer knows this, the proposer will offer more; and that is what we observe in the playing of the ultimatum game.) The likeliest reason, but one that behavioralists resist because they do not think in evolutionary terms, is that the offer would signal to the respondent the proposer's belief that the respondent holds a low supposal of his own worth, that he is grateful for scraps, that he accepts being ill-used, that he has no pride, no sense of honor.³³ This weak-spirited creature is just the type who in a prepolitical, vengeance-based policy would have been stamped on by his aggressive neighbors and, thus deprived of resources, have left few offspring. The neighbors would have trampled on his rights because they would have known that he had no sense of *having* any rights and was in any event too diffident to act in defense of them. It is from the aggressive neighbors that most of us moderns are descended, and we reveal our prideful genetic heritage in a wide variety of settings, one of which is the ultimatum game. The game itself shows that this heritage continues to be rational in a range of instances—it is what enables the respondents in the ultimatum game, and their counterparts in analogous real-world situations, to avoid complete defeat. The vengeful spirit was the basis of the nuclear deterrent that contributed, perhaps vitally, to avoiding a world war for half a century. It is also the basis of most reporting of crime in those situations, which are common, in which neither the victim of the crime nor any other

potential reporter or witness of it anticipates a selfish gain from reporting.³⁴

I mentioned signaling, but not to suggest that the respondent in the ultimatum game is seeking a reputation for toughness because he expects to be playing the game again with the same proposer. That would be an easy case for rational-choice analysis. The difficult case is where there is no prospect of repeat play—indeed, no observation at all of the player's moves. In that case, the response of turning down an insultingly low offer is in a narrow sense emotional rather than rational, but in a broader sense rational because the emotion that generates it is part of a cognitive-emotional complex that enables the making of commitments that are rational *ex ante*.³⁵

The acid test of the rationality of the players in the ultimatum game would be if the offer, though a minute percentage of the stakes, were large in absolute terms.³⁶ Suppose the stakes were \$1 million, and the offer was \$10,000. Although the offer would be meager in relative terms (only 1 percent of the stakes), one would be surprised if many respondents refused it. The demand of a rational person for the “goods” of pride, self-esteem, and vengeance is not perfectly inelastic, and so as the price of these goods rises from \$10 (say) to \$10,000, we can expect the amount demanded to decline and thus more offers to be accepted.

Another focus of behavioral economics, loosely related to fairness in representing a kind of primitive sense of entitlement, is the already mentioned “endowment effect.” When students in an experiment are given coffee mugs that they are free to sell to other students, few sales take place, because the “owners” of the mugs mysteriously charge a higher selling price than the other students are willing to pay.³⁷ Indeed, students given the mugs at the outset of the original experiment demanded a price roughly twice as high as the others were willing to pay, which was why there was very little trading even though the initial distribution was random rather than being based on how much the students valued the mugs. But this is not such a mystery from an evolutionary perspective. The only “rights” in prehistoric society were possessory rights, and so people who didn't cling to what they had would have been at a disadvantage, and maybe the outcome of the coffee-mug experiment is a vestige of a rational adaptation to a vanished situation.

But we do not need evolutionary theory to explain the outcome of the coffee-mug experiment. In any bargaining situation, the potential seller (a student who had been given a mug) is likely at first to demand a higher price than the potential buyer is willing to offer. Bargaining proceeds until the demand and offer cross or the parties give up. When the good in question has a very low value, abandonment of the bargaining process

before its completion may be common. That is why in real-life rather than experimental settings, low-value items are usually sold at a fixed, take-it-or-leave-it price rather than being subject to bargaining.

As we have begun to see, evolutionary biology offers the hope of a rapprochement between rational-choice economics and behavioral economics. This is the theme of Paul Rubin's book *Darwinian Politics*, to which I have already referred, and of important work by Owen Jones.³⁸ Rubin explains that many of the phenomena identified by the behavioralists reflect cognitive limitations and emotional dispositions stemming from the fact that human beings reached their current state of neurological development in the very different conditions of life thirty or more thousand years ago.³⁹ He gives the important example of voting: even though, as I noted earlier, voting in a political election is not costless (although the costs are low) yet the benefits in terms of the probability of swinging the election are essentially zero, many people vote. Are they irrational? One possibility, mentioned earlier in this paper, is that voting is not perceived as an instrumental act, but rather is a form of consumption, perhaps most closely analogous to applauding a performance. But Rubin suggests another possibility—that individuals overestimate their likely impact on the outcome of an election because they do not have a good intuitive sense for probabilities, the reason being that in the ancestral environment the equivalent of a “vote” (the expression of a preference) would have been cast in a setting of very few people, where one vote could well be decisive.⁴⁰ Probabilities are a useful way of thinking about large numbers; but large numbers were not a characteristic feature of the ancestral environment.

In like vein, commenting on our “preference for benefiting known individuals at the expense of unknown or statistical individuals,” Rubin remarks that “in the evolutionary environment the notion of one thousand deaths—let alone one million—was not something that could exist, or that our ancestors could grasp. Our ancestors simply would not have known or known about that many people,”⁴¹ and so their brains may not have developed to the point where they could handle statistical phenomena with ease.

This may also be why it is “difficult for humans to grasp the significance of certain tragedies or of certain political practices, including genocide.”⁴² And why people do much better on the kind of tests that behavioralists administer in an effort to demonstrate cognitive limitations when questions are posed in terms of frequencies rather than probabilities.⁴³ It is no surprise that the most impressive findings of behavioralism concern the difficulties that ordinary people have in assessing risks (probabilities) of harm.⁴⁴

(Speaking of how questions are framed, I believe that many of the

findings of the behavioralists reflect misunderstandings rather than mistakes. One of the most common such findings is that when people are asked whether they are a better than average driver, more than half say they are. I suspect that many people asked such a question interpret “average” to mean minimally competent.)

The more common case in which the endowment effect is observed is where the good that one is asked to part with has been in one’s possession for a long time; and here the reluctance to part with it can be understood in straightforward rational-choice terms.⁴⁵ Any rational person who owns a good, except the marginal owner, values it above the market price—otherwise he would sell it. This implies that owners of the good as a class value it more than nonowners do.

An alternative explanation draws on the idea of rational adaptive preference. We rationally adapt to what we have, and would incur new costs to adapt to something new. A person who is blinded in an accident must incur costs to adapt to being blinded. But a blind person who through a doctor’s negligence fails to regain his sight has already adapted to being blind, so his loss of (prospective) sight imposes a smaller cost than the sighted person’s loss of sight.

It may be objected that to speak of adaptive preferences violates the rational-choice economist’s normal assumption of stable preferences. But obviously people’s preferences change, so all the assumption can signify is that ordinarily it is facile and uninteresting to explain a change in behavior (for example, a fall in demand for some good as a result of an increase in its relative price) by saying that peoples’ preferences changed. That’s like “explaining” irrational behavior by saying that people aren’t always rational—a true statement, but not a helpful one. Rejecting the facile invocation of preference changes doesn’t place beyond the bounds of economics the explanation for why certain undoubted preference changes do occur.⁴⁶

Surveys of attitudes toward national parks and other recreational public lands reveal what behavioralists deem dramatic endowment effects. Asked how much money they would demand to sell their rights to use such lands, people give much higher figures than when they are asked what they would offer to buy the rights. This disparity need not be thought irrational. It may reflect the remoteness of such a hypothetical transaction from our experience with markets.⁴⁷ In this respect, opinion surveys share the artificiality of behavioral experiments, the other principal source of empirical support for behavioralist hypotheses. Or the disparity may simply reflect the absence of close substitutes for access to national parks.⁴⁸ Absence of a close substitute implies that a good could not be replaced easily if it were lost. So the owner demands a high price to part with it. But if he doesn’t

own it, he may be unwilling to pay much for it because he doesn't know what he is missing; by hypothesis, nothing he owns is much like it.

