2007

A Liberal Challenge to Behavioral Economics: The Case of Probability

Alex Stein
Brooklyn Law School

NYU JOURNAL OF LAW & LIBERTY

A LIBERAL CHALLENGE TO BEHAVIORAL ECONOMICS: THE CASE OF PROBABILITY

ALEX STEIN*

This paper argues that the Behavioral Economics movement (*BE* for short) needs to be liberally inclusive in formulating the criteria for rationality against which to evaluate people's cognitive ability and performance. It is illiberal to require that individuals use a single cognitive methodology where more than one sound methodology is available. This prescription is appropriate both for evaluating people's probabilistic reasoning and for studying people's utility valuations. This paper focuses primarily on BE's assessment of people's probabilistic performance. Specifically, it examines the famous "Blue Cab" experiment upon which several behavioral studies have grounded their assessment of an average person as "probabilistically challenged."

* Visiting Professor of Law, Yale Law School. Professor of Law, Cardozo Law School. LL.B., LL.M. Hebrew University of Jerusalem; Ph.D. University of London. E-mail: alex.stein@yale.edu. Website: www.professoralexstein.com. I thank Samuel Buell, David Carlson, Claire Hill and J.D. Trout for comments and suggestions.

Part I

BE challenges liberalism by questioning its "free agency" assumption that people are generally able to make rational choices as to what is right for them. This assumption postulates that, by and large, people perform reasonably well in maximizing their welfare and in determining facts needed for their decisions. This dual competence—economic and epistemic—comes along with two normative implications. First, government must not interfere with people's choices as to how to go about their lives, except for special reasons. If a person's mental capacity is not substantially impaired and his actions do not externalize harm upon others, the government should leave that person alone. Second, government must not impose upon people any rationality criteria for fact-determination. Facts upon which people proceed in their lives should be free from official dogma (an extreme version of which is portrayed by the "Ministry of Truth" in Orwell's dystopia[1]). By the same token—and coming closely to this paper's specific point—government must not condition one's designation as a juror in a trial upon any special intellectual qualifications. To properly function as a juror, a person need not be a logician, a statistician or a scientist. A lay person can adequately determine what is "beyond all reasonable doubt" and which factual allegations are "more probable than not."[2]

BE challenges the "free agency" assumption by empirical demonstrations that people systematically err. According to BE, people form value-preferences that are economically irrational and combine those preferences in ways that reduce their welfare, rather than augmenting it.[3] They also mishandle fact-finding tasks by

---

[1] *See* GEORGE ORWELL, 1984 7 (1949).

[2] *See* L. Jonathan Cohen, *Freedom of Proof*, 16 ARCHIV FÜR RECHTS-UND SOZIALPHILOSOPHIE 1, 21 (1983).

[3] *See, e.g.*, Christine Jolls, Cass R. Sunstein & Richard Thaler, *A Behavioral Approach to Law and Economics*, 50 STAN. L. REV. 1471 (1998); Russell B. Korobkin & Thomas S. Ulen, *Law and Behavioral Science: Removing the Rationality Assumption From Law and Economics*, 88 CAL. L. REV. 1051 (2000); Russell B. Korobkin, *Bounded Rationality, Standard Form Contracts, and Unconscionability*, 70 U. CHI. L. REV. 1203 (2003). See Richard A. Posner, *Rational Choice, Behavioral Economics and the Law*, 50 STAN. L. REV. 1551

breaking the most basic rules of probabilistic reasoning.[4] These cognitive failures are both systematic and harmful. Government, therefore, ought to step in and do what is right.

This, in a nutshell, is BE's challenge to liberalism. Liberalism elevates individual autonomy over authority, but an autonomous person's bounded rationality will bring him down. To prevent this downfall, the person clings to his benevolent government that keeps him up. He acts like a medical patient, putting his faith in a doctor's prescription of Prozac without investigating its chemistry and impact. This person subordinates his reason and decisions[5] to those that his authority (another person or agency) develops for him. He knows that insubordination is good for his autonomy—which happens to be an important liberal value—but he also knows that this celebration of autonomy would not last for long if he really needs Prozac but takes aspirin instead.

