

From Greenspan's Despair to Obama's Hope: The Scientific Bases of Cooperation as Principles of Regulation

Yochai Benkler

"Those of us who have looked to the self-interest of lending institutions to protect shareholders' equity (myself especially) are in a state of shocked disbelief."

—Written testimony of Dr. Alan Greenspan,

Committee of Government Oversight and Reform, October 23, 2008

HENRY WAXMAN: *In other words you found that your view of the world, your ideology was not right. It was not working.*

ALAN GREENSPAN: *Precisely, no I, that's precisely the reason I was shocked because I've been going for forty years or more with very considerable evidence that it was working exceptionally well.*

—Committee of Government Oversight and Reform hearing,

October 23, 2008

"We have never been just a collection of individuals. . . ."

—Victory speech of President Barack Obama, November 4, 2008

Morning, November 4, 2008. We're standing in line, my two sons, my wife, and I; waiting to volunteer at the Obama campaign headquarters in Raleigh, North Carolina; waiting to be told how to be useful in this battleground state. We have come down from Cambridge, Massachusetts for the last few days of the campaign, and have met many people. This morning the campaign headquarters

is brimming with people, black and white, young and old. A well-turned-out middle-class white woman with a lilting Southern accent is sitting at a phone-bank table next to a young African American student in his early twenties; both are calling people to say, *Go vote!* Voting, the standard rational-actor model would normally say, is irrational. Volunteering: inexplicable. My nine-year-old, on the phone, suddenly lights up—“Yes,” he says, “of course I’m real!” The person on the other end of the line is clearly warming to the young voice full of seriousness, and they take obvious pleasure from talking to each other, if only for a few seconds: “I voted,” the person seems to say, as Ari smiles broadly and replies, “That’s wonderful!” A human connection.

The campaign of candidate Obama—and, if he can convert it into a well-designed, thoughtful practice of governance, the presidency of President Obama—must give us the answer to the questions that all of us looking at the enormous challenges of tomorrow, and specifically at the global economic crisis, must face: How do we think about practical governance now that the model of well-designed incentives for selfish beings has collapsed around us? How do we convert what has inspired millions in the Obama campaign into a practical replacement for the economic model that Alan Greenspan says he has been relying on for forty years, and which has now failed him in the fall of 2008?

The answer is that we already have the building blocks of a new approach to organizing production and consumption, to organizing governance and practical problem solving in both the public and private arenas. It is this approach that is responsible for developing the free and open-source software that runs the vast majority of Web sites and Internet functions we use every day. It is the approach that allowed Toyota to build better and more productive relations with the very same employees and suppliers that General Motors had alienated. It is the approach that has led contemporary evolutionary biology to see cooperation, rather than competition alone, as a fundamental force of evolution. It is the approach that has motivated the rapid spread of community-policing initiatives in the majority of police departments around the country. In none of these cases is the alternative to the self-interest model that animated Greenspan’s ideology a starry-eyed reliance on hope, or a generalized belief in human benevolence. Rather, in every case, a carefully designed or organically grown system structures and gives form to a set of basic assumptions very different from Greenspan’s: the assumption that a majority of us care not only about ourselves, but also about others with whom we interact and the groups we see ourselves as belonging to; that we care not only about what is in it for us in a given situation (although we care about that too), but also about what is the right and fair thing to do; and that most of us care about the social context we live in, and respond to our understanding of a social situation as social beings, and not merely as a

collection of individuals, relentlessly driven to maximize our own returns and to coordinate for our common good only when the incentives are set just right to lead us by the nose to that result.

During the week or so after Alan Greenspan's testimony before Congress, we saw several editorials and comments about behavioral economics, or what some have derisively derided as merely "the economics of stupid people." Behavioral economics mainly studies how people fail to think rationally or to act upon what would theoretically be in their rational self-interest, supposedly because human beings pervasively lack self-discipline. Nonetheless, the discipline's literature would tend to lead policymakers to consider themselves to be somehow immune to this human failing. Rather, legislators, for example, who are influenced by behavioral economics, may flatter themselves that they alone can think clearly, and could set up systems that manipulate the population into doing what their leaders, from an Archimedean standpoint outside the psychological universe of cognitive biases, know is the right outcome for everyone.