Selection effects, and the closely related phenomenon of specialization, are important here. In a modern economy the sale of goods and services (other than labor) is to a considerable extent professionalized. Most individuals, including most respondents to opinion surveys and virtually all university students—the principal experimental subjects of behavioral economics, which relies much more heavily than standard economics on experiments—are buyers but not sellers, and moreover are buyers from stores and other institutions rather than, with rare exceptions, individuals. When we have something to sell we usually sell through middlemen, such as real estate brokers, rather than directly to the ultimate consumer. Experimental situations in which the subjects are asked to trade with each other are artificial, and so are surveys asking us to value national parks in pecuniary terms, and so we cannot have much confidence that the results generalize to real markets.

Here is another example of the importance of the difference between hypothetical and real situations. When married people are asked the probability that they will divorce, they give very low numbers, usually around 5 percent; yet the current divorce rate is 50 percent or higher. Does this show that people are irrationally optimistic? Not necessarily. People do not behave as if they *really* believed their probability of divorcing was low. The rise in pre- and postnuptial agreements, the high age at first marriage, the fact that women increasingly hedge against the risk of divorce by establishing themselves in the market before they marry, the low birth rate, the increased incidence of pre-marital cohabitation (a kind of “trial” marriage), and the low marriage rate are all responses (in part anyway) to “gut feeling” recognition that the risk of divorce is in fact very high. Social scientists who used the divorce estimates obtained by the behavioralists either to predict the divorce rate or to dismiss marriage as an irrational institution would understand modern marriage less well than social scientists who model prospective spouses as rational actors.

The broader point which this example brings out is that a traditional strength of economics is skepticism of rhetoric as an explanation of behavior. It is apparent from the rhetoric of politics, advertising, and for that matter ordinary conversation that, admittedly for reasons not well understood by economists but possibly connected with the importance of language as a signal of intentions, that speech is a good deal more moral and conventional than behavior. The divorce example is a perfect illustration. A person about to get married who was asked what he thought the probability of divorce was and answered 60 percent (roughly the national aver-

age) would be thought cynical, depressed, frivolous (not taking marriage seriously), or a pessimist. Rather than seem weird, he is likely to give the conventional answer, which is expected; it costs him nothing.

Rhetorical conventions seem to influence and be influenced by certain not very well understood characteristics of human subjectivity or phenomenology. The language in which we not only communicate with others but communicate with ourselves, evaluating our behavior and examining our intentions, seems poorly descriptive of that behavior, and of those intentions as realized in action. Social scientists interested in predicting behavior as distinct from the structure of thought and feeling are likely to be misled if they take at face value the way in which we describe our motives and actions to ourselves and to others.

Coming back to the coffee-mug experiment, I doubt that the outcome illustrates anything more mysterious than the operation of habit. Habitual behavior occurs when cost and benefit are time-dependent and cost is negatively related to time and benefit positively related to it.⁴⁹ Not only is it cheaper to brush one's teeth after brushing has become habitual, but to stop brushing (maybe in response to convincing evidence that it was actually bad for one's teeth) would make one uncomfortable. Breaking a habit, like throwing off an addiction (an extreme example of habit), causes withdrawal symptoms, though in the case of a mere habit they usually are slight and fleeting. Habit formation is one way in which "learning by doing" operates; tasks are performed more quickly and with less effort when they become habitual. If acting in accordance with the endowment effect is rationally habitual because of the real-world examples given earlier (such as rational adaptive preference), this may explain the outcome of the coffee-mug experiment even though that outcome is irrational if habit is ignored.

Habit may explain a good deal of "fair" behavior. Benjamin Franklin famously said not that honesty is required by the moral law, but that honesty is the best "policy." A reputation for honesty is an asset in personal and business relations, and a habit of honesty is a method of assuring steady compliance with this policy at low cost. It resembles self-binding as in the chocolate example. This is still another side of the voting paradox. Because democracy depends on a certain percentage of the population voting, failing to vote is in some quarters reprobated as uncivic-minded. People who wish to maintain a good reputation in those quarters will therefore want to be able to tell people that they voted. It is easiest to do this if they actually did vote, and therefore don't have to incur the costs of lying, slight as they are (making the lie credible and avoiding its being unmasked)—but then the costs of voting are slight too. It is no surprise that

the incidence of voting is greatest among those who travel in circles in which civic-mindedness is valued.

Signaling and Methodology

There is a further problem with the behavioralist interpretation of the results of the coffee-mug experiment. I pointed to signaling as a possible explanation for the results of the ultimatum game, but dropped the suggestion on the assumption that play was anonymous. Even if that assumption is correct, the conclusion that signaling cannot possibly explain the outcome of the game need not follow. For if the *type* of behavior that is the subject of the experiments is characteristically observed by other people rather than being anonymous, the experimental behavior may be similar as a matter of habit or inertia. In addition, however, the assumption of anonymity is questionable. Most behavioralists' descriptions of their experiments do not indicate what if any efforts were made to protect the experimental subjects' anonymity. Yet if a player's behavior is observed by the teacher and by the other students, the experiment will be an occasion for the subjects to signal the possession of traits, for example generosity in the ultimatum game and wealth in the coffee-mug experiment (the recipient of the mug signals by his refusal to sell it that he doesn't need money), that may enhance his reputation with the teacher or the other students.⁵⁰ The incentive to engage in such signaling may distort the results of the experiment. When the students in the coffee-mug experiment were given tokens redeemable in cash, rather than mugs, there was no reluctance to trade. The difference in outcomes may be explained by reluctance to sell a consumer good (the mugs) in response to a small increase in price, because willingness to sell would signal an elastic demand and hence that the person is not affluent.

Loewenstein and Issacharoff added the following twist in their version of the coffee-mug experiment: whether one received the mug depended on how well one scored on a short test administered at the beginning of their experiment.⁵¹ When it was the low scorers who received the mugs, they did not value them any more highly than the high scorers who did not receive them. But when it was the high scorers who received the mugs they valued them much more highly than the low-scoring nonreceivers did. The signaling interpretation of these results is that when received by low scorers the mugs were in the nature of booby prizes, and this lowered their value because possession of them conveyed a negative signal; it marked the possessor as a low scorer. The low scorers would have been eager to part with this symbol of their poor performance on the test.

In a well-known paper, Richard Thaler presents several cases that he says falsify the predictions of rational-choice economics, yet all can be

seen to involve signaling. The first case is that of Mr. R, who “bought a case of good wine in the late ’50’s for about \$5 a bottle. A few years later his wine merchant offered to buy the wine back for \$100 a bottle. He refused, although he has never paid more than \$35 for a bottle of wine.”⁵² Thaler thinks that R’s reaction is due to the endowment effect, but a more plausible explanation is that R wants others to view him as being well-to-do and having high opportunity costs of time. A professional wine trader would be proud to make a large profit by selling the wine at a high price. But R is a consumer, who by refusing the high offer of \$100 signals that he can afford to consume the wine even though it (opportunity) costs him that much to do so. He also signals that he is too busy to be bothered by such a trivial matter as returning a bottle of wine to the dealer in order to get \$100, while his unwillingness to pay more than \$35 a bottle signals not poverty but a disdain for hedonistic consumption. By signaling affluence, busyness, and frugality, he makes himself a more attractive potential transacting partner, whether in markets or in personal interactions.

Thaler’s second example, again offered to show the operation of the endowment effect, is that of Mr. H, who “mows his own lawn. His neighbor’s son would mow it for \$8. He [Mr. H] wouldn’t mow his neighbor’s same-sized lawn for \$20.”⁵³ Indeed, very few people would consider asking a neighbor or colleague to mow their lawn or clean their house for money. Such a request would signal a belief that the neighbor or colleague is so necessitous, or perhaps so lacking in pride or self-esteem—or so greedy—that he is willing to perform low-status work. In other words, the offer conveys a negative signal about the offeree. It is insulting, and will be resented, and so will not be made unless the offeror *wants* to insult the offeree.

