

Altruism, Righteousness, and Myopia
T. Clark Durant and Michael Weintraub*

Published in *Critical Review* 23, 3: 257–302.

January 10, 2012

Abstract: Twenty years ago Leif Lewin made the case that altruistic motives are more common than selfish motives among voters, politicians, and bureaucrats. We propose that motives and beliefs emerge as reactions to immediate feedback from technical-causal, material-economic, and moral-social aspects of the political task environment. In the absence of certain kinds of technical-causal and material-economic feedback, moral-social feedback leads individuals to the altruism Lewin documents, but also to righteousness (moralized regard for the in-group and disregard for the out-group) and myopia (disregard for distant consequences). The mix of altruism, righteousness, and myopia increases the focus on winning the next high-stakes election rather than on discovering or enforcing socially productive institutions.

* T. Clark Durant, formerly post-doctoral fellow, Department of Economics, New York University, New York, NY 10012, tc Durant@gmail.com, and Michael Weintraub, Department of Government, Georgetown University, Washington, DC 20057, mlw63@georgetown.edu, thank Mario Rizzo and participants at the Colloquium on Market Institutions and Economic Processes, as well as the participants at the Critical Review Foundation seminar in San Antonio, students in the Reading Group in Behavioral Political Economy at New York University, and Jeff Friedman for many helpful comments. The usual disclaimer applies.

Honor is a great check upon mankind; but where a considerable body of men act together, this check is in a great measure removed, since a man is sure to be approved of by his own party for what promotes the common interest, and he soon learns to despise the clamors of adversaries.

— David Hume (1953[1742]), “On the Independence of Parliament”

Leif Lewin’s great contribution in *Self-Interest and Public Interest in Western Politics* (1991) was to examine whether voters, bureaucrats, and politicians are really as selfish as public-choice scholars had supposed. He drew on extant evidence to show that they were not and, for the most part, we agree. Particularly important, for our purposes, is the now long-established fact that most citizens vote “sociotropically,” not selfishly: that is, they aim at their understanding of the good of all, not their material self-interest (Kinder and Kiewiet 1979 and 1981).

At the same time that Lewin was challenging the public-choice hypothesis, other scholars were beginning to examine whether people were so instrumentally rational as economists, including public-choice scholars, had supposed. How rational are people really? Economists, political scientists, psychologists and sociologists have built a convincing case that the answer is: “It depends.” What it depends upon are the details of the task environment, i.e., the situation or context in which people are engaged. Where behavior is rational, it is not because of the mind alone, but because of productive mind-environment interactions.

One contribution of this paper is to offer a framework of citizen belief formation that builds on these scholars’ challenge to rational-choice theory, and on the challenge to public-choice theory posed by the scholars upon whom Lewin drew. Research on the economics and psychology of belief formation and decision making, especially in social contexts, provides a common framework for explaining how both beliefs and motives are formed. This allows us to put together an integrative framework that explains certain stylized realities of mass and elite psychology better than does a framework that focuses on motives or beliefs only.

A second contribution of this paper is to pose a dilemma. We argue that certain political task environments lead to potentially productive disregard for oneself

(“altruism”), but to socially destructive disregard for out-groups (“righteousness”) and for distant consequences (“myopia”). More specifically, by “righteousness,” we mean disregard for the insights or welfare of out-groups, driven by a feeling of moral superiority that has been validated by a group of like-minded others. By “myopia,” we mean a disregard for distant consequences, whether in space or time. As a result of these biases, people place disproportionate attention on winning the next high-stakes election so as to put “the good guys” in power. The task of discovering, enforcing, and adapting institutions that are robust, fair, and mutually productive goes relatively ignored.

The paper proceeds as follows. The first section puts our arguments in the context of Lewin’s empirical challenge to the first generation of public choice theory, and presents a brief history of the relevant political-economy literature. The second section defines what we see as the fundamental political task: forming true beliefs about the likely consequences of alternative people, parties, policies, and policy-making processes. The third explores extant frameworks for explaining how people arrive at political motives (whether selfish or sociotropic) and political beliefs. The fourth section presents a “cybernetic” account of the endogenous formation of motives and beliefs. The fifth section proposes a functionalist framework for reasoning, which boils down to the claim that the moral-social context can cue us to think like intuitive prosecutors, politicians, or theologians rather than intuitive scientists. The sixth section considers how political task environments foster altruism, the seventh how they foster righteousness, and the eighth how they foster myopia.

1. Public Choice and the Comparison of Institutional Alternatives

Public-choice theory is often characterized as the application to politics of the core behavioral assumptions of economics: selfishness and rationality. It is a fair characterization of the vast majority of the public-choice research of the past 50 years.

For many in the first generation (e.g., James Buchanan, Gordon Tullock, Ronald Coase, Harold Demsetz), however, this would be missing the point. The primary point was that scholars should compare and contrast the performance of institutional

alternatives, not blithely assume that the real shortcomings of one institution could be remedied by the idealized capabilities of another.¹ Specifically, they took issue with the notion that market institutions could be automatically corrected by a benevolent despot who took his cues from welfare economists. In a world with real voters, policy makers, and bureaucrats, sometimes the political cure is worse than the market disease. And sometimes not. Institutional choice should be comparative and practical; citizens exclusively face choices between imperfect alternatives.

Models with selfish, instrumentally rational decision makers were considered (by many scholars) to be a good way to make headway on this project. The models were tractable. They provided a common conceptual apparatus (beliefs and desires drive choices; choices lead to an equilibrium outcome; that outcome has certain stability, efficiency, and equity properties) that made it easier to compare institutional performance. Moreover, assuming that motives are the same in market and political contexts (what Buchanan called the “behavioral symmetry” assumption) seemed to avoid stacking the deck in favor of politics.

Buchanan ([1972] 2000) allowed that this was a contingent research strategy; it could be true that our motives and beliefs stem from details of our institutional context, rather than from a universal tendency toward the same degree of self-interestedness, whether in political contexts and market contexts. Even if behavioral symmetry is a good place to start the inquiry, behavioral asymmetry (in the form of institutionally dependent mind-environment interactions) is a good place to end up.²

¹ This point was made initially in Coase 1960 and Buchanan and Tullock 1962, all of whom were at the University of Virginia at the time. This point is made most

² This seems to have been Buchanan’s position in 1972: “What is needed is considerably more research to ascertain the explanatory power of competing behavioral hypotheses. . . . One result of such research will surely be that the relative applicability of the competing hypotheses will vary from one institutional-environmental setting to another. And indeed a central part of the research may be the identification of those institutional characteristics that seem to exert an influence on personal behavior” (Buchanan 1972 [2000], 36). Similarly, some in the behavioral tradition argue that the rational-choice program was the right way to begin the inquiry, even if it would also be a

Many public-choice theorists would defend the assumption of self-interest not because it is true always and everywhere, or even because it is true to the same degree always and everywhere, but because it is effectively true some of the time in some contexts. The classic example – referenced by Buchanan and Tullock (1962, 17) – is of the apostle Paul’s tent-making: “Paul may be acting out of love of God, the provincial church, friends, or self without affecting the operational validity of the theory of markets.” This is the assumption of “non-Tuism” – that people do not internalize the purposes or interests of the strangers with whom they engage. Similarly, in politics, the assumption of self-interest is effectively true insofar as (1) individuals take on group or partisan interests that can be set into conflict with other group or partisan interests, rather than taking on the general interest, and/or (2) individuals arrive at different notions of the general interest. Many of the analyses shake out the same regardless of whether they involve the perceived interest of humanity, one’s fellow citizens only, one’s industry or profession, one’s union, one’s household, or oneself. Showing that voters, bureaucrats, and politicians do not pursue their individual interests is not the same as showing that they pursue the general interest.³

Even if we were to concede that the typical case were, as Lewin proposed, that participants in political processes seek the general interest, it could be a mistake, from an institutional-design perspective, to accommodate only typical cases. Institutional designs that are “robust” or “fault-tolerant” must accommodate (among other impediments) self-serving or group-serving behaviors (Boettke and Leeson 2004; Pennington 2011), even if these are the exception rather than the rule.

In this context, it is helpful to see Lewin’s challenge as part of a broader conversation about the relative appeal of grand institutional alternatives: politics and markets. Space and knowledge constraints permit us to provide only a potted history of

regrettable way to end it. See for example, Bendor et al. 2011 (1-5) in their exposition of a “behavioral theory of elections.”

³ Technically, what Lewin shows is that voters do not vote their pocketbooks, bureaucrats do not seek to maximize their budgets, and politicians do not create political business cycles.

the relevant political economy. Moreover, our potted history comes with a bias. We choose to highlight the tacking back and forth between left-leaning implications (where political solutions to market problems look relatively better) and right-leaning implications (where they look relatively worse). This does some violence to the complex aims, conclusions, and implications of the work we discuss, but—with apologies—we find it illuminating nonetheless.

Once upon a time, some economists developed arguments about how markets could fall short of idealized efficiency criteria. The perfect market is a myth, they argued. The real world was, is, and always will be teeming with externalities, public goods, less-than-perfect competition, natural monopolies, and so forth. The often-implicit assumption was that economic policies made by a benevolent despot could solve these problems. For many, this—the “welfare economics” of the 1930s, ’40s and ’50s—was an attempt to correct what seemed to be overly optimistic appraisals of the likely consequences of laissez-faire capitalism.

Public-choice economists, however, saw welfare economics as an overcorrection. Comparing real markets to ideal governments created an overly optimistic bias in favor of political solutions to market problems. These economists developed arguments about how politics, too, could fall short of ideal efficiency criteria. Public policies emerged from the interactions of voters, lobbyists, politicians, and bureaucrats, all of whom have purposes and prejudices that could lead them to use their influence to make economic policies very different than those recommended by economists. Thus, they argued that economists should try to identify and recommend the “least bad” of less-than-ideal institutional alternatives. For many, the “public choice” of the 1960s, ’70s, and ’80s was an attempt to correct what seemed to be one-sided appraisals of the likely consequences of political solutions to market problems.

In turn, other economists, psychologists, political scientists, and philosophers saw public choice as an overcorrection. Some argued that markets worked even worse than the earlier generations of welfare economists had suspected. Beginning in the 1970s and 1980s, information economics highlighted the information asymmetries that could cause

problems in markets (Akerlof 1970 and 2002; Spence 1973 and 2002; Stiglitz 2002). At the same time, those working in behavioral economics showed how the behavior of market participants departed from ideal conceptions of rationality (Kahneman et al. 1982; Loewenstein and Elster 1992; Gilovich et al. 2002). In turn, Lewin and others argued that politics worked better than public-choice economists alleged, because voters, politicians, and bureaucrats were in fact typically motivated to behave altruistically (or “sociotropically”) rather than selfishly (Sears et al. 1980; Kinder and Kiewiet 1981; Mansbridge 1990; Lewin 1991). Many saw this work as (among other things) a necessary corrective for the biases of public-choice theory.

Beginning in the 1990s, other scholars argued that the problems with markets highlighted by information economics and behavioral economics were more likely to persist and to be destructive when people acted as voters or politicians than when they acted as consumers or producers. Moreover, it was held that while many voters and politicians are not selfish (in the sense of being short-term wealth maximizers), their altruism is mediated by group identities and/or ideologies that can cause a lot of trouble, and maybe more trouble than mere selfishness would (Bennett and Friedman 2008; Brennan and Lomasky 1997; Caplan 2007; Cowen 2005; Friedman 1996, 1998, 2003, 2005, 2006a, and 2007b; Friedman and Kraus 2011, ch. 4; Glaeser 2004; Hardin 2009; Kirchgassner and Pommerehne 1993; Somin 1998, 2000, 2004, and 2006).

The conversation continues. In this telling, at least, the conversation could be understood as a mutually conciliatory one, where each side concedes the last point of the other, and then adds a new piece rather than challenging the old one. Such conversation increases the quality of arguments and the level of sophistication available to students of markets and politics. Today, rather than taking the process of motive- and belief-formation as exogenously given in “markets” or “politics,” the aim of many scholars is to explain how motives and beliefs emerge from the interaction of individual psychology and the social environment. The idea is to make the nature of motive and belief formation endogenous rather than exogenous.

2. The Fundamental Political Task

This section considers alternative frameworks for explaining how people arrive at self-regarding or other-regarding motives, and given those ends, how they arrive at conclusions about which means – what people, parties, policies, and processes – would be better or best to pursue.