However, the key to constructing future regulatory systems does not lie in perfecting the economic theory of how stupid people behave. Nor does it lie in attentive examination of how people make systematic errors in perception and judgment, or experience periodic failures of will. At least, we cannot rely on such studies alone, although they are important. The term in Greenspan's statement that we need to focus on is *self-interest*. If self-interest were the universal motivator, there would be no Wikipedia, no Linux or Apache, or any of the free software applications that run the Internet that we all depend on. If self-interest were the universal motivator, millions of volunteers would not have gone out canvassing for their candidates, nor would they be out in more normal times, serving meals to the homeless, or cleaning up their city parks. If self-interest led to success in business, then it would be Toyota asking for a government handout to help keep it afloat, not General Motors—because it was GM that perfected a production model based on self-interest, as shown in its conflict-ridden negotiations with its unionized employees, in the aggressively competitive bidding that the company has encouraged among its suppliers, and in its incentive compensation schemes for managers that should (according to agency theory, at least) have led to the best management decisions, but plainly did not. By contrast, it was Toyota that took the management-science world by surprise a quarter of a century ago by building a system based on trust, teamwork, and commitment to a set of shared values, instead of trying to perfectly align the self-interest, and only the self-interest, of workers, managers, suppliers, and distributors.

A growing body of work, in disciplines as wide ranging as experimental economics and psychology, human evolutionary biology and neuroscience, and

political science, organizational sociology, and management theory, has given us a basis for rejecting, not rationality, but selfishness, as the prime, universal motive of human action. What we find instead, in thousands of experiments and real-world studies, is what we all actually already know from our day-to-day lives: people are not universally selfish. Some are, to be sure. But many, a majority, are not. Many of us care about what happens to others, about doing what is right and fair and appropriate. We care about being trustworthy, and are affected by the social dynamics of the situations in which we find ourselves. The challenge we face is to take what we know from everyday life and from the fact of successful cooperation all around us, combine this knowledge with what are learning at the cutting edge of the social and human sciences, and develop a new model of human action and motivation that will allow us to design cooperative human systems—like Wikipedia and Linux, like Toyota or the community policing movement.

What are the tools that will make up this new approach to cooperative human-systems design? First, a tremendous amount of work is being done in the experimental study of human behavior, mostly in economics and psychology. In these studies, total strangers are brought into laboratory environments where they sit in front of computers and interact with other people, who they typically have never seen, do not see during the experiment, and will never see after the experiment is done. In many of these experiments, people interact with one another in ways that have real economic consequences for them, gaining or losing from a few dollars to (in one experiment) as much as three months' salary. Because the setups are so spare, people are expected to conform as closely as anyone ever would to the predictions of the standard economic models in which Greenspan put such faith. One fact stands out from these studies, above all else: *in no human society examined under controlled conditions have the majority of people consistently behaved like selfish rational actors.* This has been proved true in hundreds of experiments, in more than two dozen countries.

These experiments allow researchers to be very precise in their assumptions about human cooperation, in what they alter under experimental conditions about the nature of social relationships, and in determining what their results might imply. Such experiments gain precision from their reductiveness. Thus a long-term relationship between people might be represented in an experiment by having a pair of people, who are represented to each other only by onscreen icons, repeatedly play a game, ten or even forty times. Participants might be allowed, at most, to spend five minutes face-to-face during a break before going back to sit in front of their computer screens.