Context is all-important here. Many teenagers do mow lawns for money and so it would not be rude to inquire whether a neighborhood teen would do yardwork. Few teenagers, even those possessing considerable status within their cohort, are able to find part-time, high-paying, prestigious jobs, and even if his parents are wealthy the teenager may be expected to work so as to learn to shoulder responsibilities or earn pocket money to buy things his parents refuse to buy for him. As it is already assumed by everyone that the teenager is not a high-class worker, a proposal that he do yardwork is not a challenge to his status and therefore is not offensive and likewise, of course, if the proposal is to a person known to make his living as a yardworker. Also different is asking your neighbor or colleague to do something for you as a “favor,” for here the signal is that he is altruistic, not that he is poor, desperate, or lacking in a sense of self-worth.

Thaler gives a third example of what he considers irrational behavior:

Two survey questions: (a) Assume you have been exposed to a disease which if contracted leads to a quick and painless death within a week. The probability that you have the disease is 0.001. What is the maximum you would be willing to pay for a cure? (b) Suppose volunteers were needed for research on the above disease. All that would be required is that you expose yourself to a 0.001 chance of contracting the disease. What is the minimum payment you would require to volunteer for this program? (You would not be allowed to purchase the cure.)⁵⁴

The typical response to (a) is only \$200 but to (b) is \$10,000, and Thaler again ascribes the difference to the endowment effect. With students as experimental subjects, there may well be a liquidity effect that would explain the difference; but a more interesting interpretation is that being deliberately exposed to a disease (b) is degrading. It resembles experiments on laboratory animals, prisoners of war, and inmates of prisons and insane asylums—all “persons” of low status. There is no comparable connotation to refusing to purchase an expensive cure. So while the probability of death is the same in (a) and (b), the latter involves an additional cost, that of emitting a negative signal about one’s status.

Thaler gives an example of how regret can induce irrational reactions. “Mr. A is waiting in line at a movie theater. When he gets to the ticket window he is told that as the 100,000th customer of the theater he has just won \$100. Mr. B is waiting in line at a different theater. The man in front of him wins \$1,000 for being the 1,000,000th customer of the theater. Mr B wins \$150. Would you rather be Mr. A or Mr. B?”⁵⁵ Some people, according to Thaler, choose to be Mr. A and Thaler’s explanation is that “the knowledge that he just missed winning causes regret to Mr. B.”⁵⁶ Signaling again provides an alternative explanation. Mr. A is a winner; Mr. B is a loser who received a large booby prize. Although no accomplishments are involved in this particular example, being a winner draws attention to oneself and may create a reputation for being lucky. An even clearer example is a competition of ability or skill. The fastest runner receives more status from being Number One than from being faster by some particular number of seconds than the second-best runner. Rank order, because it conveys the *essential* information so efficiently (which is why many prospective employers would prefer to know a candidate’s class rank than his grade-point average), can have a value in itself that a small sum of money cannot equal.

Signaling is *to* someone and therefore in experimental settings, where the reaction of the professor and the other students to any signals conveyed by the subject’s response could be important, the *degree* of anonymity may affect the outcome. Few experiments control for this possibility. Most are “anonymous” only in the sense that the answers that the experimental

subjects give to the questions they are asked are not publicly announced. Most take place in a single class period and are discussed by the students and professor afterward, and in those discussions the veil must often slip. Since the items traded or not traded in these experiments are usually worth only a few dollars but the opinions of the professor and fellow students are likely to matter greatly to each student, it is plausible that the students would demonstrate some degree of self-consciousness in their answers and hence that the answers would be likely to contain a signaling component. The prospect of future classroom discussions or informal chats between students after the experiment can affect the students' responses even if the responses were completely anonymous when made. For example, if somebody claimed to have answered a question in a certain way but classroom discussion revealed that all the answers given fell in a different range, the lie would be unmasked.

In an experiment on the endowment effect using lottery tickets rather than coffee mugs,⁵⁷ the experimenters found to their surprise that the tickets generally traded for much higher prices than the expected payoff. But a follow-up survey revealed that the students had been anxious to participate in the lottery because the "social pleasure of participating in the lottery with the group outweighed the value of the prize."⁵⁸ If social influence has such a large effect on the value of the stakes, it is plausible that signaling, a form of social interaction, could also explain much of the seemingly irrational behavior of the experimental subjects. A student in one experiment asked "Is this 'Candid Camera'?" and a student involved in a study that was much like the coffee mug one but involved pens instead said "I feel silly waiting around to exchange one pen for another for just a shekel [a small unit of Israeli currency], but it seems to be the 'correct' thing to do."⁵⁹ It is "silly" to sit around for a long time to make a small transaction because it signals low opportunity costs of time.

A recent study tested for the effects of anonymity by rerunning experiments using double-blind procedures to eliminate the possibility of cuing by the experimenter. Sure enough, the results showed a marked, indeed a dramatic, diminution in behavior motivated by a sense of fairness.⁶⁰

The interpretive problem created when experimental subjects lack complete confidence in their anonymity is acutely presented by a recent study of undergraduates' view of whether lying is a good or a bad thing.⁶¹ Undergraduates would be concerned that if the experiment were not truly anonymous, they would be revealing untrustworthiness and harming their careers if they failed to condemn lying.

The ultimatum game, as we saw, tests for the influence of altruism and concepts of fairness on the behavior of the experimental subjects and

likewise the “dictator” game. The players in that game are told to divide a sum with another person, as in “if you get \$10, how would you split it with another subject in a different classroom?” A high fraction of the “dictators” give the other player a nontrivial fraction of the stakes and in the ultimatum game the fractions are even higher, the difference being explained by reference to the offeror’s concern that the offeree will reject a highly unfavorable split because he will consider it “unfair” (there is no similar concern in the dictator game, because the dictator’s share does not depend on the response of the other player to the dictator’s “offer”). The difference can be explained without reference to “fairness.” An individual who gives something *voluntarily*, as in the dictator game, signals that he is an altruist or at least that he cares whether others *think* he is an altruist,⁶² while a small offer in the ultimatum game not only signifies a lack of altruism (or lack of caring about being thought an altruist) but also is an offensive challenge to the offeree’s status. Knowing that acceptance of the penny offer will convey a negative signal about the offeree, the offeror will fear rejection and will therefore have an incentive, quite apart from his own signaling motives, to make a much more generous offer. By refusing the ungenerous offer the offeree is also signaling his affluence. The price of signaling pride, self-esteem, or related personal characteristics by such a refusal is forgoing the money offered by the offeror. The wealthier the offeree, the higher his marginal rate of substitution of nonpecuniary for pecuniary goods.

There is experimental evidence that the degree of anonymity affects the outcome of the dictator game.⁶³ The more anonymous the game, the smaller the “dictator’s” gift. This result supports the signaling hypothesis, as does the accidental finding in the same study that in an experiment in which a high degree of anonymity was supposed to be guaranteed by requiring each “dictator” to place his gift in a sealed envelope, many students did not do so and their choice was not random. “There was a pronounced tendency for those leaving no money to seal their envelope, and for those leaving positive amounts of money to not seal their envelopes.”⁶⁴ In like vein it has been found that “if fairness is a social norm guiding distribution behaviour, it needs monitoring, i.e., it can prevail only in environments where behavior is observable.”⁶⁵ Another study found that anonymity increases the willingness of offerees in the ultimatum game to accept the offering player’s “ultimatum.”⁶⁶

All this is not to deny that altruism, and not just a desire to be thought an altruist, is at work in the dictator and ultimatum games. Rational-choice theorists accept that altruism is a common argument in people’s utility functions. Especially when the players in the ultimatum or dictator game are students in the same class, acquainted with each other and to some

extent, at least, friends, or when the prospective offerees are traditional objects of charity,⁶⁷ we would expect a degree of altruistic behavior even if none of the players had, as posited by the behavioralists, a concept of “fairness.” But we also expect the tendency to engage in such behavior to be amplified by the benefits for one’s reputation of appearing to be generous and trustworthy.⁶⁸

It has been claimed that “the outcomes of ultimatum, dictatorship and many other bargaining games have more to do with manners than [with] altruism” because by turning down small offers the respondents in the ultimatum game show that they are not altruistic toward the offeror.⁶⁹ I agree that altruism is not the complete explanation for these outcomes, but not because there is any contradiction between the behavior of the “dictators” in the dictator game and the behavior of the respondents in the ultimatum game. Only saints are altruistic toward people who insult or otherwise mistreat them. Hence we are not surprised that, as I noted earlier, respondents in the ultimatum game do not turn down stingy offers when they know that the offer is made by a computer,⁷⁰ since there is neither a revenge motive nor a signaling motive for rejecting such an offer. Notice also that the “revenger” who turns down the small offer in order to keep offerors “honest” is a kind of altruist; his self-sacrificing behavior contributes to the maintenance of an efficient social order. All this has nothing to do with “manners,” a term that connotes an unthinking reflex based on early training or “culture” and remote from considerations of self-interest.