Following Michael X. Delli Carpini and Scott Keeter (1996), we believe that the key task faced by citizens is to form beliefs about the likely consequences of alternative politicians or parties taking office, and thus about the effects of alternative policies championed by the different parties. Delli Carpini and Keeter (*ibid.*, 14) note that “effective citizens” should have a general familiarity not only with “the rules of the game (the institutions and processes of elections and governance),” but with “people and parties (the promises, performances, and attributes of candidates, public officials, and political parties)” and with “the substance of politics (the major domestic and international issues of the day, current social and economic conditions, key policy initiatives, and so forth).” As a shorthand, we will refer to forming true beliefs about people, parties, policies, and policy-making processes as the “fundamental political task.”⁴

Citizens’ beliefs affect whether and how they vote and whether and how they seek to mobilize others. However, a citizen’s task goes beyond picking the better of two candidates in an election. Party platforms are constrained by the beliefs of voters. If all citizens appear to share the belief that ethnic outsiders (or the rich, or the poor, or the media, etc.) are to blame for some recent trouble, then candidates and parties may converge to an identical platform on this point. In such circumstances, voters cannot help but be “right” (in the sense of picking the party with the stance closest to their own on this issue) even as they are all very “wrong” relative to the truth. Therefore, Bryan

⁴ The relative importance of each of these depends upon the time horizon we consider. We would agree with those that propose the most fundamental task is that of forming true beliefs about alternative policies and policy-making processes.

Caplan (2008, 458-59) maintains that scholarship that confines itself to asking whether voters can correctly match platforms to parties or candidates suffers from a “binary fallacy.”⁵

Of course, the fact that voters’ beliefs constrain platforms may be true even if platforms do not converge, or if there are more than two parties. A more apt name, then, would be the “exogenous alternatives fallacy.” This puts more weight on the importance of knowing about the effects of policies than most public-opinion scholars would. In our reading, a voter does not merely need to orient himself within a fixed world of vote choices well enough to vote “correctly” between them, but rather needs to know whether policies are desirable relative to technically (if not politically) feasible alternatives. We assume that public opinion can constrain what candidates see as politically feasible without assuming that it always does so, or that it always does so through the same mechanism.⁶

The fundamental political task is difficult, and part of its difficulty is that it can seem easy. Complex social problems involve many parts, indirect and delayed effects, and interactions that are more obscure than simple additivity (Dörner 1997; Jervis 1998). As a result of this complexity, outcomes do not necessarily follow from intentions. As well, people may find confirming evidence for simple models that capture some aspects of complex systems; the behavior of the part may thus be confused for the behavior the whole. In this light, we might say that there are four kinds of technical-causal systems: those that are both objectively and subjectively “simple”; those that are objectively simple but seem complex; those that both are and seem complex; and those that are

⁵ Caplan’s response is to Arthur Lupia (2008, 445-49), but the idea that the task is simply to pick the better of two candidates is commonplace. Even Delli Carpini and Keeter (1989, 50), whose criteria for an “effective citizen” we borrowed above, note in passing: “In the typical two-person race [the randomizing] decision rule does leave a 50 percent chance that the voter’s interests will be served (at least as well as allowed by the choices available).”

⁶ Politicians could be constrained by public opinion as they apprehend it in casual encounters, polls, the results of previous elections, or the popular press.

complex but do not seem so. The last of these are systems that are “deceptively simple.”⁷ When citizens face deceptively simple systems, their intuitions lead them astray.

Many others have written about how citizens address (or fail to address) the fundamental political task. In the next section we try to capture the evidence that seems (to us) to have been most fertile as a set of “stylized facts.” We use the stylized facts to capture the inadequacies as well as to appreciate the insights of existing frameworks (and economic frameworks in particular). We use these to lay the groundwork for our proposed framework.

3. Economic Accounts of Belief and Motive Formation

Economic theories of voting tend to discuss three types of behavior: (1) whether citizens turn out to vote, (2) how they choose to vote, and (3) how much additional effort they choose to put into choosing how to vote. Both motives and beliefs come into play for each of the three. An adequate framework for motive and belief formation should therefore be consistent with stylized facts that pertain to at least these behaviors. We believe the accounts we survey fail this test. Nonetheless, they fail in interesting ways. We use these accounts to pull together a framework that is consistent with not only these special cases (and the stylized facts that pertain to them), but more generally with how people form motives and beliefs outside the political domain.

It seems important to distinguish between three levels of economic explanation that are often conflated: rational-choice explanations, intention-based explanations, and opportunity-based explanations. Rational choice, as Jon Elster (2007, 191) summarizes it, involves three optimizing operations: (1) choosing the best action given one’s opportunities, desires, and beliefs, (2) forming the best grounded beliefs based on given evidence, and (3) collecting the right amount of evidence given opportunities, desires, and prior beliefs. The second and third operations drive belief formation.

⁷ Psychologists refer to the “illusion of explanatory depth” (Rozenblit and Keil 2002; Tasic 2009).

Caplan (2007) refers to a failure of the second operation as “irrationality,” in part to contrast it with “ignorance,” which is commonly seen as a failure of the third operation.⁸ While some see this as a problematic distinction (Bennett and Friedman 2008), especially insofar as ignorance of how to form the best grounded beliefs for given evidence may cause “irrationality,” we find it nonetheless useful.⁹ Using Herbert Simon’s terminology (1985, 294), we can define rationality as substantive or procedural. Whereas substantive rationality is optimally adapted to the evidence regardless of any constraints internal to the chooser, procedural rationality takes “into account the limitations of knowledge and computing power of the choosing organism,” which can prevent it from making objectively optimal choices (ibid.). Procedural rationality is a human standard to replace the superhuman standard of substantive rationality.¹⁰

Intention-based explanations (rational choice or otherwise) have three elements: a set of opportunities available to an agent (i.e., behaviors consistent with physical, economic, legal and other constraints), a set of beliefs about how opportunities map to outcomes, and a set of motives (i.e., desires or preferences) as to what makes an outcome desirable. An explanation in terms of an agent’s opportunities, beliefs, and motives is intention-based, even if the three optimality conditions do not hold. Behavior may still be intelligible in these terms even if it is not substantively or procedurally rational. Intention-based accounts include rational belief-formation as a special case, but they are also consistent with inadvertently under- and over-reacting to unintended consequences (Jones 2001) as well as with downright (procedural) irrationality. Motivated and counter-motivated thinking—where hopes and fears improperly influence beliefs—are examples

⁸ Economists, following Downs (1962), would argue that the “right” amount of evidence is less than all the evidence when gathering it is costly, but in common parlance, to be ignorant is to have little or less evidence, knowledge, or information than the speaker judges to be appropriate.

⁹ It is especially regrettable that both of the first two operations can be confused with the whole. The notion of rationality which means “forming the best grounded beliefs for given evidence” (the second optimizing operation) is distinct from that which consists of weighing costs and benefits (the first optimizing operation), as well as from the most general notion of rationality that consists of all three optimizing operations.

¹⁰ As we use it, procedural rationality—while a lower standard—is still a standard that behavior can fail to meet. It does not mean that whatever a chooser does is by definition the best they could have done.

of procedurally irrational behavior that we will consider below.¹¹

Opportunity-based explanations are the most general. They need not consider beliefs, motives, or even conscious thought in the explanation of behavior (Becker 1962). Economists typically focus on monetary budget constraints as defining the opportunity set, but other key constraints are those on time and attention (Hardin 2009). In double auctions, so-called “zero intelligence traders”—traders who make offers randomly but are constrained by their budget and market rules—realize most of the available gains (Gode and Sunder 1997). Regardless of whether we arrive at our behavior through surrender to cravings and compulsions, adherence to local custom, or optimizing choice, we cannot exceed the bounds of the opportunity set. I can only give away all that I own, not more. I can only spend all my time in service to others, not more. And, as Nathan Hale regretfully noted, “I have but one life to give for my country.”

As we consider economic accounts of motive and belief formation, it is good scholarly hygiene to check whether an explanation that fails as a rational-choice explanation might work as an intention-based explanation, or if one that fails there might work as an opportunity-based explanation.

Stylized Fact 1. Most citizens are ignorant of even basic political facts. The early public-opinion literature, especially Berelson et al. 1954 and Campbell et al. 1960, established the first precise profiles of the nature and extent of ignorance of basic political facts.¹² Widespread and deep public ignorance was inconvenient for the prevailing normative emphasis on the importance of being well informed. This normative perspective was further challenged by Anthony Downs’s explanation of how ignorance—which he defined as a failure to invest in gathering political knowledge—was the

¹¹ It would be “rational” for desires (i.e. hopes and fears) to influence beliefs indirectly insofar as they effect the amount of evidence that is gathered, which in turn influences beliefs. The improper influence would be more direct: having a bias toward believing what one hopes, or what one fears. We discuss mechanisms in which this is consistent with the illusion of personal objectivity in the section on righteousness.

¹² While others intuited the nature and extent of public ignorance through careful observation (Lippmann [1922] 1997; Schumpeter [1942] 1950), the fact base they had to systematically study the phenomena was not as “thick” as it would become.

individually rational (albeit socially problematic) response in the face of (a) the costs of gathering knowledge, and (b) the near-zero probability that one person's more well-informed voting would affect the outcome in a large electorate. Downs proposed that the same logic applied to the question of whether one should bother voting at all. Given the costs of voting and the near-zero probability that one person's vote would affect the outcome, the individually rational response was to stay home.

Stylized Fact 2. Many citizens participate in mass elections. The problem for theories of “rational abstention” is that, while many citizens do stay home, many others choose to vote. Depending on the details of the model, the theory seems to predict zero or near-zero turnout. The disconnect between that prediction and the reality of relatively high rates of turnout has been dubbed the “paradox that ate rational choice” (Fiorina 1990, 334).

Some theorists (e.g., Riker and Ordeshook 1968) try to escape the paradox by suggesting that voters may gain procedural utility from the act of voting itself, regardless of instrumental impact on the outcome. This type of explanation can be abused (“Why did the man do X? Because he gets procedural utility from doing X, whatever X might be.”), but that alone does not make it wrong. Responsible use of this kind of explanation requires a commitment to explaining which mechanisms imbue actions with procedural utility.

The theory of expressive voting (Brennan and Lomasky 1993) is one attempt to go down this path. Brennan and Lomasky propose that the satisfaction of expressing an opinion—especially in solidarity with others—drives participation. People who will pay good money to sit and cheer for hours at sporting events can be expected to wait for an hour (and pay nothing) to cheer for their preferred political team with a vote. Like Downs, Brennan and Lomasky argue that citizens anticipate that the probability of their vote deciding the election is near zero, and that this plays into not only that but how they participate.

Stylized Fact 3. Some citizens are well informed about basic political facts. An additional problem for the theory of “rational ignorance” is that, while many voters are

ignorant of basic political facts, a politically passionate minority of voters is highly informed (Campbell et al. 1960). Brennan and Lomasky's theory can explain this, because the same motives that enable people to vote could motivate them to become highly informed. Like sports fans, they may meticulously gather the latest "stats" about their preferred team, even though their actions make no difference to whether the team wins. At the same time, however, most sports fans do not go that far. So this explanation does not predict that everyone will be as knowledgeable as the most ardent fans, and, indeed, most members of the public are quite ignorant politically, even though they do, like sports fans, vote. This could be seen as the exception that proves the rule. Most voters who manage to be highly informed require the aid of "belief systems" or ideologies that process otherwise overwhelming complexity into a meaningful system or narrative, filtering out the irrelevant and framing the relevant (Converse 1964).

Another problem was the emerging evidence—carefully considered by Lewin—that those who did vote were voting with altruistic or sociotropic rather than selfish motives.

Stylized Fact 4. Many voters exhibit sociotropic rather than selfish motives. A "public-spirited" explanation of voting is stronger than the Downsian explanation insofar as it predicts that many people will participate, and insofar as it predicts how they will vote when they do: They will vote in line with their perceptions of which politicians, parties, and policies will serve the public good. But there is still the problem that most citizens— even among those who choose to vote—are not well informed about the most basic political facts. This remains a problem for the public-spirited view.¹³

¹³ As Friedman's work (1996, 1997, 1998, 2003, 2005, 2006a,b, 2007a,b; Bennett and Friedman 2008) has always emphasized, the theory of rational ignorance does not get explanatory traction here either insofar as people are "inadvertently" or "radically" ignorant. Evans and Friedman 2011 discuss "radical ignorance," or ignorance of "unknown unknowns." Evans and Friedman claim that such ignorance is not as susceptible to incentives to overcome ignorance, because the agent does not know that what he does not know might prove valuable to learn. In a sociotropic context, this means that the agent does not know that there is missing information that might change his assessment of which parties, politicians, policies, etc. to favor. As we discuss below, we consider this a good place to start but a bad place to end the explanation. In some

Another problem is to explain sociotropic motives, and how they wax and wane, rather than to stipulate them (as economists stipulated selfish motives). One such explanation is that people readily form sociotropic or altruistic motives when the cost of doing so is low (Caplan 2007). Kirchgassner and Pommerehne (1993) raised this hypothesis as “a challenge to public choice.” It has been satirized as the idea of “selfish altruism”—a play on “rational irrationality”—and challenged on the grounds that people cannot purchase a “positive self-image” or a “warm glow” from mouthing sociotropic sentiments, even in the voting booth (Elster and Landemore 2008). Instead, these feelings must emerge as a byproduct of sincere altruism.