Such artificial restrictions create, at most, thin representations of what human relations really are. For this reason, it has been important to enrich experimental

studies with real-world observation, which has been undertaken in many different disciplines. Political scientists might look at the Chicago policing and educational systems reforms of the 1990s and 2000s, or they might look at cooperative irrigation systems around the world. Organizational sociologists and management scientists might study the corporate cultures and strategies of different companies to get rich case-based insights into what has worked in the real world. More recently, the Internet has provided us with a breathtaking range of cooperative activities available for study, many of which have generated rich archives, reveal clear paths of relations, and have created a virtual treasure trove of practices allowing us to study what makes cooperation succeed—or fail. Together, these sources of insight, combined with experimental and theoretical work in game theory and evolutionary dynamics that allows us to place our observations in context, are providing us with a powerful set of tools with which to build new systems for human cooperation, systems built on more than a simple mixture of Hobbes’s *Leviathan*—with the strict control hierarchies that his thinking implies—and Adam Smith’s invisible hand, with its relentlessly individual incentives-driven focus.

What are the elements of cooperative human-systems design? A useful way of organizing the conclusions of the thousands of articles in many disciplines looking at this question, like the wise men examining an elephant, is to group their findings into core “levers” that encourage cooperation. The literature strongly suggests that attention to each of these levers in the design of human systems will improve the likelihood that a cooperative dynamic will arise and self-stabilize.

Communication

Communication plays a critical role in fostering cooperation. In dozens of experiments, allowing participants to communicate with one another predictably and reliably has led to higher levels of cooperation (Sally 1995). In observational studies too, stabilizing and routinizing communication appears to be a crucial part of the new managerial processes.¹ Focusing on cooperation in the study of human relations is anchored in the tradition of dialogic theories of the self: the idea that the self comes to know its interests, desires, and meaning through communication with others, rather than through solipsistic or egocentric engagement with the self alone. Communication is therefore both a dynamic in its own right, through which people come to see their own goals, preferences, and policies in conversation with others with whom they interact, and a mechanism for achieving the cooperation dynamic in that good communication facilitates most of the other design levers. The effect of communication is a very robust finding in these literatures, and an obvious target for design interventions. It has a large effect in experimental work, and its routinization is one of the

core design principles of the organizational shift to collaborative models. In regulatory policy this means, most radically, that the term *communications* in a government agency should not be a euphemism for *propaganda* or *marketing*, but must instead stand for an genuine effort to engage stakeholders and, equally, all citizens, in a conversation about what needs to be done, how, and why.

Situational Framing

We cannot help but think of relations within frames of reference, and these frames in turn shape the remainder of any decision-making process. In sociology, Erving Goffman called this aspect of social interactions “frame analysis” (Goffman 1974). In psychology, it is often called “situational construal” or simply “framing.” The baseline phenomenon is the same: we cannot avoid interpreting a situation in which we find ourselves in social and cultural terms. By virtue of the act of interpretation we already at least partly determine the nature of the interaction and our likely behavior in it. This aspect of the interaction is like a lens through which we observe reality; there is no unmediated mechanism giving access to any situation.

One particularly evocative experiment studied whether framing a task by telling the subjects they were playing “the community game” as opposed to telling them that they were playing “the Wall Street game” would make a difference (Lieberman *et al* 2004). What the study found was that, with identical payoff structures, when subjects were told they were playing “the community game,” about 70 percent opened by cooperating and sustained cooperation for the duration of the experiment, while when subjects were told that they were playing “the Wall Street game,” 33 percent opened cooperatively, and the rest “defected” (abandoned cooperative behavior temporarily) and continued to defect throughout the game. The frame in this case may have defined for the test subjects “the right thing to do,” or it may have altered their predictions about what the other subjects would do, so as to make cooperation or defection appear to be a better strategy. In any case, the frame clearly had a real effect on behavior of otherwise similar populations encountering otherwise identical payoffs.