Nonexperimental Empirical Studies of Behavioral Phenomena

I have been emphasizing experimental and survey results, but they do not entirely exhaust the empirical evidence marshaled by the behavioralists against nonbehavioral economics. An example of real-world behavior that behavioralists point to is the phenomenon of criminals’ hyperbolic discounting of future punishments. I explained earlier why I don’t consider hyperbolic discounting necessarily inconsistent with rationality, but even if it is, the behavioralists’ analysis is unconvincing.

A peculiarity of criminal punishment, when it takes the form of imprisonment, which is the standard form it does take in this country, is that a reduction in its probability cannot easily be offset by an increase in its severity. The only way to increase severity is to add prison time at the end of the criminal’s sentence. If the sentence is already long, any increments of length will have little weight in the criminal’s calculations, simply because of ordinary, not hyperbolic, discounting. For example, lengthening a prison sentence from 20 to 25 years will increase its disutility (in “present value” terms, that is, as reckoned by the defendant when he is deciding whether to commit a criminal act that would expose him to such a

punishment) by much less than 25 percent; at a discount rate of 10 percent, the increase will be only about 6 percent.

But I am willing to grant that many criminals are hyperbolic discounters rather than merely ordinary discounters. Indeed I think it likely. For we must consider the selection effects of a criminal punishment system. If it is designed to deter, then criminals—which is to say, the part of the population that is *not* deterred—will not be a random draw from the population. We can expect the undeterrable to have peculiar traits, including, in a system in which punishment takes the form of imprisonment, an abnormal indifference to future consequences. Most criminals are not very intelligent and this may make it hard for them to imagine future pains. This does not show that a criminal justice system should be designed on the assumption that the population of *potential* criminals, the people we think we can deter from becoming actual criminals, is dominated by hyperbolic discounters.

What is true is that any personal discount rate⁷¹ higher than necessary to adjust for the risk of death is suspect from the narrowest rational-choice standpoint, as it implies an arbitrary preference for present over future consumption. But this presentist orientation can be profitably analyzed in terms of rational choice, as I suggested earlier, either through the concept of multiple selves (the present self is in control of a person's present actions, and disvalues the welfare of the person's future selves) or because of information costs that make it difficult to imagine our state of mind in the future.⁷²

Selection effects also explain some at least of the phenomenon of over-optimism (the “winner's curse”). People are more likely to want to enter an activity if they think they will do well in it. But the competition among such people will reduce the likelihood of success, so that viewed *ex post* their original expectations will seem inflated.

A better example of an apparent departure from rationality is that of mandated childbirth health insurance. A careful study found that the wages of the workers covered by the insurance fell by the full cost of the coverage.⁷³ This implies that the workers valued the coverage at its full cost⁷⁴ even though, before it was imposed, they did not value it that highly, since if they had, the employer would have provided it voluntarily. The implication is that the coverage changed the workers' preferences; something they didn't like before they had it they did like once they got it, just as in the case of the coffee mugs.

But there is again a selection problem—or rather two such problems. Women planning to have children will be attracted to employments in

which childbirth insurance is available,⁷⁵ thus competing down wages. And, given the insurance, women will be more likely to have children, and this will make them less productive and lower their wages. The author of the study speculates that the insurance may have resulted in an excessive number of caesarean sections⁷⁶—indicating a thoroughly rational reaction by the women and the medical profession to the availability of a new funding source for obstetrical procedures.

Not the study itself, which does not mention behavioral economics, but the attempt to make it a poster child for behavioral economics, arbitrarily combines the premises of behavioral economics with those of rational-choice economics. The behavioralist construal necessarily assumes that employees are governed in their employment behavior by the endowment effect, but also that if before the childbirth insurance was mandatory employees had valued it at or above its cost, employers would have offered it without any prodding by government. The fact that they did not is taken to imply that the employees did not value it at more than its cost. But this is to assume that before the law was passed, employers and employees alike were perfectly rational. It is true that until then the endowment effect was not in play. But the endowment effect is only one of a number of irrationalities that behavioralists believe pervade labor markets. Why were none of them operating before the imposition of mandatory childbirth insurance?

Ward Farnsworth has found that parties to lawsuits do not recontract after the plaintiff has succeeded in obtaining an injunction.⁷⁷ The study is based on a sample of only 20 cases, which is too small to be statistically significant. If that point is put to one side, along with the fact that if the courts in his study “got it right”—that is, granted injunctions only in cases in which the plaintiff had more to gain from the injunction than the defendant had to lose from it—there would be no basis for a subsequent renegotiation, Farnsworth’s study suggests that once people have received a court judgment they are irrationally unwilling to negotiate with the opposing party. If so, this may seem to demonstrate the endowment effect operating at the heart of the legal system. But in fact it is not unusual for parties to settle a case after judgment in the trial court, rather than take their chances with an appeal.⁷⁸ And if a case that has become final through exhaustion of appellate remedies could have been settled, because the remedy sought by the plaintiff would cost the defendant more than it would benefit the plaintiff, the case would have been settled earlier—at the latest after the judgment in the trial court and before the appeal. Farnsworth’s findings vindicate rather than challenge rationality. Had the parties waited until all appeal rights had been exhausted before working out the value-maximizing resolution of their dispute—if final judgments turned out to be the prelimi-

nary to negotiations that undid them—it would mean that the parties had irrationally failed to economize on their expenses of litigation.

In support of the claim that the availability heuristic has given rise to “legislation by anecdote,” Jolls and her coauthors offer their own anecdote, about how the highly publicized rash of illnesses of people living near Love Canal gave rise to the Superfund law: “The behavioral account of Superfund is that the availability of ‘Love Canal’ as a symbol for the problem of abandoned hazardous waste dumps greatly intensified public concern, to the point where a legislative response became nearly inevitable, no matter what the actual facts might be.”⁷⁹ It is unclear what this narrative, plausible though it is, owes to behavioral economics or to any other organized body of thought.

They cite a study which hypothesized that each side in teacher collective-bargaining negotiations, in seeking to bolster its negotiating position with data on teachers’ wages in comparable communities, would “adopt self-serving judgments about which communities are ‘comparable,’ and impasses may result from such judgments.”⁸⁰ This is no more surprising than that each side in a lawsuit will make self-serving judgments about which cases provide the closest analogies to the case at hand or which facts are most probative. It is no answer that the strategic incentive to make self-serving judgments was eliminated in the collective-bargaining study by the fact that the only audience for the responses in the study was the study’s authors. Negotiators are unlikely to drop their (rational) biases when talking to professors, especially since they may lack confidence that their disclosures will remain confidential.

Jolls and her coauthors offer as examples of the operation of “fairness” thinking in the law usury laws; the avoidance of price gouging, which is not a legal imposition but a presumably compensated buffer of the risks faced by customers; and laws against ticket scalping, which are in force in fewer than half the states and which coexist mysteriously with laws permitting ticket brokers to buy in bulk from the theater and resell at “scalpers” prices. These are all claimed to be cases in which sellers are reined in from charging what the market will bear by a concept of fairness.