We could not agree more. It is entirely plausible to us that most people default to sincere altruistic motives, that acting on these motives creates a “warm glow” that positively reinforces the tendency to do so, that self-interest is activated only when the altruistic motive imposes enough of a material cost, and that this may only be discovered *ex post*. The effect, however—that people act more altruistically when the cost of doing so is lower—is the same.

Stylized Fact 5. Many citizens are confident in their political preferences, despite having low levels of political knowledge. There are more partisan voters than there are people who score high on tests of political knowledge. Moreover, they vehemently disagree with one another, and they cannot all be right (Cowen 2005). Some of us, and perhaps all of us, are too confident. This phenomenon, familiar to psychologists, was called out as a “novel fact” in the debate about rational-choice theories of voter behavior by Caplan (2001, 2003, and 2007).¹⁴ This fact contradicts the predictions of the theories of rational ignorance and expressive voting, since if those theories were correct, the ignorant majority of partisan voters would know that they are underinformed.

circumstances, feedback from the technical-causal environment is available to make “unknown unknowns” into “known unknowns” or “known knows,” and in others it is not. The persistence of inadvertent ignorance is different across different contexts, and this is something to be explained.

¹⁴ The claim is not that Caplan was the first scholar to highlight the over-confidence of voters, but rather that he turned the attention of scholars to a phenomenon that seemed to contradict the predictions of the theories of rational ignorance and expressive voting.

One way to explain the general public's overestimation of its knowledge is to view it as a version of the Dunning-Kruger effect (Dunning and Kruger 1999). Those without knowledge often lack the means to assess their level of knowledge, and thus to the means to set the appropriate level of confidence. It could also be understood as a version of "motivated reasoning." As Elster and Landemore (2008) note, there are typically two kinds of motivated reasoning: wishful thinking (believing what I prefer to be true) and prideful thinking (preferring my beliefs just because they are mine). Both of these are typically seen as reality-constrained. That is, mentally healthy adults cannot strictly believe whatever they prefer to be true, nor can they defend whatever beliefs they may have held. But within the subset of beliefs that seem sufficiently plausible, we may be biased towards the beliefs that seem desirable.¹⁵ And within the subset of beliefs that seem sufficiently defensible, we may be biased toward our prior beliefs. Notably, wishful thinking can contribute to prideful thinking. One common meta-belief that many would consider desirable is that they have every reason to be confident in their beliefs.

The theory of rational irrationality (Caplan 2007) goes a step further than psychological accounts of motivated thinking, and claims that this effect persists more when the material cost of error is low. In contrast to the theory of rational ignorance, the question is not whether evidence is sought after (operation three above), but whether immediately available evidence is used responsibly (operation two above). In contrast to the theory of expressive voting, the theory of rational irrationality is not just that voters—like sports fans—derive psychological benefits from backing what feels good, but also from believing in what they back (Caplan 2007, 138-39). Like the theories of rational ignorance and expressive voting, the starting place of the theory of rational irrationality is that voters believe the probability of their vote deciding the election is near zero. As Caplan (2007) would have it, the rationally ignorant and rationally expressive must make conscious trade-offs, but the rationally irrational know not what they do. Rationally ignorant citizens with little knowledge would offer political opinions with the caveat that

¹⁵ We discuss below, in the section on righteousness, how this bias is consistent with the illusion of personal objectivity. There are several ways in which motivated belief and counter-motivated belief are thought to fit with the phenomenology of belief that we believe things simply because they are true.

“I don’t know, because it doesn’t make sense for me know,” and rationally expressive voters would say, “I think policy X is the right answer, but it feels better to root for policy Y.” But rationally irrational voters would believe what they want, and then be confident that they are right.

On its face, this is no more psychologically plausible than the notion that people could derive utility from rooting for policies they know to be bad policies. In the most literal (and least defensible) version, then, rational irrationality would claim that people knowingly discard the truth when the cost of doing so is sufficiently low (Bennett and Friedman 2008, pt. I). A more defensible version, suggested by both Caplan (2007) and by Elster and Landemore (2008), is that the default is to be miserly with one’s cognitive and meta-cognitive effort, while the perceived possibility of making a high-cost error triggers more effort.¹⁶ In some cases, however, merely perceiving the possibility of error *ex ante* is not enough; sometimes one must actually commit an error and feel the cost directly (perhaps several times) before greater meta-cognitive effort will be forthcoming. This type of explanation—in which reaction to negative reinforcement is more important than anticipation of costs—figures into the cybernetic account we provide below. We think it is much more plausible than the notion that people knowingly choose to consume false beliefs, so long as the price is low. But as with the theory of selfish altruism, the effect is the same: people are more (procedurally) irrational when the cost of being so is lower.

In an absolute sense this helps the theory of rational irrationality, but relative to its peers, it hurts it; the same reasoning could apply to the theory of rational ignorance and that of expressive voting, and doing so allows these theories to explain the “novel fact”

¹⁶ This is more psychologically realistic—for reasons we will discuss more below— but may be less appealing to the kind of economist who would affirm departures from the rationality assumption only with the express permission of the rationality assumption. In this context, the contribution of Caplan’s more stylized, less realistic version is to open the door for these economists to learn from social psychologists. Lewin (1991, 1) notes that “to establish cross-fertilization is the aim of [his] book,” and we read Caplan (2007) through the same lens.

emphasized by the theory of rational irrationality.¹⁷ For example, suppose the default is to be ignorant and even to be radically ignorant, unaware of our own ignorance. We may only realize we were radically ignorant after the fact, when our radical ignorance proves costly. But in the absence of that feedback, instead of saying, “I don’t know because it doesn’t make sense for me to know,” we might well be proud and confident in our (blissfully unaware) ignorance. Similarly, suppose the default is to express our moral intuition (i.e., what feels right), unless our moral intuition turns out to be costly after the fact, at which point our reasoning is activated. When the price is low, we simply express what feels right (and believe it to be truly right).

Though we have somewhat rehabilitated the psychological plausibility of the theories of rational ignorance, expressive voting, and rational irrationality by recasting them as intention-based or opportunity-based accounts, we still think they are inadequate.

Stylized Fact 6: Many highly motivated and knowledgeable citizens nonetheless disagree with one another, often dogmatically. A different challenge to these theories is that they reduce only to a motivational problem: If only people were motivated to be rational, they would be rational. But by this logic, many powerful elites should be highly motivated to obtain true, not feel-good, beliefs, yet they still disagree with one another, indicating that at least some of them must be wrong (Friedman 2006a, xxxiii). Even people with real political power often disagree with each other, and they give every indication of believing what they say (Bennett and Friedman 2008, pt. II).

What all the economic hypotheses have in common is the idea that people will be more responsible—they will gather more evidence and use the evidence they have more carefully—when they are pivotal decision makers and the stakes are high. In these accounts, higher stakes have a way of getting our attention and motivating us to “be our best,” cognitively speaking. But this depends on the extent of technical-causal complexity

¹⁷ We do not think that this makes the theories hard to distinguish from one another. Each theory concerns a different mental operation—acquiring knowledge is distinct from using given knowledge, which is distinct from expressing moral intuitions—even though they often operate in concert.

in the task of interest to us, as well as the feedback from the task environment. When we face complex, unfamiliar, and hard-to-observe problems, even the understanding of the most capable and highly motivated among us is not likely to be very good. Where decisive choosers would falter, we cannot fault the voter's lack of decisiveness.¹⁸

Moreover, those with the most knowledge may also be the most dogmatic and resistant to belief changes (Friedman 2006a). In this light, one explanation for stylized fact number 6 is that the social and political world is complex enough to accommodate the divergent interpretations of observers who are highly motivated to get things right.¹⁹

Stylized Fact 7. Many citizens believe in things that they do not find pleasant. Another challenge to the theory of rational irrationality is that people are burdened by “believing what they fear” as well as by “believing what they hope” (Elster 2007, 39).²⁰ While it may be possible to consider some irrational fears as a product of prideful or wishful thinking, the political world is full of citizens believing things they wish weren't so. Many Americans, for example, seem to have exaggerated beliefs about threats from crime, drugs, minorities, immigrants, and terrorists (Glassner 1999; Glaeser 2005; Mueller 2006).

Stylized Fact 8. Many citizens say their individual vote matters. The traditional starting place of all the rational-choice explanations (rational ignorance, expressive voting, and rational irrationality) is that each voter subjectively knows that he objectively faces a near-zero probability of deciding an election. Yet many voters turn out to vote, and, when asked, profess the belief that their vote could decide the election. In a recent

¹⁸ This suggests an alternative (and complementary) explanatory benchmark to the non-decisive chooser: the behavior of a chooser who subjectively believes he faces and objectively does face the reality of being the decisive chooser (Cowen 2005). This point is considered more in the discussion of Stylized Fact 8 below.

¹⁹ High stakes need not help, and may hurt. When the time frame is short, at some point increasing the stakes has the effect of increasing stress (Ariely 2010).

²⁰ As we discuss below, unlike the theories of rational ignorance, expressive voting, and rational irrationality, the theory of prideful thinking or self-deception is not predicated on a voter knowing his vote will not matter (Cowen 2005). However, that theory is challenged by instances of counter-motivated reasoning.

poll, for example, 83 percent of likely voters say that they believe that one person's vote really matters (Rasmussen 2011). This brute fact (cf. Zukin et al. 2006, 223)²¹ contradicts the premise of the rational-choice theories, or at least restricts their domain to a small minority of voters. All 83 percent cannot be right, and objectively none of them are.

It is helpful to set this in the context of four ideal types:

Case 1: Choice will be pivotal in fact, and the chooser (correctly) believes it will be.

Case 2: Choice will be pivotal in fact, but the chooser (incorrectly) does not believe it will be.

Case 3: Choice will not be pivotal in fact, but the chooser (incorrectly) believes it will be.

Case 4: Choice will not be pivotal in fact, and the chooser (correctly) believes it will not be.

Case 4 is the traditional starting point for economic explanations of ignorance (of the Downsian variety), of expressive behavior (of the Brennan and Lomasky variety) and irrationality (of the Caplan variety), but in fact the evidence suggests that most voters are examples of Case 3. Whether from acceptance of "every vote counts" propaganda, ignorance of the odds, or magical thinking, they believe that their individual votes are likely to be pivotal, even in a large electorate. Case 2, in turn, is the situation posited by many of those who work to mobilize voters: "Don't be an apathetic fool who stays home when otherwise your vote would have tipped the scale." Case 1 is the situation typically faced by a consumer, by a manager inside a firm, or by an executive with discretion to

²¹ Zukin et al. 2006 provides survey results of nonvoters, of whom only 18 percent said that the reason they did not vote is that "my one vote isn't going to make much of a difference."

make public policy.²²

Stylized Fact 8 also challenges the idea that people have internally consistent beliefs. As Cowen (2005) notes, people

will claim that their vote might matter, yet without going to great lengths to vote in each and every election, or without putting great time into making the right decision. . . . Or, when asked why they vote, many voters respond with the query “What if everyone didn’t vote?” Yet this reasoning is rarely applied consistently. If a person suddenly breaks his leg and cannot easily vote, he does not fear that democracy suddenly will vanish.

For theories that hold people have internally consistent beliefs, it is a puzzle that voters do not act on the professed belief of their decisiveness consistently. Part of what we want to explain is this inconsistency.