Part II

Liberalism, however, makes an additional point that BE fails to consider seriously. Under liberalism, rationality is multifaceted and diverse, rather than single-faceted and monistic. Crucial to the present debate, this pluralist vision originates from uncertainty rather than toleration. Liberalism does not merely tolerate people's diverse viewpoints as to what counts as rational, it also supports those viewpoints affirmatively by giving them equal epistemic credentials. Rawls's "overlapping consensus,"[6] for example, is not primarily a method for developing mutual respect among members

---

(1998), for criticism. See BEHAVIORAL LAW AND ECONOMICS (Cass R. Sunstein ed., 2000), for a good collection of essays documenting BE experiments and findings.

[4] *See* Amos Tversky & Daniel Kahneman, *Judgment under Uncertainty: Heuristics and Biases*, 185 SCIENCE 1124 (1974); *see also* JUDGMENT UNDER UNCERTAINTY: HEURISTICS AND BIASES (Daniel Kahneman, Paul Slovic & Amos Tversky eds., 1982).

[5] *See* Joseph Raz, *Authority and Justification*, 14 PHIL. & PUB. AFF. 3 (1985) (developing an account of authority based on preemptive reasons).

[6] *See* JOHN RAWLS, POLITICAL LIBERALISM 133-34, 145-53 (expanded ed., 2005).

of a pluralistic society (which is not a bad thing in itself). Rather, it is a rational preference-maximizing framework of "public reason."[7]

To have a concrete example from the area of economics, consider the famous "sunk cost" problem.[8] Under the standard economic model, in estimating the utility of her prospective endeavor, a rational person must look forward, not backward. Specifically, a person must ignore her "sunk costs." For example, after investing money and effort into acting, Jane ought to ignore this investment in deciding whether to study law and become an attorney. All she needs to consider is the difference between her net expected benefits from acting and from lawyering. In reality, however, people do account for their "sunk costs." According to BE, this "sunk cost fallacy" is a paradigmatic example of bounded rationality.[9]

But what if Jane considers her ongoing acting endeavor as important to the definition of her life?[10] For her, not abandoning an important endeavor and not succumbing to temptations of the moment determine what kind of person she is in the larger scheme of things.[11] This course of action is perfectly rational, although not rational in the narrow economic sense. If the narrow paradigm of rationality fails to account for this action, why treat it as a baseline for assessing people's performance as welfare maximizers? Why consider other forms of rationality as inadmissible? Why not adopt, in other words, the inclusive liberal position that perceives rationality as a complex and multifaceted phenomenon?

This, in a nutshell, is a liberal challenge to BE. This challenge has many important implications for designers of social policy. In what follows, I examine these implications in the domain of probabilistic reasoning.

---

[7] *Id.* at 150-53, 212-13.

[8] *See, e.g.*, Hal R. Arkes & Catherine Blumer, *The Psychology of Sunk Cost*, 35 ORG. BEHAV. & HUMAN DECISION PROCESSES 124, 126-36 (1985).

[9] *Id.*

[10] *See* ROBERT NOZICK, THE NATURE OF RATIONALITY 22 (1993).

[11] *See id.* at 23.

Part III

To make this discussion manageable, I focus on a single probabilistic issue: the famous "Blue Cab" experiment. According to BE, this experiment and its findings provide a paradigmatic example of people's bounded rationality in probabilistic reasoning.[12]

Factfinders considered a car accident that occurred in a city in which 85% of cabs were Green and the remaining 15% were Blue. The factfinders heard a witness testify that the cab involved in the accident was Blue, and then accepted as uncontested background fact that the testifying witness correctly identified cabs in 80 out of 100 cases. Factfinders participating in that experiment systematically failed to determine the probability of the victim's case against the Blue Cab Company.[13] Their typical estimation—80%—coincided with the given credibility of the witness,[14] but not with the basic rules of probabilistic calculus, which would take into account the relative frequency of Blue and Green cabs as well. Under these rules, the prior odds attaching to the scenario in which the cab involved in the accident was Blue rather than Green—P(B)/P(G)—equaled 0.15/0.85. To calculate the posterior odds—P(B | W)/P(G | W), with W denoting the witness's testimony—these prior odds had to be multiplied by the likelihood ratio. This ratio had to be determined by the odds attaching to the scenario in which the witness identified the cab's color correctly, rather than incorrectly: P(W | B)/P(W | G). The posterior odds consequently equaled (0.15×0.8)/(0.85×0.2), that is, 12/17. The probability of the victim's allegation against the Blue Cab Company thus equaled 12/(17+12), that is, 41%—an outcome that fails to satisfy the preponderance-of-the-evidence standard that applies in civil litigation. This outcome therefore undermined the rationality of the factfinders' decision that the victim should prevail.[15]