Usury laws do not support the claim, because there is no “reference point” interest rate and therefore no benchmark for triggering the sense of indignation that is the relevant component of the behavioralists’ concept of fairness. Lenders do not typically refuse to lend to risky borrowers at above-market rates, whatever exactly “market” means in this context. Banks quote a prime rate, but not necessarily the same rate, to their best customers and charge everyone else—that is, the riskier borrowers—more. (All a “prime rate” means is the bank’s best interest rate.) Mortgage lenders

charge varying numbers of points. Bonds, a form of loan, are rated for risk, and the lower-rated bonds pay higher interest rates without anyone crying “usury!” Credit card interest is much higher than bank-loan interest. Long-term interest rates usually differ from short-term rates. Interest rates on secured loans are lower than those on unsecured loans. Interest rates fluctuate with inflation, and of course with the demand for and the supply of capital. Even in consumer credit transactions, the focus of modern usury laws, there is no uniformity in interest rates, as many of my examples show. And can these laws have *any* effect today, when one considers that credit card and installment credit interest rates often approach 20 percent yet are perfectly lawful?

What is true is that if a borrower has a *really* high risk of default, there may be no interest rate that will make the loan worthwhile to either lender or borrower. This is especially likely because the higher the interest rate, the greater the risk of default, since an interest rate is a fixed rather than a variable cost of the borrower. This is why the riskiest loans and resulting astronomical interest rates are the domain of the loan shark, who, facing an unusually high risk of default, employs the threat of force in lieu of the milder remedies that are all that are available to the legal lender.

The same concept of fairness that is said to explain usury laws, price-gouging laws, and ticket-scalping laws is also said to explain laws forbidding prostitution and refusing to enforce surrogate-motherhood contracts, laws forbidding the sale of body parts and political votes, and laws refusing to enforce the contracting around of laws against race and sex. This is a heterogeneous collection of laws indeed, and to refer them to “pervasive judgments about fairness”⁸¹ is not to explain them. Behavioralists must explain what all these laws have in common, must give some form and content to their idea of “fairness,” and must consider more carefully the possibility of competing explanations for the laws, such as that they serve politically powerful special interests or are a product of misunderstandings unrelated to any of the three “boundeds.” For example, limits on adoption prices are supported by nonprofit adoption agencies, which are concerned about competition from profit-making adoption agencies, and by public ignorance of the consequences of price ceilings.

There have been other efforts to explain legal rules as responses to the insights of behavioralism. An article on cognitive biases of judges argues that Rule 407 of the Federal Rules of Evidence, which forbids use of subsequent remedial measures as evidence that the accident which they would have prevented was caused by the defendant’s negligence, and the common law rule that compliance with custom is a defense to medical malpractice, are both designed to combat hindsight bias.⁸² This is implau-

sible. The obvious office of Rule 407, and the justification traditionally offered for it, is to enhance safety by removing the disincentive to taking corrective measures that would exist if the taking of those measures could be used to show that the failure to take them earlier had been negligence. As for the defense of custom, the logic of the behavioralist argument is that compliance with custom should be a defense in all negligence cases; nothing in behavioralism suggests that it be confined to medical malpractice.

Normative Implications of Behavioralism

I have been focusing on the significance of behavioral economics for positive analysis, but I wish to consider briefly its possible normative implications. On the one hand, the picture of the human being that behavioralists draw is one of unstable preferences and (what turns out to be related) infinite manipulability. If you give a worker childbirth coverage, she'll like it (endowment effect); but if you don't give it to her, she'll dislike it (more precisely, won't want to pay for it in lower wages). If you force her to take an HIV test, she'll thank you afterwards, even though she fought kicking and screaming against having to take it. If you describe the threat of breast cancer to a woman in one way, she'll want a mammogram, but if you describe it in another although logically equivalent way, she won't. It seems then that the politically insulated corps of experts that many behavioralists (especially law professors) favor would be charged with determining the populace's authentic preferences, which sounds totalitarian to me. On the other hand, nothing in the behavioralist analysis exempts "experts" from the cognitive quirks, from weakness of will, or from concerns with fairness. The expert, too, is behavioral man. Behavioral man behaves in unpredictable ways. Dare we vest responsibility for curing irrationality in the irrational? Have we then a standoff?

I do not contend that the alternative to the behavioralist picture of man is a simple-minded libertarianism or individualism which claims that leaving people "free to choose" will always yield optimal results because people are always competent judges of their self-interest. Rather there is a spectrum of competence depending on the nature of the choice to be made. A policy of freedom to choose works best in giving people what they want in areas in which people have clear preferences, the preferences vary across individuals so that it is hard for the government to ascertain them, mapping from choice to utility is easy, the cost of obtaining the information necessary to make a sensible choice is low, and it is difficult for a stranger to determine one's preference or how a particular choice will achieve it. The rational-choice economist need not be dogmatic about what choices to leave to the individual and what choices to assign to government.

One might have thought that behavioral economics had at least one

clear normative implication: that efforts should be made through education and perhaps psychiatry to cure the cognitive quirks and weakness of will that prevent people from acting rationally with no offsetting gains. Even if as I believe the sunk-costs fallacy has biological roots, it should not be impossible to educate people out of it. Behavior therapy has enabled many people to overcome their fear of flying, which I suspect has more tenacious biological roots. Behaviorists treat the irrationalities that form the subject matter of behavioral economics as unalterable constituents of human personality. Their suggestions for legal reform are of devices for getting around, rather than dispelling, our irrational tendencies.

Consider Professor Rachlinski's reply to an article by Robert Hillman critical of the behavioralist approach.⁸³ Rachlinski argues for a behavioralist interpretation of the rule of contract law that penalty clauses are unenforceable: parties to contracts are excessively optimistic about their ability to perform and therefore underestimate the probability of having actually to pay the penalty for nonperformance. Even if the psychological premise is correct, that people are prone to overoptimism, the conclusion does not follow, at least with regard to enterprises (a qualification that Rachlinski does not suggest). The problem is self-correcting, without need for legal doctrine. Enterprises that find themselves paying heavy penalties for failure to perform their contracts will either be driven out of business or refuse to agree to penalty clauses in their contracts. Rachlinski's failure even to consider this possibility is symptomatic of a general failure of behavioral law and economics to generate feasible proposals for legal reform.

ENDNOTES

¹ See, for example, Christine Jolls, Cass R. Sunstein, and Richard Thaler, "A Behavioral Approach to Law and Economics," 50 *Stanford Law Review* 1471 (1998), reprinted in *Behavioral Law and Economics* 13 (Cass R. Sunstein ed. 2000); the other essays in that book; Jolls, Sunstein, and Thaler, "Theories and Tropes: A Reply to Posner and Kelman," 50 *Stanford Law Review* 1593 (1998); Sunstein, "Behavioral Law and Economics: A Progress Report," 1 *American Law and Economics Review* 115 (1999); Jon D. Hanson and Douglas A. Kysar, "Taking Behavioralism Seriously: The Problem of Market Manipulation," 74 *New York University Law Review* 630 (1999). For an early discussion of the implications of behavioral economics for economic analysis of law, see Thomas S. Ulen, "Cognitive Imperfections and the Economic Analysis of Law," 12 *Hamline Law Review* 385 (1989). For a comprehensive discussion of the underlying psychological literature, see Eldar Shafir and Robyn A. LeBoeuf, "Rationality," 53 *Annual Review of Psychology* 491 (2002).

² Jeffrey J. Rachlinski, "The 'New' Law and Psychology: A Reply to Critics, Skeptics, and Cautious Supporters," 85 *Cornell Law Review* 739 (2000). "Law and psychology" is a better term, as it avoids confining the new field to a reaction against economic analysis of law. Then too, "behavioral" and "behavioralism" are misleading because of the association of these terms with B. F. Skinner's anti-mentalistic psychology remote from contemporary "behavioralism."