In our judgment, when it comes to motive formation, assuming that citizens are selfish seems consistent with low levels of knowledge among many voters (as well as the abstention of non-voters) but inconsistent with the fact that many citizens vote, vote sociotropically, and have basic political knowledge. Assuming that citizens are altruistic has the opposite strengths and weaknesses; those who spend the time and effort to inform themselves and vote are explained more readily than those who do not. (Not incidentally, we think Lewin does a better job at showing where the assumption of selfishness fails—especially for voters—than he does in explaining the domain where the proposition that citizens are altruistic holds, or in deriving further implications of that proposition and then testing those.) While it is psychologically implausible, assuming that citizens are selfishly altruistic would explain relatively high turnout and relatively low levels of political knowledge (Hardin 2008). When it comes to belief formation, assuming that citizens are well informed and rational flies in the face of low levels of political knowledge and excessive levels of confidence (whether among the poorly informed many or the well informed few). Assuming that citizens are rationally ignorant is psychologically implausible for many. That assumption would explain the low levels of

²² Other things equal, legislators and bureaucrats would be more likely to be pivotal than the typically voter, but less likely than the typical executive.

knowledge, but not the excessive levels of confidence, especially among those who are highly motivated (meaning, in this context, powerful). Assuming that citizens are rationally expressive is even less psychologically plausible than that they are rationally ignorant, though it would do more to explain why some citizens participate, acquire lots of political knowledge, and root for their team. Assuming that citizens are rationally irrational is even less psychologically plausible than assuming they are rationally expressive, but it would do more to explain why non-decisive citizens are overconfident. It would not, however, explain the overconfidence of the decisive, or the prevalence of counter-motivated reasoning (e.g., fearful thinking).

On the motive side, the simplest versions of the selfishness and altruism stories are inadequate, as is the selfish altruism story. On the belief side, the simplest versions of the rational and well-informed voter theories are inadequate, as are the accounts of rational ignorance, rational expressiveness, and rational irrationality. However, intention and opportunity-based accounts of selfish altruism, rational ignorance, rational expressiveness, and rational irrationality do not suffer from the same limits as those that assume the mind is governed by some higher-order selfishness or higher-order rationality. We draw on these intention and opportunity-based accounts of motive and belief formation in our discussion of the cybernetic framework below.

4. A Cybernetic Framework of Mind-Environment Interactions

Our account begins with the bottleneck of attention. All human beings lack the ability to attend to most of what goes on their immediate environment, to say nothing of the broader world. They also lack the desire to attend to as much as they could. People pay attention to matters that are surprising, or that seem to be materially or socially relevant.

When we pay attention, we engage first with our intuitions, which psychologists call “System 1,” and only engage our reasoning capacities, “System 2,” when our

intuitions prove inadequate relative to expectations.²³ If and when our reasoning capacities are engaged, the quality of reasoning continues to depend upon the task environment. Although a complex environment limits our ability to predict the future accurately (Jervis 1998; Tetlock 2005), we can sometimes steer a reasonable path by reacting to feedback. Rather than anticipating what would be “optimal” (or even “adequate”) ex ante, we are prompted to realize what was inadequate ex post. We call this a cybernetic alternative to the rational-choice accounts because, rather than focusing on people’s anticipation of costs and benefits, we focus on people’s reaction to feedback from the environment.

Metaphorically, we can think of the mind as a student with three instructors: (1) “cause,” (2) “cost,” and (3) “other people.” Each instructor offers a distinct kind of feedback. The first kind is “technical-causal,” by which we mean the observation of apparent cause-and-effect linkages (e.g., the consequences of using a hammer to strike a nail). The second kind is “material-economic,” by which we mean the perception of opportunity costs in the form of commensurate bundles of benefits and burdens that can be weighed against one another (e.g., a proposed exchange or investment opportunity). The third kind is “moral-social,” by which we mean regard for the behavior of our fellow human beings, including their regard for our behavior (e.g., the perception of anger or disgust on another person’s face). Our hypothesis is that the first two instructors—“cause” and “cost”—have a relatively limited ability to impose themselves upon the mind in political task environments, so feedback from “other people” predominates in the formation of political motives and beliefs. Since other people also teach moral lessons, create imagined communities, and define sacred identities, a cybernetic view can also add an account of motive formation.

Feedback can be present or absent, and if present, it can be negative or positive. Negative feedback dampens or reverses changes to the system. If feedback is negative on

²³ There are many variants on the notion of the mind as a dual-processing system. The first system is automatic, fast, and (in evolutionary terms) old; the second controlled, slow, and young (Kahneman 2011). Many of the perspectives on dual-processing systems are collected and contrasted in Chaiken and Trope 1999.

net, the system will stay within a certain range. The classic example is that of a thermostat. A simple thermostat turns on the air conditioning once the temperature exceeds an upper threshold, and then turns it off once the temperature drops below a lower threshold. This design keeps the temperature between the two thresholds. The ideal temperature (e.g. 70 degrees Fahrenheit) is typically somewhere between the two thresholds (e.g. 65 and 75). Relative to an ideal, the system seems to under-react as the temperature climbs above 70, and then to over-react once it hits 75, pushing the temperature past 70 to 65.

Positive feedback amplifies a change, so that a small change leads to a larger change. When feedback is positive on net, it tends to push a system away from its starting point. An example is a stampede, where the panic of one animal can lead its neighbors to panic, which leads their neighbors to panic, and so forth, until the whole herd is on the run.

Negative and positive feedback are related to, but distinct from, the notion of positive and negative reinforcement. Thorndike's Law of Effect, probably "the most important principle in learning theory" (Bower and Hilgard [1966] 1981, 481), is that organisms tend to try alternatives that generate positive reinforcement and tend not to retry those associated with negative reinforcement. This is also known as operant conditioning.²⁴ An aspiration threshold sets what counts as "positive" or "negative" reinforcement (Simon 1955; Bendor et al. 2011, 10).

It is easy to confuse feedback and reinforcement, and not without reason. One important difference is that negative reinforcement can be understood as positive feedback (e.g., an alternative that fails relative to aspirations can lead to a destabilizing search for better alternatives), and positive reinforcement can be understood as negative feedback (e.g., an alternative that exceeds aspirations can lead to a stabilizing settling in

²⁴ For example, Friedman 2006b (479-81), makes operant conditioning the basis of an explanation for why consumer choice can correct consumer purchasing errors: Negative reinforcement from an unsatisfactory purchase impels the consumer to use the power of exit to try a different product.

to routine).²⁵ Another important difference is that feedback from the task environment must be perceived and processed before positive and negative conditioning can happen.²⁶

But feedback is often absent or omitted. Wiener (1948, 95) begins his classic chapter on feedback by considering patients at a neurological clinic with ataxia. As a result, “though their muscles are strong and healthy, they are unable to organize their actions.” To walk, one patient must stare at his feet and kick out his legs; he cannot stand—much less walk—when blindfolded. The cause is that “the part of the spinal cord which ordinarily receives sensations has been damaged or destroyed . . . the receptors in [his legs] send no messages which his central nervous system can pick up and transmit, and for information concerning his posture he is obliged to trust his eyes and the balancing organs of his inner ear” (ibid., 95-96). Analogously, the lack of technical-causal and material-economic feedback can cause the formation and persistence of inaccurate beliefs. Without feedback, we do not register the errors that would provoke us to discard or refine our beliefs.²⁷

A person may engage a particular task environment, political or otherwise, with different degrees of depth. At one extreme, there is inattention to a task. At the other extreme, there is the regular use of reasoning capabilities to gather and integrate new information and to revise previous beliefs and behavior to better solve the problems presented by the task. A cybernetic theory proposes that feedback from the task environment is required to pass an agent from one spot along the continuum to the next (see Figure 1). The negative reinforcement of falling short of aspirations activates attention, and if intuition also falls short, this activates reasoning. The positive

²⁵ That is how we will use the terms here. On the other hand, negative reinforcement can be understood as negative feedback (e.g., trying a new alternative that fails relative to aspirations can lead to a return to a previous routine), and positive reinforcement can be understood as positive feedback (e.g., trying a new alternative that exceeds aspirations can lead to further search).

²⁶ There are occasions when the absence of feedback is itself meaningful (e.g., those for which one might say “no news is good news”), but these are exceptional cases.

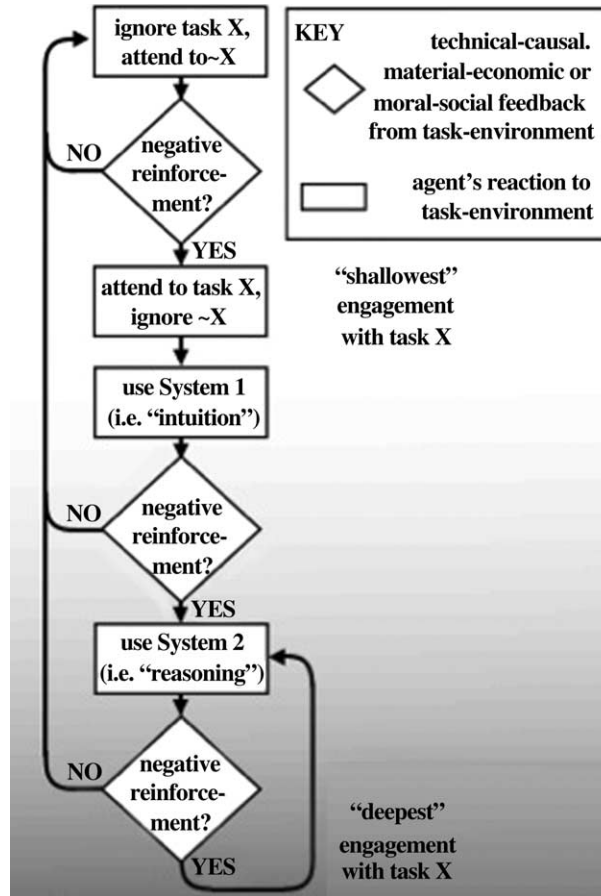
²⁷ In some cases, however, the lack of feedback can be a very good thing. We argue that the absence of material-economic feedback is crucial to unleashing the formation of altruistic motives.

reinforcement of exceeding aspirations de-activates reasoning, intuition, and attention as they relate to task X. Moreover, today's reasoning can become tomorrow's intuition, such that processing the same problem again requires only routine intuition.

This is a simple framework, but even so it is several degrees more psychologically realistic than the rational-actor model. Two points are of particular importance: (1) attention is scarce, and (2) intuition happens before reasoning, if reasoning happens at all. We add these to our list of stylized facts.

Stylized Fact 9: Attention is scarce. People live lives with many purposes and pastimes, which for most of us, most of the time, have little to do with politics. “Most experiences that make life joyful, poignant, boring, or worrisome are not part of the news: the grounds for personal concern, frustration, encouragement and hope; the conditions that matter at work, at home, and with friends; the events people touch, as distinct from those that are ‘reported’; the experience of financial distress or of opulence; children in trouble; lovers; alienating or gratifying jobs” (Edelman 1988, 35). As a result of the busyness of our own lives, we may never form the intention to address the fundamental political task.

Figure 1. The Interaction of the Mind and Task Environments Drives Attention, Intuition, and Reasoning.



Even those who do form the intention, registering it on the mental to-do list of “prospective memory,” may find their attention captured by other tasks. Mundane examples of this are provided in Human Error (Reason 1990, 70-72), among them:

- “On starting a letter to a friend, I headed the paper with my previous home address instead of my new one.”
- “I intended to stop on the way to work to buy some shoes, but ‘woke up’ to find that I had driven right past.”
- “I picked up my coat to go out when the phone rang. I answered it and then went out the front door without my coat.”
- “I walked to my bookcase to find the dictionary. In the process of taking it off

the shelf, other books fell onto the floor. I put them back and returned to my desk without the dictionary.”

James Reason calls these “double capture errors” because first, “the greater part of the limited attentional resource is claimed either by some internal preoccupation or by some external distractor at a time when a higher-order intervention (bringing the workspace into the control loop momentarily) is needed to set the action along the currently intended pathway”; second, “the control of action is usurped by the strongest schema leading onwards from that particular point in the sequence” (ibid., 68). Distraction drives deference to defaults.

Notably, these are not complex systems in technical-causal terms. The only problem is the bottleneck of attention; once attention is forced back to the problem, it is obvious what has gone wrong. In all of these cases, the actor is “inadvertently ignorant” or “radically ignorant” of his error, but feedback from his task environment is able to trigger his attention. With sufficient feedback, many inadvertent errors are noticed and intuitively corrected; without it, they persist. In a marginally more complex task environment, the intuitive corrections would not work as intended.

In contrast to the idea of rational ignorance, the lack of investment in more political information is here driven by the fact that many citizens are simply otherwise engaged and therefore unable to attend to the question of how much political information they should have, much less actually deciding to gather that information (or not). “A wealth of information creates a poverty of attention and a need to allocate that attention efficiently among the overabundance of information sources that might consume it” (Simon 1971, 40-41). Political happenings are geographically and conceptually remote (Lippmann [1922] 1997; Schumpeter [1942] 1950). Information may be freely available but nonetheless ignored because attention is captured by more immediate concerns.