Expanded Utility: Empathy and Solidarity

One of the most important ways in which we deviate from pursuing pure self-interest is by caring about others, and caring about the welfare of groups when membership in those groups constitutes at least part of our identity. These emotions are, respectively, empathy and solidarity. One clear experimental finding is that a process of “humanization”—using mechanisms to assure that participants know and recognize the humanity of their counterparts—improves the number of cooperators and the degree of “generosity” they are willing to show others (Bohnet

and Frey 1999). Neurological studies support the proposition that agents' brains respond differently to cooperation with humans than to "cooperation" (that is, playing strategies that in game theory count as cooperative) with computers (Rilling *et al* 2002, 2004).

But generally treating other people as human beings worthy of our concern is only one of two signals we give that we consider another worthy of our cooperation or, at least, consider them highly likely to reciprocate. The other major signal we offer to trigger recognition in others that we consider them close to ourselves is the expression of group solidarity. There has been substantial research in social psychology that supports the finding that people increase the degree to which they cooperate with strangers whom they perceive to be part of even very weakly defined solidarity groups. Experimental subjects have long shown greater generosity to and cooperation with others who merely claimed to share their preferences, for example, for paintings by Klee over paintings by Kandinsky (Tajfel and Turner 1979; Yamagishi 1999). Several researchers continue to determine just how minimal the feeling of solidarity must be to trigger cooperation, and to what degree this feeling functions mainly as a heuristic for reciprocity as opposed to an essential constituent of identity with a group (Yamagishi and Mifune 2008).

The role of symbolically marked groups in fostering cooperation is an important field of study in human evolutionary biology, especially in anthropological research on the co-evolution of genes and culture (Boyd 1986, 2004). Moreover, the importance of "affiliation-based trust," in organizational sociology (Zucker 1986) is consistent with the key role of homophily in the formation of social networks (McPherson *et al* 2001), and similar concepts play significant roles in organizational psychology (Haslam 2004). The basic intuition is that either (a) the more people have a sense of being part of a team, or a clan, the more they are willing to sacrifice their own good for the group; or (b) the clearer the "groupness" of the group is to all its members, the more likely that cooperative action by any member will be reciprocated.

Both empathy and solidarity, and their encouragement through face-to-face meetings or detailed descriptions of the background of participants together constitute, then, another important mechanism for the design of human systems. The modern nation-state has a particularly powerful but also ambiguous relationship to group solidarity. At one level, it represents perhaps the most powerful instance of invented solidarity in history, and has proven capable of leading people to great sacrifice. On the other hand, it has also provided the excuse for some of humanity's worst atrocities. As government in particular aims to harness cooperative dynamics, its agencies must be extremely careful in how they deploy appeals to solidarity.

Normativity: Fairness, Rights, and Norm Compliance

In his now classic article “Rational Fools,” Amartya Sen emphasizes the importance of what he called “commitment” to human motivation, and the failure of economics, by and large, to account for the possibility that people act out of commitment (Sen 1977). Commitment should cover at least two distinct concerns: what is fair and what is right. And, indeed, a consistent finding of the experimental literature is that fairness is endogenous to the cooperative dynamic. Experimental mechanisms whose designs are based on a selfish-rational-actor model put fairness of outcomes aside, focusing on whether individuals are made better or worse off by a given interaction as a way of predicting their behavior. Likewise, fairness is usually separated in policy analysis from efficiency, and left to be addressed after the desired level of activity has been induced through egocentrically defined incentives. A consistent finding of the experimental literature is that this approach fails to consider that people care about the fair distribution of outcomes, the perceived fairness of the intentions of others, and the fairness of the overall process (Fehr and Schmidt 2003). There is mounting evidence in cognitive psychology and neuroscience supporting the contention that we have both emotional and subconscious cognitive capacity to do what we understand to be moral (Hauser 2006). Government policy that aims to harness cooperative social dynamics must include, as an integral part of the legislative process, public debate about why any new policies are right and fair, and lawmakers must understand that only regulations that those affected believe to be right and fair will be successful.