³ See, for example, Richard A. Posner, *Frontiers of Legal Theory*, chs. 11–12 (2001); Posner, *Law, Economics, and Pragmatism*, ch. 5 (forthcoming).

⁴ These are numerous and severe, and I merely touch on a few of them. For a comprehensive analysis, see Gregory Mitchell, "Taking Behavioralism Too Seriously? The Unwarranted Pessimism of the New Behavioral Analysis of Law" (forthcoming in *William and Mary Law Review*). I draw heavily on Mitchell's article in this paper.

⁵ See note 1 above.

⁶ See Richard A. Posner, *Aging and Old Age* 148–152 (1995).

⁷ Adam Hirsch, "Lawmaking and Bounded Rationality" (Florida State University Law School, n.d.).

⁸ See, for example, Thomas C. Schelling, "Self-Command in Practice, in Policy, and in a Theory of Rational Choice," 74 *American Economic Review Papers and Proceedings* 1 (May 1984); Richard A. Posner, "Are We One Self or Multiple Selves? Implications for Law and Public Policy," 3 *Legal Theory* 23 (1997). For a similar approach by one of the pioneers in behavioral decision theory, see George Ainslie, "Will as Intertemporal Bargaining: Implications for Rationality" (forthcoming in *University of Pennsylvania Law Review*).

⁹ The ability of rational-choice economics, as in Gary S. Becker and Kevin M. Murphy, *Social Economics: Market Behavior in a Social Environment* (2000), to explain at least some of the phenomena with which behavioral economics is concerned is acknowledged in Christine Jolls's interesting paper "'Social-Behavioral' Economics," 5 *Green Bag* (2d ser.) 321 (2002).

¹⁰ Paul H. Rubin, *Darwinian Politics: The Evolutionary Origins of Freedom* 158 (2001).

¹¹ Id. His discussion is based on Gerd Gigerenzer, Peter M. Todd, and ABC [Center for Adaptive Behavior and Cognition] Research Group, *Simple Heuristics That Make Us Smart*, ch. 6 (1999). See also *Bounded Rationality: The Adaptive Toolbox* (G. Gigerenzer and R. Selten eds. 2001). Throughout *Darwinian Politics*, which I shall cite frequently in this paper, Rubin relies heavily on the work by Gigerenzer and his associates, who argue, in opposition to the emphasis that other cognitive psychologists place on the fallacious reasoning to which our cognitive biases give rise, that the various cognitive shortcuts (“heuristics”) that we use because of our cognitive limitations (“bounded rationality”) are generally well adapted to the real-world problems for which we use them. See, in addition to *How Simple Heuristics Make Us Smart*, Peter M. Todd and Gerd Gigerenzer, “Précis of *Simple Heuristics That Make Us Smart*,” 23 *Behavioral and Brain Sciences* 727 (2000). A related literature, based in psychology but also in philosophy, considers the emotions as an alternative method of cognition to conscious reasoning, a method that in many settings is more efficient. I discuss and review this literature in *Frontiers of Legal Theory*, footnote on p. 1, ch. 7.

¹² See Rubin, note 10 above, at 178–179.

¹³ Matthew Rabin, “Psychology and Economics,” 36 *Journal of Economic Literature* 11, 36 (1998).

¹⁴ See Gary S. Becker, Michael Grossman, and Kevin M. Murphy, “Rational Addiction and the Effect of Price on Consumption,” 81 *American Economic Review Papers and Proceedings* 237 (May 1991), reprinted in Becker, *Accounting for Tastes* 77 (1986).

¹⁵ See Tomas J. Philipson and Richard A. Posner, “Sexual Behaviour, Disease, and Fertility Risk,” 1 *Risk Decision and Policy* 91 (1996).

¹⁶ Gary S. Becker, “Irrational Behavior and Economic Theory,” 70 *Journal of Political Economy* 1 (1962), reprinted in Becker, *The Economic Approach to Human Behavior*, ch. 5 (1976).

¹⁷ See Mitchell, note 4 above.

¹⁸ With the partial exception of jury service—which is a reason both to repose greater confidence in the results of mock-jury studies, as I do in chapter 12 of my book *Frontiers of Legal Theory*, footnote on p. 1, and to worry that cognitive quirks may indeed distort jury decisionmaking.

¹⁹ See Daniel Houser, Michael Keane, and Kevin McCabe, “Behavior in a Dynamic Decision Problem: Evidence from the Laboratory” (George Mason University Dept. of Economics, Dec. 5, 2001, unpublished).

²⁰ Illustrated by a study in which college economics students were found to behave more in accordance with the rational-choice model than noneconomist students in playing the ultimatum game—and the authors attribute the difference to selection rather than to learning. See John R. Carter and Michael D. Irons, “Are Economists Different, and If So, Why?” *Journal of Economic Perspectives*, Spring 1991, p. 171.

²¹ See Christine Jolls, “Fairness, Minimum Wage Law, and Employee Benefits,” 77 *New York University Law Review* 77 (2002).

²² It is the implicit modern liberal conception of the average person—good, but inept, and for both reasons not very responsive to incentives, though perhaps rather plastic. The implicit conservative view of the average person, in contrast, is that he is competent but bad; hence conservatives place emphasis on incentives and constraints.

²³ To repeat an earlier point, he need only be rational with respect to the particular choice confronting him. People who are frightened about flying are assumed to respond rationally to changes in airline ticket prices, even though it is difficult to give a rational account of their fears.

²⁴ See, for example, James Konow, "Which Is the Fairest One of All? A Positive Analysis of Justice Theories" (forthcoming in *Journal of Economic Literature*).

²⁵ Though this seems to be changing. See, for a notable contribution, Adam Gifford, Jr., "The Role of Emotions and Language in Choice" (California State University Dept. of Economics, July 2002, unpublished).

²⁶ Even so, one is surprised to read in Jolls, Sunstein, and Thaler, "Theories and Tropes: A Reply to Posner and Kelman," note 1 above, at 1600, that "evolutionary biology is an unfalsifiable theory." This is the kind of thing one is more likely to hear from creationists than from liberals. They quickly backtrack, however, acknowledging "that much of the behavior we discuss can be considered a vestige of our long held instincts." Id.

²⁷ So, other things being equal, having three nephews (each a 25 percent genetic copy of you) will contribute more to your inclusive fitness than having one child (a 50 percent genetic copy of you). The qualification ("other things being equal") is vital. If your three nephews were much less likely to survive to reproductive age than the one child, they would contribute less, at least on an expected basis, to your inclusive fitness than the child.

²⁸ See, for example, Susan M. Essock-Vitale and Michael T. McGuire, "Predictions Derived from the Theories of Kin Selection and Reciprocation Assessed by Anthropological Data," 1 *Ethology and Sociobiology* 233 (1980).

²⁹ Cf. Charles J. Morgan, "Natural Selection for Altruism in Structured Populations," 6 *Ethology and Sociobiology* 211 (1985); Morgan, "Eskimo Hunting Groups, Social Kinship, and the Possibility of Kin Selection in Humans," 1 *Ethology and Sociobiology* 83 (1979). This analysis is questioned, however, in Allan Gibbard, *Wise Choices, Apt Feelings: A Theory of Normative Judgment* 258 n. 2 (1990).

³⁰ Rubin, note 10 above, at 173–174.

³¹ See Cristina Bicchieri, "Learning to Cooperate," in *The Dynamics of Norms* 17, 39 (Cristina Bicchieri, Richard Jeffrey, and Brian Skyrms eds. 1997); Oded Stark, *Altruism and Beyond: An Economic Analysis of Transfers and Exchanges within Families and Groups* 132 (1995). Generalization (less grandly, pattern recognition) seems an innate, and very valuable but of course fallible, capacity of the human animal.