Stylized Fact 10: Intuition happens before reasoning, if reasoning happens at all. The intuitive system is ancient, automatic, and very fast; the reasoning system is younger, consciously controlled, and relatively slow (Zajonc 1980; Haidt 2007). The exact boundaries between System 1 and System 2 in dual-processing accounts of the mind are

not always the same, and not always clear (Gilbert 1999). We consider “intuition” to include gut feelings, heuristics, and rules-of-thumb that are cued as appropriate by features of the task environment. In the moment, one may be aware that one is applying a rule or one may not.

Our intuitions come from “experience” in previous task environments. One can speak of three time scales of past task environments (Jones 2001). The first is the species level, at which our capabilities for attention, intuition, and reasoning, among others, evolved (Flack and de Waal 2000; Cosmides and Tooby 2007; Pinker [1997] 2009). The second is the social level, through which the knowledge of our predecessors and peers is made available to us (Boyd and Richerson 2005; Hardin 2009). Most knowledge is of this type: second hand. As Avishai Margalit has summarized it, we are not so much “caught in a web of beliefs” as “caught in a network of witnesses” (quoted in Schulz 2010, 142). Critically, our ability to absorb local customs (i.e., “biased transmission” or “imitating the majority”) is itself a part of species-level learning (Boyd and Richerson 2005; Simon 1996; Jones 2001). The third level is our own individual experience. This is a combination of formal training and the ad-hoc accumulation of habits and rules of thumb, facts, and theories about the world that come through cultural exposure and through trial and error (Gigerenzer 1999; Smith 2008; Hardin 2009). Typically, if one’s initial intuition fails relative to implicit aspirations, positive feedback (or negative reinforcement) amplifies one’s level of engagement, switching on reasoning.

Haidt (2007, 999) notes that there are three ways to “override our immediate intuitive responses.” The first is to imaginatively reframe the situation, “thereby triggering a second flash of intuition that may compete with the first.” The second is to use “conscious verbal reasoning, such as considering the costs and benefits of each course of action.” The third is to get feedback from other people, “which then trigger in us new flashes of intuition followed by various kinds of reasoning.” Critically, these ways of overriding our intuitive responses—or integrating them into a more sophisticated perspective—must be triggered by the mind’s interaction with the task environment. Otherwise, the immediate response predominates.

The cybernetic framework leads to the following questions: What would force someone's attention to the fundamental political task? If she ignored it through inattention, what would trigger her attention? If she suffered from ignoring it through inattention, would she know it? Would she be able to attribute that suffering to her having ignored it?

If she attended to the fundamental political task, and then did a sloppy (i.e., irresponsible) job of addressing it, would she know it? If she suffered from that sloppy performance, would she know it?

The cybernetic framework can reconcile why so many people vote and why many, if asked, maintain that their individual votes could be decisive. First, the fact that a voter turns out to vote does not mean he believes his choice will be (or would have been) pivotal. A voter may believe his choice would be pivotal either through quasi-consequentialist Kantianism, asking "What if everyone else did what I did?" (Elster 2007; Hardin 2008), or through overestimating the probability that an election will be decided by one vote (Somin 2006). A voter may also turn out for non-instrumental reasons, such as patriotic or partisan duty or the expressive value of participation. We find these to be more plausible than proposed instrumental reasons for turnout.

Many who vote say, or would say if queried, that their vote "matters." Even if all of those who responded in this way meant they thought their vote would be pivotal with a high probability, none of them is compelled to deduce, much less act upon, all the implications of that belief. It takes training, time, and effort to get even a small subset of our beliefs to hang together as a consistent whole (Hardin 2009). Most citizens do not seem to gather, evaluate, and integrate knowledge to the extent they would if they truly expected to decide the election.

Inspired by Caplan (2007), we propose a field experiment to test whether voters stick by their assertion that their vote matters when the stakes are higher. First, citizens going into the polling place are asked how likely is it that their vote decides the election—perhaps with choices of (a) less than .001, (b) between .001 and .01, (c) between .01 and .1, (d) between .1 and .5, and (e) greater than .5. If they choose

something other than option (a), they are offered a \$100 bet with odds better than what is implied by their initial choice. We conjecture that most of those who would choose (b), (c), (d) or (e) would not take such a bet. Such an immediate reversal would suggest that they had (understandably) not taken the time to think the problem through. But many errors would not be so responsive to simply raising the stakes in the present moment.

Inspired by Tetlock (2005), we propose a second field experiment to test whether greater material incentives and/or social accountability would affect the extent to which citizens form an intention to take up the fundamental political task, take up the task, and execute it competently. Suppose we told voters a year in advance that they would be given a “test” after voting. Suppose the test covered non-controversial content on the subjects set out by Delli Karpini and Keeter (1996, 14): roughly, the identities of alternative candidates and parties, the policies they favor, and the processes through which policies are enacted. The dependent variable is the level of knowledge as measured by the score on the test. One independent variable is time available to improve performance. We might consider giving the test without any prior notice or giving six month’s notice in advance of the test. Another independent variable is the nature of motivation. We might consider relying on whatever accidental or intrinsic motivation citizens have, without any additional material reward or social accountability; or announcing that those scoring in the top half would be paid \$1,000; or announcing that each person might be called upon to explain her answers to a panel of unknown or unbiased authorities, or broadcast on television; or combining material reward and social accountability. We conjecture that those with notice would outperform those without it, and that those with material incentives, social motivation, or both would outperform those without. Raising the stakes alone is not enough to improve performance—in the short run it may increase stress and thereby hurt performance—but it is potent when combined with time to learn.

Consider a third “field experiment”—though this one we have no desire or intention to implement—to test whether the motive-formation process is responsive to incentives. Suppose we offered registered partisans \$500 in cash (or even \$500 donated to the charity of their choice) to switch their vote. This is not unlike the way bribery and

blackmail were used to influence voters in the nineteenth-century United States and in parts of the developing world today. The anonymous vote is a safeguard intended to keep voters from suffering a material and social cost for voting sociotropically. We would predict that many voters would take a bribe, as many voters used to do.

These field experiments—for now, thought experiments—are meant to shed light on the standard case. In the standard case, where participation is socially encouraged and there are only weak material incentives or social sanctions encouraging dispassionate inquiry (either by getting technical-causal details right or by admitting one’s ignorance), other forms of moral-social reasoning can flourish. The context of “other people” not only provides feedback as to where we should pay attention, and what intuitions we should have when we do; it also affects the motives for reasoning and the beliefs that emerge from it.

5. A Functionalist Framework for Political Reasoning

Our hypothesis is that those who seek to address the fundamental political task have little economic-material or technical-causal feedback to constrain their beliefs. This allows social feedback to play a greater role than it would when tempered by obvious causal relationships and material incentives.

In his account of the “new synthesis in moral psychology,” Jonathan Haidt (2007, 999) proposes that “moral thinking is for social doing.” This formulation comes originally from William James’s pragmatist dictum that “thinking is for doing.” Whereas the assumption in much of rational choice and behavioral economics alike is that the mind is or should be an intuitive scientist (seeking truth), we can explain more of how the mind works if we see the mind as an intuitive politician, prosecutor, and theologian, engaged in adaptive tasks like defending our behavior and beliefs in front of diverse constituents, detecting cheaters in cooperative relationships, assimilating the perspective of our peers, and/or creating a sacred identity that binds insiders together against outsiders. What would be cognitively “bad behavior” for the scientist can be reckoned as cognitively “good behavior” for the politician, prosecutor, or theologian. Specifically, we

propose that the nature of the mind's interaction with its moral-social task environment depends critically upon whether one engages with those one takes to be insiders or outsiders, and if insiders, whether they are superiors (or principals), subordinates (or agents), or peers.

In engaging with *superiors*, i.e., the “principals” to whom one is accountable as an “agent,” people act as “intuitive politicians,” which means that

they are accountable to a variety of constituencies, they suffer consequences when they fail to create desired impressions on key constituencies, and their long-term success at managing impressions hinges on their skill at anticipating objections that others are likely to raise to alternative courses of action and at crafting accounts that preempt those objections. (Tetlock 2002, 454)

Accountability should not be limited to the threat of losing one's job, as it is in standard principal-agent models. When expecting to be accountable to an unknown or divided audience, people engage in more preemptory self-criticism that increases the “integrative complexity” of their reasoning. That is, they are more likely to entertain alternative contradictory hypotheses and to express interest in generating, testing, and revising boundary conditions to integrate them (Lerner and Tetlock 1999). But these circumstances, while desirable, are rare. If instead the perspective of the audience is known and unified, people engage in attitude shifting to emphasize areas of agreement. When held accountable for commitments already made, people tend to engage in defensive positioning and self-justifying reasoning (Tetlock, Skitka, and Boettger 1989). Our claim is not that all such behavior is sincere, but rather that some of it is. Those who are able to sincerely and intuitively shift attitudes and/or rationalize choices would tend to come out ahead in many social interactions (Wright 1994).

In engaging with subordinates or agents one holds accountable, people act as “intuitive prosecutors,” which means that

as transmitters of accountability pressures onto others, people try . . . to detect cheaters and free riders who seek the benefits but shirk the responsibilities of membership in the collective. A key function of thought becomes closing loopholes in accountability regimes that unscrupulous intuitive politicians might otherwise exploit. (Tetlock 2002, 452)

This mindset is activated by a perception that cheating or misbehavior is on the rise and goes unpunished, and leads to the desire to tighten standards of accountability.

In engaging with peers—those who are neither superior nor inferior— the mind is drawn by the “shared reality” or “social proof” of local custom (Hardin and Higgins 1996), as well as the “social pressure” of attracting the approbation and avoiding the disapprobation of one’s peers (Gilovich 1991). For the purposes of sincere belief formation, this peer pressure must have its effect in a way that is consistent with the illusion of personal objectivity (e.g., by setting up a default hypothesis that is then validated via confirmation bias). One may also engage in “status seeking” that attempts to turn peers into inferiors, either by imitating the best local role models or otherwise signaling dominance (Frank 1987; Ridley 1995).

Finally, in engaging with outsiders, the mind is an “intuitive theologian,” which means that

people are posited to . . . have an existential need to believe that the rules governing their social world are not just the arbitrary preferences of currently dominant interest groups but rather are anchored in sacred values that confer legitimacy on collective practices. A key function of thought becomes protecting sacred values from secular encroachments. (Tetlock 2002, 452)

This framework helps explain why initially contingent means (e.g., “the way we do things around here”) can become a source of identity or pride, becoming ends in themselves. Herbert Simon called this “identification with means” (quoted in Jones 2001, 46). Identification with means slows or blocks the adoption of good ideas that were “not invented here” (also known as the “NIH syndrome”).

Clearly, electoral politics offers voters and politicians alike the opportunity to engage both insiders (superiors, peers, subordinates) and outsiders. Our hypothesis is that “we the people” are all too capable of defending our behavior and beliefs in front of diverse constituents, detecting cheaters in cooperative relationships, assimilating the perspective of our peers, and/or creating a sacred identity that binds insiders together against outsiders. Unfortunately, this has unintended and tragic consequences for the way in which citizens engage the fundamental political task.

6. Altruism and the Disregard for Oneself

Bearing burdens to bring joy to others can be an immense source of gratification and meaning. At the species level, it is clear that human beings have both the appetite and aptitude for other-regarding behavior. Several explanations of other-regarding behavior are often given. First, kin selection explains other-regarding behavior within an extended family (Hamilton 1964). Second, reciprocal altruism explains other-regarding behavior between friends (Trivers 1971; Cosmides and Tooby 2010). Third, “indirect reciprocity,” or building a reputation, explains other-regarding behavior between strangers within the same community (Haidt 2007).

There is a tension between the sincerity of altruistic behavior and the material costs and benefits of altruistic behavior. Other things equal, the greater the material cost that an altruistic actor must bear, the more sincere he is regarded as being. He who expects to receive a substantial material benefit for his altruistic behavior, however, is less likely to be regarded as sincerely altruistic. Risking one’s own life to help a stranger is rightly seen as more altruistic than risking one’s own life to help a stranger when there is known to be a substantial reward for doing so. When self-regarding and other-regarding motives would point in the same direction, it is harder to decipher the contribution of each.