In addition to seeking to do what is right and fair, we also tend to do what we regard as normal; that is, we conform to social norms (Ellickson 1991; McAdams 1997). Much research into the phenomenon of social conformity has dealt with long-standing, usually tightly knit communities that rely upon many of the “design levers” I describe here. When thinking of design for such recently invented systems as a collaborative wiki, a musician’s Web site that seeks voluntary donations, or a process aimed at engaging citizens in supporting a public good (such as recycling or making their homes more energy efficient), social norms must play a different role. At a minimum, they refer not to long-standing internalized norms, but to instances of more or less clearly specified behavioral expectations about what counts as “cooperative” in a given system. Once participants know what counts as cooperation and what counts as failure to cooperate, they can adjust their own actions, as well as judge the actions of others. These expectations function as Schelling coordination norms—arbitrary coordination points that allow people to coordinate their actions without excessive negotiation, such as driving on the left or right, or meeting by the clock in Grand Central Station. Though they lack substantive content, these norms

provide focal points for coordinating behavior. Beyond that, they can be explicitly stated expectations about behavior, like those that anchored Wikipedia in its early days and made it unique among cooperative models in being purely norms based. There is evidence that norms that are self-consciously chosen by a group enjoy high adherence with minimal enforcement requirements (Ostrom *et al* 1994). Where these norms evoke background norms that are already culturally ingrained, they may enjoy the status of those already internalized norms, or the norms may themselves be the object of enforcement through another design lever, punishment.

Trust and Authenticity

Trust is the subject of its own immense literature, and the term has been used in many different ways. Often, it has been used to characterize the success of a system that removes the possibility of defection or human error. When used thus, *trust* is not a design lever at all, but rather a description of the outcome of a system, which signifies confidence in its performance. Trust as a design lever should be viewed as a belief that people in a given relationship have about how others in the system will behave when those others may choose to act in ways harmful or helpful to those who trust them (in the absence of genuine choice, the concept of trust has no meaning). It is precisely because trust here refers to an empirical belief people have about the state of the social system within which they act that we treat it as distinct from normativity. *Trust* as I use the term here is not the same as *trustworthiness*. Rather, a system of trust is a system in which people can reasonably hold the belief that some substantial number of others will not take advantage of them whenever they can. A system designed to foster trust in this sense will usually be improved by breaking down cooperative actions into observable steps, so that participants can reduce their vulnerability to one another while observing the proclivities of others to cooperate or defect. A requirement for the creation of trust is that the person constructing the cooperative system act authentically and be seen as acting authentically. Empty promises of community and cooperation may trick others for a short time, but not over the long term. Thus, for example, a government agency that aims to harness cooperation among the citizens with whom it interacts will need to make its commitments to a cooperative dynamic credible, behave in ways that exhibit trustworthiness, and express by its actions trust in citizens.

Transparency and Reputation

Another important design element, the transparency of a system, bears powerfully on the issues of both trust and punishment. Critically, many of the other

design features I have discussed depend on participants' knowing who has done what, to and with whom, to what effect, by which mechanism. Recognition of this dependence lies behind the argument that biologists Nowak and Sigmund make about the evolutionary impact of moral accounting (though they do not call it that): such accounting, they suggest, was necessary to sustain indirect reciprocity among our forbears, which in turn may have been the driving force behind the evolution of human intelligence (Nowak and Sigmund 2005). Whether or not they are correct about human evolution, studies in experimental economics typically show that reputation-rich games lead to cooperation more quickly and robustly than anonymous games (Fehr and Gächter 2000). Similarly, reputation systems play a significant role in social-software platforms, ranging from commercial systems such as eBay and Amazon (Resnick and Zeckhauser 2002) to the wide range of commons-based peer-production projects that rely upon the Internet's resources for observing online behavior, building enduring reputations, and influencing opinion.² For a government agency, this implies that its internal deliberations and its interactions with interested parties must all be made much more transparent than they typically have been.