³² See, for example, Bernard Williams, "Moral Luck," in his book *Moral Luck: Philosophical Papers 1973–1980* 20 (1981); also Williams, "Moral Luck: A Postscript," in his book *Making Sense of Humanity, and Other Philosophical Papers 1982–1993* 241 (1995).

³³ Consistent with this suggestion, an experimental study found that "proposers [in the ultimatum game] do not want to be fair, but rather want to appear fair, in order to prevent the responder from rejecting the offer." Werner Güth and Eric van Damme, "Information, Strategic Behavior, and Fairness in Ultimatum Bargaining: An Experimental Study," 42 *Journal of Mathematical Psychology* 227, 242 (1998). My interpretation is further supported by the fact that players in the ultimatum game who are told that offers are randomly generated (and so do not emanate from a potentially competitive human being) act as classic utility maximizers. See Sally Blount, "When Social Outcomes Aren't Fair: The Effect of Causal Attributions on Preferences," 63 *Organizational Behavior and Human*

Decision Processes 131, 135 (1995); Colin Camerer and Richard H. Thaler, "Anomalies: Ultimatums, Dictators and Manners," *Journal of Economic Perspectives*, Spring 1995, pp. 209, 215.

³⁴ One would like to see a series of ultimatum-game experiments in which the proposers make the same offers to respondents who differ both among themselves and from the proposers in age, sex, income, and education, viewed as proxies for or sources of differences in status, self-esteem, or other plausible correlates of the sense of pride that causes respondents in the game to reject chintzy offers. We might learn how closely the ultimatum game corresponds to status struggles among chimpanzees and other monkeys who resemble our proto-human ancestors.

³⁵ See Jon Elster, "Emotions and Economic Theory," 36 *Journal of Economic Literature* 47, 68 (1998); Jack Hirshleifer, "The Emotions as Guarantors of Threats and Promises," in *The Latest on the Best: Essays in Evolution and Optimality* 307 (John Dupr ed. 1987)..

³⁶ L. G. Telser, "The Ultimatum Game and the Law of Demand," 105 *Economic Journal* 1519 (1995).

³⁷ Daniel Kahneman, Jack L. Knetsch, and Richard Thaler, "Experimental Tests of the Endowment Effect and the Coase Theorem," 98 *Journal of Political Economy* 1325 (1990).

³⁸ See, for example, Owen D. Jones, "The Evolution of Irrationality," 41 *Jurimetrics* 289 (2001); Jones, "Science and Human Behavior: A Reply," id. at 371; Jones, "Time-Shifted Rationality and the Law of Law's Leverage: Behavioral Economics Meets Behavioral Biology," 95 *Northwestern University Law Review* 1141 (2001).

³⁹ See Rubin, note 10 above, at 156–181.

⁴⁰ Id. at 163.

⁴¹ Id. at 168.

⁴² Id. This brings to mind a quotation attributed to Stalin: "One death is a tragedy; one million deaths is a statistic."

⁴³ Rubin, note 10, at 156–158. This is a central thesis of Gigerenzer's work. See note 10 above. To illustrate, if you tell a person that he has a chance of 1 in 1000 of catching a fatal disease, this is likely to impress him much more than if you tell him that he has a probability of one-tenth of 1 percent of catching the disease; for the 1-in-a-1000 formulation brings vividly to the hearer's mind that some people do get the disease. Conversely, if you tell a person that eating sardines will increase his chance of contracting a rare disease by 25 percent, you will frighten him much more than if you tell him that eating sardines will increase his chance of contracting the disease from 4 in 4 million to 5 in 4 million. In both cases, stating the "odds" in frequency rather than probability terms is more meaningful to the average person.

⁴⁴ See, for example, Paul Slovic, *The Perception of Risk* (2000).

⁴⁵ The analysis that follows is drawn from Richard A. Posner, *Economic Analysis of Law* 20, 95–96 (5th ed. 1998).

⁴⁶ On the economics of preference formation, see generally Gary S. Becker, *Accounting for Tastes* (1996).

⁴⁷ Amartya Sen, "The Discipline of Cost-Benefit Analysis," 29 *Journal of Legal Studies* 931, 949 (2000).

⁴⁸ See Daniel S. Levy and David Friedman, "The Revenge of the Redwoods? Reconsidering Property Rights and the Economic Allocation of Natural Resources," 61 *University of Chicago Law Review* 493 (1994).

⁴⁹ Gary S. Becker, "Habits, Addictions, and Traditions," 45 *Kyklos* 327, 336 (1992), reprinted in Becker, note 46 above, at 118; Marcel Boyer, "Rational Demand and Expenditures Pattern under Habit Formation," 31 *Journal of Economic Theory* 27 (1983). The obverse case—cost positively, benefit negatively, related to time—is that of boredom.

⁵⁰ This methodological criticism of behavioral economics is, it should be noted, limited to experiments designed to test for "fairness" or other moralistic attitudes, not to experiments concerned with cognitive defects or weakness of will. I should further note that experimental economics and behavioral economics should not be equated; experiments are increasingly used as a method of empirical research by economists committed to the rational-choice approach.

⁵¹ George Loewenstein and Samuel Issacharoff, "Source Dependence in the Valuation of Objects," 7 *Journal of Behavioral Decision Making* 157 (1994).

⁵² Richard Thaler, "Toward a Positive Theory of Consumer Choice," 1 *Journal of Economic Behavior and Organization* 39, 43 (1980).

⁵³ *Id.* at 43.

⁵⁴ *Id.* at 43–44.

⁵⁵ *Id.* at 52.

⁵⁶ *Id.*

⁵⁷ Jack L. Knetsch and J. A. Sinden, "Willingness to Pay and Compensation Demanded: Experimental Evidence of an Unexpected Disparity in Measures of Value," 99 *Quarterly Journal of Economics* 507 (1984).

⁵⁸ *Id.* at 513 n. 7.

⁵⁹ Maya Bar-Hillel and Efrat Neter, "Why Are People Reluctant to Exchange Lottery Tickets?" 70 *Journal of Personality and Social Psychology* 17, 23–24 (1996). Mitchell, note 4 above, points out that most behavioralist experiments are designed to elicit irrational responses, and he presents evidence of "cuing" by experimenters. He also points out that studies which fail to elicit such response are less likely to be published; the result is a misleading impression of the robustness of the experimental evidence of irrationality. This is a general problem in economics, perhaps in social science generally; results that contradict the author's hypothesis are less likely to be published than those that are consistent with it.

⁶⁰ See James C. Cox and Cary A. Deck, "On the Nature of Reciprocal Motives" 16, 18 (University of Arizona Department of Economics and University of Arkansas Walton School of Business, July 2002, unpublished).

⁶¹ Uri Gneezy, "Deception: The Role of Consequences" (University of Chicago Graduate School of Business, May 22, 2002, unpublished).

⁶² Another possibility is habitual behavior. The dictator game resembles tipping, see Bradley Ruffle, "More Is Better, But Fair Is Fair: Tipping in Dictator and Ultimatum Games," 23 *Games and Economic Behavior* 247, 258 (1998), a social norm "enforced" by the dirty looks that a waiter will give a person who fails without cause to leave a tip within the

customary range.

⁶³ Elisabeth Hoffman, Kevin McCabe, and Vernon L. Smith, "Social Distance and Other-Regarding Behavior in Dictator Games," 86 *American Economic Review* 653 (1996).

⁶⁴ Id. at 656.

⁶⁵ Güth and van Damme, note 33 above, at 243.

⁶⁶ Gary E. Bolton and Rami Zwick, "Anonymity versus Punishment in Ultimatum Bargaining," 10 *Games and Economic Behavior* 95 (1995). See also Iris Bohnet and Bruno S. Frey, "Social Distance and Other-Regarding Behavior in Dictator Games: Comment," 89 *American Economic Review* 335 (1999); Bohnet and Frey, "The Sound of Silence in Prisoner's Dilemma and Dictator Games," 38 *Journal of Economic Behavior and Organization* 43 (1999); Catherine C. Eckel and Philip J. Grossman, "Are Women Less Selfish than Men? Evidence from Dictator Experiments," 108 *Economic Journal* 726 (1998); Duncan K. H. Fong and Gary E. Bolton, "Analyzing Ultimatum Bargaining: A Bayesian Approach to the Comparison of Two Potency Curves under Shape Constraints," 15 *Journal of Business and Economic Statistics* 335 (1997).