When the net benefit or net cost of doing the right thing is zero, then material considerations do not sway behavior one way or the other. Conversely, a judge must recuse himself from a lawsuit where he has a material interest in one side winning, since—if justice demands that the other side win—he would then face a cost to serve justice (Kirchgassner and Pommerehne 1993). The higher the price of altruism, the more we worry—justifiably—that it will be overwhelmed by other motives.²⁸

Other things equal, the greater the material cost that an altruistic actor must bear,

²⁸ Montaigne (1993, 1157-1158) opines that “the more glittering the deed the more I subtract from its moral worth, because of the suspicion aroused in me that it was exposed more for glitter than for goodness: goods displayed are already halfway to being sold.”

the less likely he is to choose (and sustain) altruism. As considered above, this need not be the result of optimizing choice. It could be the result of activating reasoning only when costs are perceived ex post. And even if reasoning is never activated, there is the budget constraint. If you must give up your life to save the lives of a city of strangers, then you can do so only once. If you can save a city of strangers with one month of your attention, then you may be able to do so (at most) a dozen times a year. Other commitments, whether inviolable (e.g., to one's family, friends, employer and colleagues) or merely unexamined (e.g., to one's habits, to local norms), would take more time off the table.

If the price of voting is a couple of hours every couple of years, and the price of supporting the party one thinks best for the polity as a whole is effectively zero, then many may be sufficiently altruistic to show up and vote sociotropically. Fewer will show up when it rains or snows, but still, many do.²⁹ Unfortunately, it is several orders of magnitude more work to arrive at informed beliefs about people, parties, policies, and policy-making processes.

Nor is more effort alone sufficient. Even those who are fully motivated as a matter of professional and personal identity to arrive at informed beliefs will not be able to do so for all relevant issues. Effort helps with the simple problems. Sufficiently complex problems can sustain multiple interpretations, however, such that even fully motivated experts may end up in righteous disagreement. In such cases, we might hope not to motivate people to "know," but rather to embrace the Socratic insight that "I know that I don't know." And where that realization is not forthcoming, we might hope to motivate the mutual engagement of those who reasonably disagree with one another. But as we will see in the next section, it not easy to embrace the Socratic insight, nor is it easy to see those who disagree with us as reasonable.

²⁹ Brady and McNulty 2011 shows through a natural experiment that changing polling places in Los Angeles County during California's 2003 gubernatorial recall election substantially decreased voter turnout by raising transportation and search costs for voters.

7. Righteousness and the Disregard for Out-Group Members

Our enemies see us better than we see ourselves.

— La Rochefoucauld, 1851

Things are neither so good nor so bad as we suppose.

— La Rochefoucauld, 1851

As we define it, “righteousness” is disregard for the insights or the welfare of out-groups that flows from a feeling of moral superiority validated by a group of like-minded others. To understand righteousness, one must understand how differences of opinion arise and then how they are transformed into a moral conflict between well-informed, smart, good in-group members and poorly informed, stupid, bad out-group members.

We see differences of opinion as arising from problematic mind-environment interactions. The first step is for the environment to be sufficiently complex to sustain multiple hypotheses with the evidence at hand. Second, without knowing that they are doing so, people can make use of what we call an “epistemic double standard” to evaluate alternative hypotheses.

The epistemic double standard consists of both (cognitive) confirmation bias and (affective) disconfirmation bias. Confirmation bias occurs when individuals seek out confirming evidence for a proposition—a relatively easy test for a proposition to pass—and then stop, rather than seeking out disconfirming evidence as well, the harder test. One class of propositions fails the easy test, a second class can pass the easy test but fails the hard test, and a third class can pass both tests. Notably, confirmation bias comes from two non-motivated errors in the way we test hypotheses (Kunda and Nisbett 1986; Kunda 1987). First, we tend to frame matters by asking ourselves whether an agreeable question is true (Miller and Ross 1975): for example, “Am I intelligent and well-meaning?” Second, we systematically tend to seek confirmation rather than disconfirmation (Klayman and Ha 1987). To those without scientific training, this seems like a neutral way to proceed. Even those with scientific training find it difficult to avoid framing things in this way (Tetlock 2005, xv):

When advocates of a point of view are far off the mark, readers have an array of options, including concluding that (a) forecasters misinterpreted the theory; (b) forecasters had the right theory but no real-world savvy, so they fed the wrong antecedent conditions into the deductive machinery of their theory which, in the tradition of garbage in, garbage out, duly spat out idiotic predictions; (c) the theory is flawed in minor ways that tinkering can fix; (d) the theory is flawed in fundamental ways that require revising core assumptions. True believers in a theory will reach option (d) only after they have been dragged kicking and screaming through options (a), (b), and (c), whereas debunkers should leap straight to option (d) at the first hint of a glitch.

When we encounter an agreeable proposition in conversation with others, the standard is “can I believe this?” When we encounter a disagreeable proposition, the standard becomes “must I believe this?”

If confirmation bias consists of lowering the bar that privileged propositions must clear (and raising the bar that non-privileged propositions must clear, by seeking disconfirmation of them), disconfirmation bias (also known as the quantity-of-processing view) consists of trying harder to get the privileged propositions over the bar than to get the non-privileged propositions over. Disconfirmation bias occurs because challenges to one’s existing attitudes create negative affect (that is, concern or anxiety) that intensifies cognitive processing (Schwarz, Bless, and Bohner 1991; Ditto and Lopez 1992; Ditto et al. 1998). As with confirmation bias, people do not realize that disconfirmation bias leads to unreliable evaluations of alternative hypotheses.

While the epistemic double standard explains why we would usually see motivated reasoning, it also suggests why we would sometimes see counter-motivated reasoning (Elster 2007, 39), or the confirmation of unpleasant beliefs. In some cases, what we fear is set up as the default hypothesis that we then seek to confirm.

There is a third step after the epistemic double standard does its work. Although in reflective equilibrium our beliefs are likely to be incompletely de-biased, we may feel unbiased. It is important to emphasize that we form beliefs in a manner consistent with

the illusion of personal objectivity.³⁰ Our casual introspection shows no sign of remaining bias in our beliefs. This is what Jones 2002 refers to as the “First-Person Constraint on Doxastic Explanation.” While I may be able to explain my past beliefs in terms of biases, and while I may allow my future beliefs to be explained as a result of bias, my present beliefs—in order to be my beliefs—cannot be explained in that way. After all, they are my beliefs because I think they are true. When it comes to evaluating propositions like “I am good, well-informed, and insightful,” we have privileged access to a special kind of evidence: our own interior life. Even though we would not be impressed by somebody else saying, “I have looked inside my heart and found no bias,” when we look inside our own hearts and find no bias, we find it hard to believe we could be biased (Pronin et al. 2002).

After one has processed immediately available feedback, one may plausibly believe that “I see the world as it is.” This is the premise of what Ross and Ward 1996 and Pronin et al. 2004 refer to as naïve realism. Naïve realism is a near-universal psychological stance, not a philosophical position.³¹ A corollary of the premise that “I see the world as it is” is that I expect other well-informed, smart, capable, well-meaning people to see things the same way.

Our tendency to see our own perspective as objectively true is thus reinforced by “social proof” in the form of the agreement of our fellow partisans (Hardin and Higgins 1996).³² Indeed, rather than making ourselves vulnerable, we tend to think that those who agree with us are well informed, smart, and impartial.³³ The disagreement of those in the

³⁰ Demosthenes overstates the case with his dictum that “what a man wishes, that also will he believe.” Overly charitable self-assessments are not a matter of directly believing whatever one wants to believe.

³¹ Popper (1963, 7) calls this “the doctrine of manifest truth.”

³² Robinson et al. 1995: “Partisans on both sides of a given issue saw their own side’s views as more reflective of objective evidence and rational concerns, and less influenced by political ideology, than the other side’s views.”

³³ This has been tested in the laboratory by Pronin, Lin, and Ross (2004): “In other words, participants assumed that peers whose views mirrored their own had arrived at those views through a process that was just as objective and attuned to the realities of the situation at hand as the process they had followed themselves. By contrast, the participants attributed support of the position that was ‘next most acceptable’ more to

opposition therefore poses a problem of inference. They claim that we are overconfident and self-serving, but we hold that we are well informed, smart, and impartial. We might think to ourselves that as a community we have come to the same conclusions about a number of important topics, thereby inflating our confidence in our conclusions. Conversely, we infer that those in the opposition must be uninformed, misinformed, stupid, irrational, ideological, greedy, evil, or some combination thereof. Once we have labeled others in this way we must also discount their criticism of our beliefs, as well as underestimate—on the margin—the quality of the opposition’s leaders and policies.

Psychologists call the “fundamental attribution error” the tendency to overuse explanations based on character (“he is an angry person”) rather than context (“he is upset because he overslept and missed the bus”). The fundamental political attribution error is to impugn the character of those who disagree with us rather than seek to understand how mind-environment interactions could lead smart, well-informed, well-intentioned people on both sides to believe what they do.

The alchemy of righteousness has three steps. The complexity of the world makes alternative interpretations possible. The epistemic double standard inadvertently leads us to pick interpretations that frame ourselves in an overly charitable light. Naïve realism rules out the simplest answer to why others disagree with us (i.e., that we are biased or in error), and so pushes us to mark the error in their account rather than our own. Thus, our vanity breeds distrust for those who disagree with us. This transforms tractable disagreements about the consequences of alternatives into good versus evil conflicts.

8. Myopia and the Disregard for Distant Consequences

For the first few hundred million years after their initial appearance on our planet, all brains were stuck in the permanent present, and most brains still are today. But not yours and not mine, because two or three million years ago our ancestors began a great escape from the here and now, and their getaway vehicle was a highly specialized mass of gray tissue, fragile, wrinkled, and appended. This

biases and less to normative factors than they attributed their support of their own position. And, of course, they attributed the position they deemed least acceptable much more to biases and much less to normative factors than their own position.”

frontal lobe—the last of the human brain to evolve, the slowest to mature, and the first to deteriorate in old age—is a time machine that allows each of us to vacate the present and experience the future before it happens.

— Daniel Gilbert (2005, 16)

This section explores the causes and consequences of myopia, with particular focus on how the feedback we receive from the social world makes it difficult to overcome our myopic tendencies. For present purposes we use a weak notion of myopia, in which myopia means only failing to consider future consequences. Note that this is considerably more relaxed than the typical economist's definition, in which an individual is myopic if he fails to optimally consider all future periods and weigh them correctly, i.e., fails to discount exponentially.

Perhaps because the ability to imagine the future depends on the section of the brain that was last to evolve, it remains extremely difficult for people to imagine the future and to bring present actions into accord with the imagined future. To do this well, we need very good feedback regarding how our actions taken at t_0 match up with future consequences at t_n . As we argue here, such feedback is extremely hard to obtain.

The consequences of most actions are distributed across time. When the valence of distant consequences matches that of immediate consequences (i.e., they are both positive or both negative), then taking into account immediate consequences will lead to the same behavior as taking into account both immediate and distant consequences.

Evaluating relatively immediate alternatives and picking the best one—often called “gradient climbing”—allows us to find local maxima in the fitness landscape (Beinhocker 2006). Gradient climbing does not work as well when the fitness landscape has many peaks and valleys, and the local peak may be far from the best one. In such a landscape, one has to descend into a valley before one can climb a new peak. If we take into account only immediate consequences, then we forego many “investment-like” activities where up-front costs yield greater benefits in time, and indulge in too many “temptation-like” activities where immediate benefits yield greater costs in time.

Most species are gradient climbers, “stuck in time” (Roberts 2002), living in the

permanent present. Many animal species have the ability to project the immediate future (e.g., to anticipate the trajectory of objects in motion), but only human beings have the cognitive ability to time-travel beyond the immediate future.³⁴ The frontal lobe gives us the ability to generate representations of remote, possible futures (Wheeler, Stuss and Tulving 1997).³⁵ This allows desirable possibilities to attract us despite immediate burdens, and undesirable possibilities to repel us despite immediate temptations.

Our ability to take the future into account is substantial but imperfect. It takes a great degree of skill to generate the set of relevant alternatives, to project their distant consequences, to evaluate the costs and benefits of these consequences, and then to collapse the streams of expected costs and benefits into a summary score or scores. In many cases, finding an alternative that looks good for the immediate future is good enough. Even in this satisficing sense, however, our ability to form somewhat accurate expectations depends upon feedback available from the task environment. Our expectations are likely to be of higher quality (1) where the cause-and-effect linkages are simple, immediate, and responsive to our control, (2) where the material costs of bad decisions are clear, and (3) where we have access to successful and unsuccessful examples of others' experiments that are further along than our own, and where the successful reap approbation and/or the unsuccessful reap disapprobation.