---

[12] *See* Kahneman, Slovic & Tversky, *supra* note 4, at 156-57.
[13] *Id.* at 156-59.
[14] *Id.*
[15] *Id.*

This experiment is faulty in one crucial respect.[16] The critical datum that the factfinders had to process was unclear. Reportedly, "The court tested the reliability of the witness under the same circumstances that existed on the night of the accident and concluded that the witness correctly identified each one of the two colors 80% of the time and failed 20% of the time."[17] But did the court consider the prior odds in reaching this conclusion about the witness? Did it calculate, in other words, that "the witness correctly identified each one of the two colors 80% of the time and failed 20% of the time" *after* taking into account the distribution of Blue and Green cabs in the city? If the court did so, the factfinders' ascription of an 80% probability to the victim's case would then be obviously correct. This scenario would exhibit no probabilistic errors and no bounded rationality whatsoever.

This scenario, however, may be unduly optimistic. What if the court did not consider the prior odds in calculating the probability of the witness's correctness? Failure to consider these odds makes the factfinders' assessment of the probability of the victim's case defective—but then one should inquire about the reason for *not* telling the factfinders about the need to take the prior odds into account. The factfinders processed the witness's 80% credibility as invariant across all possible cab distributions. They assumed, in other words, that the distribution of Blue and Green cabs in the city was *causally irrelevant* to the witness's ability to identify colors correctly.[18]

This understanding of the evidence was not irrational. In formal terms, the factfinders assumed that $P(B|W)=0.8$. The fact that the prior probability—P(B) —was relatively low (0.15) does not make this assumption irrational, because the likelihood ratio may have been estimated as sufficiently high. This ratio may have been

---

[16] I have made this point briefly in ALEX STEIN, FOUNDATIONS OF EVIDENCE LAW 60-63 (2005).

[17] *See* Kahneman, Slovic & Tversky, *supra* note 4, at 156.

[18] *See* L. Jonathan Cohen, *Can Human Irrationality be Experimentally Demonstrated?*, 4 BEHAV. & BRAIN SCI. 317, 328-29 (1981).

high enough because, once again, the cab distribution in the city was causally irrelevant to the credibility of the witness's account. Be that as it may, bounded rationality was not the factfinders' problem. The problem was the experiment's design.

Having said this, I am still ready to assume that the experimenters did have a good reason for not telling the factfinders that the cab distribution was (somehow) causally relevant to the accuracy of the witness's account. Perhaps the factfinders were expected to come to this understanding by themselves. The factfinders then may have decided the case erroneously due to their failure to come to this understanding. The exact origin of the factfinders' error, however, is still crucial for determining what went wrong in their decision. The experimenters and their BE followers interpret this error as a probabilistic miscalculation. They claim that the factfinders failed to count the base rates in determining the probability of correctness that attached to the witness's account. They also argue, quite convincingly, that this cognitive phenomenon is widespread. This interpretation, however, is questionable at best because the factfinders did not miscalculate any *probability*. Their mistake (if they committed one) was epistemic rather than operational. The mistake was the factfinders' decision to cast away the information about cab distribution as causally irrelevant to the witness's accuracy. Under the irrelevancy assumption, the prior odds of the victim's case were 1/1. But if the cab-distribution was in fact relevant to the accuracy of the witness's account—a possibility that the factfinders evidently failed to consider—these prior odds could then be 0.15/0.85. The posterior probability of the victim's case would consequently decline.

This understanding is the most that BE can make of the "Blue Cab" experiment. If so, the problem that BE has identified is not systematic and can be fixed. The "Blue Cab" and similar findings are about people's cognitive *performance*, rather than cognitive *competence*[19]—a crucial distinction that BE systematically fails to

---

[19] *Id.*

consider. The point of the distinction is this: people remain cognitively competent even when they err. This means that people can avoid errors in probability computations (or other decisions) when properly alerted to those errors. Admittedly, people often fall into traps set by BE experimenters in order to test their rationality. These traps, however, can only function as the conjurer's sleight of hand: each trick can be played only once. As in the teleological paradox,[20] the play uncovers and thereby destroys the trick.