⁶⁷ See Catherine C. Eckel and Philip J. Grossman, "Altruism in Anonymous Dictator Games," 16 *Games and Economic Behavior* 181 (1996). Another general problem with using students as the subjects of experiments designed to test for a sense of fairness is that young people tend to be more idealistic than the population as a whole.

⁶⁸ For evidence of this, which further supports a signaling interpretation of the results of the behavioral experiments, see Kevin A. McCabe, Stephen J. Rassenti, and Vernon L. Smith, "Reciprocity, Trust, and Payoff Privacy in Extensive Form Bargaining," 24 *Games and Economic Behavior* 10 (1998).

⁶⁹ Camerer and Thaler, note 33 above, at 216.

⁷⁰ See note 33 above.

⁷¹ A term I use to distinguish *interest* rates, which are a function not only of time preferences, default risk, and administrative costs but also of the supply of capital, from nonmonetary discount rates. Interest rates might be high not because people had a strong preference for present over future consumption but because capital was scarce for unrelated reasons.

⁷² See Becker, note 46 above, at 11–12; Becker and Casey B. Mulligan, "On the Endogenous Determination of Time Preference," 112 *Quarterly Journal of Economics* 729 (1997).

⁷³ Jonathan Gruber, "The Incidence of Mandated Maternity Benefits," 84 *American Economic Review* 622 (1994).

⁷⁴ If they didn't, the employer would be forced to "swallow" some of the additional cost of the program.

⁷⁵ It is unclear from Gruber's article what exceptions if any the state laws that were his principal subject made in coverage. But even if all employed people were fully covered, the laws would tend to attract women of childbearing age and intentions from the home into paid employment.

⁷⁶ Gruber, note 73 above, at 640.

⁷⁷ Ward Farnsworth, "Do Parties to Nuisance Cases Bargain after Judgment? A Glimpse inside the Cathedral," 66 *University of Chicago Law Review* 373 (1999), reprinted in

Behavioral Law and Economics, note 1 above, at 302.

⁷⁸ Federal courts of appeals have settlement officers to facilitate the settlement of cases on appeal. The settlement offices, usually of two or three officers, are expected to reduce the judges' caseload by about 10 percent, Richard A. Posner, *The Federal Courts: Challenge and Reform* 239–240 (1996), though this is probably overoptimistic. Many cases are settled on appeal without the help of the settlement officers, even though the cost savings from such a settlement are limited, since most of the expenses of litigation will have been incurred already. About a quarter of all cases filed in the federal courts of appeals are disposed of, before full briefing, without judicial action. Computed from *id.* at 72 (tab. 3.6). An unknown, but not trivial, percentage of these are settled, along with a very small percentage of cases briefed and argued but not yet decided.

⁷⁹ Jolls, Sunstein, and Thaler, "A Behavioral Approach to Law and Economics," note 1 above, at 1521.

⁸⁰ *Id.* at 1502.

⁸¹ *Id.* at 1516.

⁸² Chris Guthrie, Jeffrey J. Rachlinski, and Andrew J. Wistrich, "Inside the Judicial Mind," 86 *Cornell Law Review* 777, 828 (2001).

⁸³ Rachlinski, note 2 above, replying to Robert A. Hillman, "The Limits of Behavioral Decision Theory in Legal Analysis: The Case of Liquidated Damages," 85 *Cornell Law Review* 717 (2000).

AIER PUBLICATIONS CURRENTLY AVAILABLE

<i>Personal Finance</i>	<i>Prices</i>
THE A-Z VOCABULARY FOR INVESTORS (ISBN 0-913610-02-X)	\$ 7.00
COIN BUYER'S GUIDE	10.00
COPING WITH COLLEGE COSTS	6.00
THE ESTATE PLAN BOOK 2000—with 2001 supplement (ISBN 0-913610-13-5)	10.00
HOMEOWNER OR TENANT? How To Make A Wise Choice (ISBN 0-913610-23-2)	8.00
HOW SAFE IS YOUR BANK?	8.00
HOW TO AVOID FINANCIAL FRAUD	3.00
HOW TO AVOID FINANCIAL TANGLES (ISBN 0-913610-19-4)	8.00
HOW TO BUILD WEALTH WITH TAX-SHELTERED INVESTMENTS (ISBN 0-913610-14-3)	6.00
HOW TO COVER THE GAPS IN MEDICARE Health Insurance and Long-Term Care Options for the Retired (ISBN 0-913610-04-6)	5.00
HOW TO INVEST WISELY (ISBN 0-913610-22-4)	12.00
HOW TO MAKE TAX-SAVING GIFTS (ISBN 0-913610-10-0)	3.00
HOW TO PLAN FOR YOUR RETIREMENT YEARS	6.00
HOW TO PRODUCE SAVINGS IN THE ADMINISTRATION OF AN ESTATE	3.00
HOW TO READ A FINANCIAL STATEMENT	9.00
HOW TO USE CREDIT WISELY	6.00
INFLATION OR DEFLATION: What Is Coming?	6.00
INTERNATIONAL INVESTING: Theory, Practice, and Results	5.00
LIFE INSURANCE FROM THE BUYER'S POINT OF VIEW	8.00
SENSIBLE BUDGETING WITH THE RUBBER BUDGET ACCOUNT BOOK	5.00
WHAT WILL RECESSION MEAN TO YOU? (ISBN 0-913610-21-6)	6.00
WHAT YOU NEED TO KNOW ABOUT MUTUAL FUNDS (ISBN 0-913610-18-6) .	6.00
WHAT YOUR CAR REALLY COSTS: How to Keep a Financially Safe Driving Record (ISBN 0-913610-20-8)	6.00
 <i>Economic Fundamentals</i>	
THE AIER CHART BOOK (ISBN 0-913610-17-8)	4.00
BREAKING THE BANKS: Central Banking Problems and Free Banking Solutions	12.00
CAUSE AND CONTROL OF THE BUSINESS CYCLE	6.00
THE COLLAPSE OF DEPOSIT INSURANCE	4.00
FORECASTING BUSINESS TRENDS (ISBN 0-913610-16-X)	5.00
GOLD AND LIBERTY	8.00
KEYNES vs. HARWOOD — A CONTRIBUTION TO CURRENT DEBATE	6.00
MONETARY POLICY IN DEMOCRACIES: Four Resumptions and the Great Depression	3.00
MONEY: Its Origins, Development, Debasement, and Prospects (ISBN 0-913610-11-9)	10.00
THE POCKET MONEY BOOK A Monetary Chronology of the United States	2.00
RECONSTRUCTION OF ECONOMICS	6.00
USEFUL ECONOMICS	6.00
 <i>General Interest</i>	
AMERICA'S UNKNOWN ENEMY: BEYOND CONSPIRACY	9.00
CAN OUR REPUBLIC SURVIVE? Twentieth Century Common Sense and the American Crisis	6.00
 <i>Behavioral Research Council Division of AIER</i>	
THE BEHAVIORAL SCIENCES: Essays in Honor of George A. Lundberg*	8.00
A CURRENT APPRAISAL OF THE BEHAVIORAL SCIENCES*	15.00
USEFUL PROCEDURES OF INQUIRY	15.00

Note: Educational discounts for classroom use are available for all of the above publications.

* Hardbound.

ECONOMIC EDUCATION BULLETIN

AMERICAN INSTITUTE FOR ECONOMIC RESEARCH
Great Barrington, Massachusetts 01230

Periodical postage paid at
Great Barrington, Massachusetts