In the political task environment, however, the cause-and-effect linkages are condition-specific, interdependent, lagged, and dispersed. The material costs of bad decision-making are not clear. And instead of having thousands or millions of peers (i.e., polities) to learn from, we have at most several hundred.³⁶ Finally, political myopia may be justified by our lack of control. It is nearly impossible to control what happens many years down the road due to an inability to credibly commit to actions in the future. Therefore, it is quite reasonable to focus attention on the here and now, or at least on the

³⁴ The discussion in Gilbert 2005 is singularly instructive and entertaining. See also Roberts 2002 and Haith 1997.

³⁵ Children seem to develop the ability to time-travel in this sense when they are between four and six years old (Friedman, Gardner, and Zubin 1995).

³⁶ Moreover, the ease of "imitative learning" or transferring lessons from one to the other is limited insofar as citizens see their own polity as fundamentally dissimilar, or insofar as they fall prey to the "Not Invented Here" syndrome mentioned above.

relevant choice horizon within which credible commitments can be made. Even if we could perfectly discern the consequences of party, policy, and policy-making alternatives for decades into the future, we may act in a way that is indistinguishable from those who think only of the present. If we cannot bind future decision makers—as in many cases we cannot—it may make more sense to focus on what we can control.

In sum, our miraculous ability to “vacate the present and experience the future before it happens” does not work well when we lack the right kind of technical-causal, material-economic, and moral-social feedback. This makes it difficult to appropriately calibrate the relationship between present decisions and future consequences.

* * *

We have tried to make two contributions to the literature on public opinion, political knowledge, and political behavior. First, we have offered a novel framework of citizen belief formation that synthesizes key insights from economic models of belief formation, scholarship on the complexity of social and political task environments, and work on the psychology of belief formation and decision making. Rather than focusing on the anticipation of costs and benefits, we have zeroed in on reaction to feedback from the environment. This feedback can be “technical-causal,” “material-economic,” or, most importantly for our purposes, “moral-social.”

The second contribution is to pose a dilemma. The same features of political task environments that enable altruism also enable righteousness and myopia. This leads citizens to focus on winning the high-stakes election that is immediately and concretely before them rather than on discovering, enforcing, and adapting institutions that are robust, fair, and mutually productive. The epic battle of good versus evil captures our hearts, and as a result, our minds neglect the impersonal, dull, and technical task of continuously improving the governance processes and structures that allow us to live well together despite our differences.

Madison’s focus on such processes was justified in part by the notion that “enlightened statesmen will not always be at the helm.” But were we to buy a drink for

each helmsman throughout American history, most would confess to being enlightened, or at least more enlightened than the alternatives. Most of their opponents would have the opposite perspective. They cannot all be right. A pessimist might suppose that all are right about their opponents and wrong about themselves. This seems too pessimistic to us, but it is probably closer to the truth than our own naïve assessments. We would conjecture — following La Rochefoucauld — that “our enemies see us better than we see ourselves.” If they think we are a “0” and we think we are a “10,” the truth is likely to be that we are less than a 5.

One task for future research is to better conceptualize and study the causes, nature, and consequences of altruism, righteousness, and myopia. Another is to craft institutions that harness righteousness and myopia to mutually productive ends.³⁷ What institutions, if any, would lower the stakes of electoral conflict when contestants live in a deeply complex but deceptively simple Manichean world? What institutions, if any, would enable myopic and righteous participants to attend to discovering, enforcing, and adapting better rules of the game? And even if there are institutions that would work well, if chosen, what reason do we have to believe they could in fact be chosen from within the status quo? These are questions we can ask only once we acknowledge the empirical realities of how citizens form motives and beliefs in a complex world.

³⁷ For our efforts so far, see Durant 2011, Durant and Weintraub 2011a, and Durant and Weintraub 2011b.

REFERENCES

- Akerlof, George. 1970. "The Market for 'Lemons': Quality Uncertainty and the Market Mechanism." *The Quarterly Journal of Economics* 84(3): 488-500.
- Akerlof, George. 2002. "Behavioral Macroeconomics and Macroeconomic Behavior." *American Economic Review* 92(3) (June): 411-33.
- Ariely, Dan. 2010. *The Upside of Irrationality: The Unexpected Benefits of Defying Logic at Work and at Home*. New York: HarperCollins.
- Becker, Gary. 1962. "Irrational Behavior and Economic Theory." *Journal of Political Economy* 70(1): 1-13.
- Beinhocker, Eric D. 2006. *The Origin of Wealth: Evolution, Complexity, and the Radical Remaking of Economics*. Cambridge, Mass.: Harvard Business School Press.
- Bendor, Jonathan, Daniel Diermeier, David A. Siegel, and Michael M. Ting. 2011. *A Behavioral Theory of Elections*. Princeton: Princeton University Press.
- Bennett, Stephen Earl, and Jeffrey Friedman. 2008. "The Irrelevance of Economic Theory to Understanding Economic Ignorance." *Critical Review* 20(3): 195-258.
- Berelson, Bernard R., Paul F. Lazarsfeld, and William N. McPhee. 1954. *Voting: A Study of Opinion Formation in a Presidential Campaign*. Chicago: University of Chicago Press.
- Boettke, Peter, and Peter Leeson. 2004. "Liberalism, Socialism, and Robust Political Economy." *Journal of Markets and Morality* 7(1): 1-20.
- Bower, Gordon H., and Ernest Hilgard. [1966] 1981. *Theories of Learning*. Englewood Cliffs, N.J.: Prentice-Hall.
- Boyd, Robert, and Peter J. Richerson. 2005. *The Origin and Evolution of Cultures*. Oxford: Oxford University Press.
- Brady, Henry E., and John E. McNulty. 2011. "Turning Out to Vote: The Costs of Finding and Getting to the Polling Place." *American Political Science Review* 105: 115-134.
- Brennan, Geoffrey, and Loren E. Lomasky. 1997. *Democracy and Decision: The Pure Theory of Electoral Preference*. Cambridge: Cambridge University Press.
- Buchanan, James. 1972 [2000]. "Toward Analysis of Closed Behavioral Systems." In *Politics as Public Choice*, ed. Robert Tollison. Indianapolis, Ind.: Liberty Fund.
- Buchanan, James, and Gordon Tullock. 1962. *The Calculus of Consent: Logical Foundations of Constitutional Democracy*. Ann Arbor: University of Michigan Press.
- Campbell, Angus, Philip E. Converse, Warren E. Miller, and Donald E. Stokes.

1960. *The American Voter*. Chicago: University of Chicago Press.
- Caplan, Bryan. 2001. "Rational Irrationality and the Microfoundations of Political Failure." *Public Choice* 107(3-4): 311-31.
- Caplan, Bryan. 2003. "The Logic of Collective Belief." *Rationality and Society* 15: 218-42.
- Caplan, Bryan. 2007. *The Myth of the Rational Voter*. Princeton University Press.
- Caplan, Bryan. 2008. "Roundtable 2: Ignorance and Error." *Critical Review* 20(4): 445-61.
- Chaiken, Shelly, and Yaacov Trope. 1999. *Dual-Process Theories in Social Psychology*. New York: Guilford Press.
- Coase, Ronald. 1960. "The Problem of Social Cost." *Journal of Law and Economics* 3: 1-44.
- Converse, Philip E. [1964] 2006. "The Nature of Belief Systems in Mass Publics." *Critical Review* 18(1-3): 1-74.
- Cosmides, Leda, and John Tooby. 2007. "Evolutionary Psychology, Ecological Rationality, and the Unification of the Behavioral Sciences: Comment on A Framework for the Unification of the Behavioral Sciences by Gintis." *Behavioral and Brain Sciences* 30(1): 42-43.
- Cosmides, Leda, and John Tooby. 2010. "Groups in Mind: The Coalitional Roots of War and Morality." In *Human Morality and Sociality: Evolutionary and Comparative Perspectives*, ed. Henrik Høgh-Olesen. New York: Palgrave Macmillan.
- Cowen, Tyler. 2005. "Self-Deception as the Root of Political Failure." *Public Choice* 124: 437-51.
- Delli Carpini, Michael X., and Scott Keeter. 1996. *What Americans Know about Politics and Why It Matters*. New Haven: Yale University Press.
- Demsetz, Harold. 1969. "Information and Efficiency: Another Viewpoint." *Journal of Law and Economics* 12(1): 1-22.
- Ditto, Peter H., and David F. Lopez. 1992. "Motivated Skepticism: Use of Differential Decision Criteria for Preferred and Nonpreferred Conclusions." *Journal of Personality and Social Psychology* 63: 568-84.
- Ditto, Peter H., James A. Scepansky, Geoffrey D. Munro, Anne Marie Apanovitch, and Lisa K. Lockhart. 1998. "Motivated Sensitivity to Preference-Inconsistent Information." *Journal of Personality and Social Psychology* 75(1): 53-69.

- Dorner, Dietrich. 1997. *The Logic of Failure: Recognizing and Avoiding Error in Complex Situations*. Cambridge, Mass.: Perseus Press.
- Durant, Thomas Clark. 2011. "Making Executive Politics Mutually Productive and Fair." *Constitutional Political Economy* 22(2): 141-72.
- Durant, Thomas Clark, and Michael Weintraub. 2011a. "From Patron-Client Mobilization to Constitutionally Constrained Competition." Working paper.
- Durant, Thomas Clark, and Michael Weintraub. 2011b. "Enabling Constitutional Solutions to Political Problems." Working paper.
- Dunning, David, and Justin Kruger. 1999. "Unskilled and Unaware of It: How Difficulties in Recognizing One's Own Incompetence Lead to Inflated Self-Assessments." *Journal of Personality and Social Psychology* 77(6): 1121-34.
- Edelman, Murray. 1988. *Constructing the Political Spectacle*. Chicago: University of Chicago Press.
- Elster, Jon. 2007. *Explaining Social Behavior: More Nuts and Bolts for the Social Sciences*. Cambridge: Cambridge University Press.
- Elster, Jon, and H el ene Landemore. 2008. "Ideology and Dystopia." *Critical Review* 20(3): 273-89.
- Evans, Anthony, and Jeffrey Friedman. 2011. "'Search' vs. 'Browse': A Theory of Error Grounded in Radical (Not Rational) Ignorance." *Critical Review* 23 (1-2): 73-104.
- Fiorina, Morris. 1990. "Information and Rationality in Elections." In *Information and Democratic Processes*, ed. John Ferejohn and James H. Kuklinski. Urbana: University of Illinois Press.
- Flack, Jessica C., and Frans B. M. de Waal. 2000. "Any Animal Whatever: Darwinian Building Blocks of Morality in Monkeys and Apes." *Journal of Consciousness Studies* 7(1-2): 1-29.
- Frank, Robert H. 1987. *Choosing the Right Pond: Human Behavior and the Quest for Status*. Oxford: Oxford University Press.
- Friedman, Jeffrey. 1996. "Public Ignorance and Democracy." *Critical Review* 10(4): i-xii.
- Friedman, Jeffrey. 1997. "What's Wrong with Libertarianism?" *Critical Review* 11(3): 407-67.
- Friedman, Jeffrey. 1998. "Public Ignorance and Democratic Theory." *Critical Review* 12(4): 397-411.
- Friedman, Jeffrey. 2003. "Public Opinion: Bringing the Media Back In." *Critical Review*

15(2-3): 239-60.

Friedman, Jeffrey. 2005. "Popper, Weber, and Hayek: The Epistemology and Politics of Ignorance." *Critical Review* 17(1-2): i-lviii.

Friedman, Jeffrey. 2006a. "Public Competence in Normative and Positive Theory: Neglected Implications of 'The Nature of Belief Systems in Mass Publics.'" *Critical Review* 18(1-3): i-xliii.

Friedman, Jeffrey. 2006b. "Taking Ignorance Seriously: Rejoinder to Critics." *Critical Review* 18(4): 469-532.

Friedman, Jeffrey. 2007a. "Ignorance as a Starting Point: From Modest Epistemology to Realistic Political Theory." *Critical Review* 19(1): 1-22.

Friedman, Jeffrey. 2007b. "A Weapon in the Hands of the People: the Rhetorical Presidency in Historical and Conceptual Context." *Critical Review* 19(2-3): 197-240.

Friedman, Jeffrey, and Wladimir Kraus. 2011. *Engineering the Financial Crisis: Systemic Risk and the Failure of Regulation*. Philadelphia: University of Pennsylvania Press.