Part IV

As I said before, the factfinders' "mistake" may not have been a mistake after all. The factfinders knew nothing about the correlation (let alone the causation) between cab distribution in the city and the accuracy of cab-identifying witnesses. They therefore had an epistemological warrant to believe the witness by assessing the accuracy of his cab identification as 80% probable. This assessment was as rational as one that does take the cab distribution into account. The choice between the two approaches depends on how one wants to allocate the risk of error; there is more than one rational way of doing it.[21]

Furthermore, as Jonathan Cohen pointed out, the factfinders may have been focusing on the Baconian case-specific probability.[22] That is, they may have been assessing the strength of the evidential support for the victim's case-specific allegation that "this witness in this particular instance identified the cab's color cor-

---

[20] This paradox looms in ethics and economics. Targeting a particular goal (such as esteem or friendship) is sometimes counterproductive to its attainment—a predicament that places the goal out of reach. *See* T. M. SCANLON, WHAT WE OWE TO EACH OTHER 383 n.15 (1998); Zev Trachtenberg, Book Review, 115 ETHICS 809, 810 (2005) (reviewing GEOFFREY BRENNAN & PHILIP PETIT, THE ECONOMY OF ESTEEM: AN ESSAY ON CIVIL AND POLITICAL SOCIETY (2004)) (observing that "people who strive for esteem will be disesteemed on that account.").

[21] *See* STEIN, *supra* note 16, at 40-56; *see also id.* at 107-40 (arguing that evidence law is generally about apportionment of the risk of error).

[22] *See* Cohen, *supra* note 18, at 328-30.

rectly."[23] The experimenters assumed that the factfinders' only goal was to calculate the witness's chances to have testified correctly. Their frame of reference was Pascalian (aleatory) probability, which happens to be one of the two basic modes of probabilistic reasoning. The other mode is Baconian (inductivist) probability. These two modes are logically distinct from each other. As Jonathan Cohen explains,

> Baconian probability-functions . . . deserve a place alongside Pascalian ones in any comprehensive theory of nondemonstrative inference, since Pascalian functions grade probabilification *on the assumption that* all relevant facts are specified in the evidence, while Baconian ones grade it *by the extent to which* all relevant facts are specified in the evidence.[24]

The experimenters' choice of the Pascalian mode would have been absolutely appropriate had they put the factfinders "on the same page." The experimenters' failure to do so is another flaw in their experiment's design.[25]

Part V

BE offers an array of fruitful methodologies for testing the rationality of people's reasoning. Consonantly with the liberal tradition, however, BE ought to recognize that people's rationality, in

---

[23] *Id.*

[24] L. Jonathan Cohen, *On the Psychology of Prediction: Whose is the Fallacy?*, 7 COGNITION 385, 389 (1979). For detailed accounts of Baconian probability see L. JONATHAN COHEN, THE PROBABLE AND THE PROVABLE 121-216 (1977); L. JONATHAN COHEN, AN INTRODUCTION TO THE PHILOSOPHY OF INDUCTION AND PROBABILITY 4-13, 145-75 (1989). For its implications on adjudicative fact-finding and evidence law, see STEIN, *supra* note 16, at 64-106.

[25] *See* Cohen, *supra* note 18, at 325, 330; *see also* Charles M. Yablon, *The Meaning of Probability Judgments: An Essay on the Use and Misuse of Behavioral Economics*, 2004 U. ILL. L. REV. 899 (demonstrating that multiplicity of probability concepts often frustrates cognitive theories that attribute probabilistic irrationality to lay reasoners).

both economic and epistemic areas, is a complex, dynamic and multifaceted phenomenon. There are many different rationalities, and BE's experimental designs need to take this pluralism into account. "Blue Cab" and similar experiments have failed to do so. As a result, these experiments do not prove that people systematically misjudge probabilities.

BE also needs to separate cognitive performance from cognitive competence. Cognitive performance depends upon trial and error and upon the information that people have and do not have. As such, it can be improved through direct learning (as opposed to indirect debiasing techniques that counter cognitive incompetence). Cognitive competence is an immanent human condition that a simple acquisition of information cannot undo. "Blue Cab" and similar experiments are about cognitive performance, not competence. They do not establish people's cognitive incompetence in the probabilistic domain.