Friedman, William J., Amanda G. Gardner, and Naomi R. E. Zubin. 1995. "Childrens' Comparisons of the Recency of Two Events from the Past Year." *Child Development* 66: 970-83.

Gigerenzer, Gerd. 1999. "The Adaptive Toolbox." In *Bounded Rationality: The Adaptive Toolbox*, ed. Gerd Gigerenzer and Reinhard Selten. Cambridge, Mass.: MIT Press.

Gilbert, Daniel. 1999. "What the Mind's Not." In *Dual-Process Theories in Social Psychology*, ed. Shelly Chaiken and Yaacov Trope. New York: Guilford Press.

Gilbert, Daniel. 2005. *Stumbling on Happiness*. New York: Alfred A. Knopf.

Gilovich, Thomas. 1991. *How We Know What Isn't So: The Fallibility of Human Reason in Everyday Life*. New York: The Free Press.

Gilovich, Thomas, Dale Griffin, and Daniel Kahneman, eds. 2002. *Heuristics and Biases: the Psychology of Intuitive Judgment*. Cambridge: Cambridge University Press.

Glaeser, Edward L. 2004. "Psychology and the Market." *American Economic Review* 94(2): 408-13.

Glassner, Barry. 1999. *The Culture of Fear: Why Americans Are Afraid of the Wrong Things*. New York: Basic Books.

Gode, Dhananjay K., and Shyam Sunder. 1997. "What Makes Markets Allocationally Efficient?" *Quarterly Journal of Economics* 112(2): 603-30.

- Haidt, Jonathan. 2007. "The New Synthesis in Moral Psychology." *Science* 316(5827): 998-1002.
- Haith, Marshall M. 1997. "The Development of Future Thinking as Essential for the Emergence of Skill in Planning." In *Why, How and When Do We Plan? The Developmental Psychology of Planning*, ed. S. Friedman and E. Scholnick. Hillsdale, N.J.: Erlbaum Associates.
- Hamilton, W. D. 1964. "The Genetical Evolution of Social Behaviour." *Journal of Theoretical Biology* 7: 1-52.
- Hardin, Curtis, and E. Tory Higgins. 1996. "Shared Reality: How Social Verification Makes the Subjective Objective." In *The Interpersonal Context. Vol. 3 of Handbook of Motivation and Cognition*, ed. Richard Sorrentino and E. Tory Higgins. New York: Guilford Press.
- Hardin, Russell. 2008. "Ignorant Democracy." *Critical Review* 18(1-3): 179-95.
- Hardin, Russell. 2009. *How Do You Know? The Economics of Ordinary Knowledge*. Princeton: Princeton University Press.
- Hume, David. [1742] 1953. "On the Independence of Parliament." In *David Hume's Political Essays*, ed. Charles W. Hendel. New York: Liberal Arts Press.
- Jervis, Robert. 1998. *System Effects: Complexity in Social and Political Life*. Princeton: Princeton University Press.
- Jones, Bryan D. 2001. *Politics and the Architecture of Choice: Bounded Rationality and Governance*. Chicago: University of Chicago Press.
- Jones, Ward E. 2002. "Explaining Our Own Beliefs: Non-Epistemic Believing and Doxastic Instability." *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition* 111(3) (December): 217-49.
- Kahneman, Daniel. 2011. *Thinking Fast and Slow*. New York: Farrar, Strauss and Giroux.
- Kahneman, Daniel, Paul Slovic, and Amos Tversky, eds. 1982. *Judgment under Uncertainty: Heuristics and Biases*. Cambridge: Cambridge University Press.
- Kinder, Donald, and D. Roderick Kiewiet. 1979. "Economic Discontent and Political Behavior: The Role of Personal Grievances and Collective Economic Judgements in Congressional Voting." *American Journal of Political Science* 23: 495-527.
- Kinder, Donald R., and D. Roderick Kiewiet. 1981. "Sociotropic Politics: The American Case." *British Journal of Political Science* 11(2): 129-61.
- Kirchgassner, Gebhard, and Werner W. Pommerehne. 1993. "Low-Cost Decisions as a Challenge to Public Choice." *Public Choice* 77: 107-15.

- Klayman, Joshua, and Young-won Ha. 1987. "Confirmation, Disconfirmation, and Information in Hypothesis-Testing." *Psychological Review* 94(2): 211-28.
- Kunda, Ziva. 1987. "Motivated Inference: Self-Serving Generation and Evaluation of Causal Theories." *Journal of Personality and Social Psychology* 53: 636-647.
- Kunda, Ziva, and Richard E. Nisbett. 1986. "The Psychometrics of Everyday Life." *Cognitive Psychology* 18: 195-224.
- La Rochefoucauld, Francois de. 1851. *Moral Reflections, Sentences and Maxims*. New York: William Gowans.
- Lerner, Jennifer S., and Tetlock, Philip E. 1999. "Accounting for the Effects of Accountability." *Psychological Bulletin* 125: 255-75.
- Lewin, Leif. 1991. *Self-Interest and Public Interest in Western Politics*, trans. Donald Lavery. Oxford: Oxford University Press.
- Loewenstein, George, and Jon Elster, eds. 1992. *Choice over Time*. New York: Russell Sage Foundation.
- Lippmann, Walter. [1922] 1997. *Public Opinion*. New York: Free Press.
- Lupia, Arthur. 2008. "Roundtable 2: Ignorance and Error." *Critical Review* 20(4): 445-61.
- Mansbridge, Jane J. 1990. *Beyond Self-Interest*. Chicago: University of Chicago Press.
- Miller, Dale T., and Michael Ross. 1975. "Self-Serving Biases in the Attribution of Causality: Fact or Fiction?" *Psychological Bulletin* 82: 213-25.
- Montaigne, Michel de. 1993. *The Complete Essays*. trans. M. A. Screech. London: Penguin Books.
- Mueller, John. 2006. "Is There Still a Terrorist Threat? The Myth of the Omnipresent Enemy." *Foreign Affairs* (September/October).
- Pennington, Mark. 2011. *Robust Political Economy: Classical Liberalism and the Future of Public Policy*. Cheltenham: Edward Elgar.
- Pinker, Steven. [1997] 2009. *How the Mind Works*. New York: W.W. Norton and Company.
- Popper, Karl. 1963. *Conjectures and Refutations: The Growth of Scientific Knowledge*. London: Routledge.
- Pronin, Emily, Daniel Y. Lin, and Lee Ross. 2002. "The Bias Blind Spot: Perceptions of Bias in Self Versus Others." *Personality and Social Psychology Bulletin* 28: 369-81.

- Pronin, Emily, Lee Ross, and Thomas Gilovich. 2004. "Objectivity in the Eye of the Beholder: Divergent Perceptions of Bias in Self Versus Others." *Psychological Review* 111(3): 78-99.
- Rasmussen. 2011. "83% Say One Person's Vote Really Matters." June 12. http://www.rasmussenreports.com/public_content/politics/general_politics/june_2011/83_say_one_person_s_vote_really_matters
- Reason, James. 1990. *Human Error*. Cambridge: Cambridge University Press.
- Ridley, Matt. 1995. *The Red Queen: Sex and the Evolution of Human Nature*. New York: Penguin Books.
- Riker, William, and Peter Ordeshook. 1968. "A Theory of the Calculus of Voting." *American Political Science Review* 62(1): 25-42.
- Roberts, William A. 2002. "Are Animals Stuck in Time?" *Psychological Bulletin* 128(3): 473-89.
- Robinson, Robert J., Dacher Keltner, Andrew Ward, and Lee Ross. 1995. "Actual Versus Assumed Differences in Construal: 'Naïve Realism' in Intergroup Perception and Conflict." *Journal of Personality and Social Psychology* 68(3): 404-17.
- Rodrik, Dani, Arvind Subramanian, and Francesco Trebbi. 2004. "Institutions Rule: The Primacy of Institutions Over Geography and Integration in Economic Development." *Journal of Economic Growth* 9: 131-165.
- Ross, Lee, and Andrew Ward. 1996. "Naïve Realism in Everyday Life: Implications for Social Conflict and Misunderstanding." In *Values and Knowledge*, ed. Edward S. Reed, Elliot Turiel, and Terrance Brown. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Rozenblit, Leonid, and Frank Keil. 2002. "The Misunderstood Limits of Folk Science: An Illusion of Explanatory Depth." *Cognitive Science* 26: 521-62.
- Schulz, Kathryn. 2010. *Being Wrong: Adventures in the Margin of Error*. New York: HarperCollins.
- Schumpeter, Joseph. [1942] 1950. *Capitalism, Socialism, and Democracy*. New York: Harper & Row.
- Schwarz, Norbert, Herbert Bless, and Gerd Bohner. 1991. "Mood and Persuasion: Affective States Influence the Processing of Persuasive Communications." *Advances in Experimental Social Psychology* 24: 161-99.
- Sears, David O., Richard R. Lau, Tom R. Tyler, and Harris M. Allen, Jr. 1980. "Self-Interest vs. Symbolic Politics in Policy Attitudes and Presidential Voting." *American Political Science Review* 74(3): 670-84.

- Simon, Herbert. 1955. "A Behavioral Model of Rational Choice." *Quarterly Journal of Economics* 69(1): 99-118.
- Simon, Herbert. 1971. "Designing Organizations for an Information-Rich World." In *Computers, Communication, and the Public Interest*, ed. Martin Greenberger. Baltimore: Johns Hopkins Press.
- Simon, Herbert. 1985. "Human Nature in Politics: The Dialogue of Psychology with Political Science." *American Political Science Review* 79(2)(June): 293-304.
- Simon, Herbert. 1996. *Science of the Artificial*. Cambridge, Mass.: MIT Press.
- Smith, Vernon. 2008. *Rationality in Economics: Constructivist and Ecological Forms*. Cambridge: Cambridge University Press.
- Somin, Ilya. 1998. "Voter Ignorance and the Democratic Ideal." *Critical Review* 12(4): 413-58.
- Somin, Ilya. 2000. "Do Politicians Pander?" *Critical Review* 14(2-3): 147-55.
- Somin, Ilya. 2004. "Pragmatism, Democracy, and Judicial Review: Rejoinder to Posner." *Critical Review* 16(4): 473-81.
- Somin, Ilya. 2006. "Knowledge about Ignorance: New Directions in the Study of Political Information." *Critical Review* 18(1-3): 255-78.
- Spence, Michael. 1973. "Job Market Signaling." *The Quarterly Journal of Economics* 87(3): 355-74.
- Spence, Michael. 2002. "Signaling in Retrospect and the Informational Structure of Markets." *The American Economic Review* 92(3) (June): 434-59.
- Stiglitz, Joseph E. 2002. "Information and the Change in the Paradigm in Economics." *The American Economic Review* 92(3) (June): 460-501.
- Tasic, Slavisa. 2009. "The Illusion of Regulatory Competence." *Critical Review* 21(4): 423-36.
- Tetlock, Philip E. 2002. "Social Functionalist Frameworks for Judgment and Choice: Intuitive Politicians, Theologians, and Prosecutors." *Psychological Review* 109(3): 451-71.
- Tetlock, Philip E. 2005. *Expert Political Judgment: How Good Is It? How Can We Know?* Princeton: Princeton University Press.
- Tetlock, Philip E., Linda Skitka, and Richard Boettger. 1989. "Social and Cognitive Strategies of Coping with Accountability: Conformity, Complexity, and Bolstering." *Journal of Personality and Social Psychology: Interpersonal Relations and Group*

Dynamics 57: 632-41.

Trivers, Robert. 1971. "The Evolution of Reciprocal Altruism." *The Quarterly Review of Biology* 46(1): 35-57.

Weingast, Barry, and Ken Shepsle. 1984. "When Do Rules of Procedure Matter?" *Journal of Politics* 46: 206-221.

Wheeler, Mark, Donald Stuss, and Endel Tulving. 1997. "Toward a Theory of Episodic Memory: The Frontal Lobes and Autonoetic Consciousness." *Psychological Bulletin* 121(3): 331-54.

Wiener, Norbert. 1948. *Cybernetics: Or Control and Communication in the Animal and the Machine*. Cambridge, Mass.: MIT Press.

Wright, Robert. 1994. *The Moral Animal*. New York: Vintage Books.

Zajonc, Robert B. 1980. "Feeling and Thinking: Preferences Need No Inferences." *American Psychologist* 35: 151-75.

Zukin, Cliff, Scott Keeter, Molly Andolina, Krista Jenkins, and Michael X. Delli Carpini. 2006. *A New Engagement? Political Participation, Civic Life, and the Changing American Citizen*. New York: Oxford University Press.