Autonomy/Efficacy

There is a significant psychological literature suggesting that people need a personal sense of competence or efficacy in their actions, and pursue activities that satisfy that need (Ryan and Deci 2000). For example, among Toyota's most important reforms of the company's production system was decreasing the number of process engineers and according greater autonomy to teams of employees on the production line (Adler *et al* 1999). The need for autonomy and personal efficacy moreover plays an important role in limiting the efficacy of reward and punishment as complementary, as opposed to competing, means for assuring cooperation. People must not only be assured of their personal autonomy, but the value of their contributions must be recognized before their peers; otherwise, people's incentive to contribute as much as they can is reduced.

Fostering a successful culture of personal autonomy as Toyota has done in the context of government regulation intended to standardize behavior across different contexts is far from trivial. It would require, particularly at the federal level, determining regulatory contexts in which regulators may set relatively broad targets for performance and optimal behavior, and certain excluded categories of intervention, while allowing groups of citizens, (initially municipalities, but in the longer run civil-society organizations and organizations that rely on both public and private funding), to exhibit their competence and to receive authority and funding to achieve a given government goal in a particular context.

Calculation: Punishment, Reward, Crowding Out, and Cost

The design levers discussed above all operate at the level of intrinsic motivations. That is, they all work to induce participants to want to cooperate for reasons that are internal to their own psychological and social needs and desires, rather than in response to external rewards or constraints. However, because both the observational and experimental literature suggest that people vary immensely in their motivations, and that many will inevitably be selfish, stable cooperation systems require some element of external compulsion to keep those who are not driven to cooperate by intrinsic motivations in line. Otherwise cooperation tends to unravel, since the presence of selfish actors may undermine efficacy, fairness, solidarity, or any of the other mechanisms that sustain cooperation even in the presence of defectors.

Mechanisms for disciplining and punishing defectors are therefore important in the design of cooperation platforms. The experimental literature finds that (a) with the right design, reciprocators will usually be willing to incur the cost of punishing defectors in order to keep them in line, without intervention from an external body, such as the state or management, but (b) that punishment can backfire if it is not properly designed (Bowles and Gintis 2002; Fehr and Gächter 2002); Fehr and Rockenbach 2002; Falk 2005). It is important to understand that introducing the idea of punishment does not collapse our analysis back to selfish rationality. Punishment is neither necessary to explain cooperation (we see cooperation without it, most importantly in the “second-order public-goods problem” created by the need to impose costly punishment on defectors) nor sufficient (we see instances in which punishment reduces cooperation, probably through “crowding out”). Indeed, punishment can impose such great costs on groups that it ceases to be worthwhile (Dreber *et al* 2008). Moreover, the degree to which its effects are beneficial or detrimental varies among cultures (Hermann *et al* 2008). Yet punishment is one design lever available to systems designers for improving compliance by selfish actors with the cooperative behavior of the other agents in the system. While punishment has been studied much more extensively, reward systems have a similar structure—participants pay a cost to keep others, who are more self-interested, in line with the common good. In an analytical sense, rewards are merely negative punishments, but rewards have the added benefit of not triggering spirals of negative retaliation.

Considering the ambiguous effects of punishment brings to the fore one more design constraint that has particular importance when government is involved: the phenomenon of crowding out. Crowding out can occur within or among systems. *Intrasystem* crowding out refers to situations in which the use of one

design lever would reduce the efficacy of another. For example, the introduction of punishment may, under certain circumstances, crowd out trust, and thereby undermine, rather than improve, cooperation (Yamagishi 1986). From the psychological literature, we know that rewards too can trigger the crowding-out effect, but not as powerfully, and that they may be susceptible to framing that will induce people to treat them as noncontrolling (Ryan and Deci 2000). *Intersystem* crowding out may occur when the designer tries to mix and match elements from cooperative systems with elements from other systems, such as market mechanisms. There is a large literature on crowding out caused by the introduction of money into otherwise cooperation-based interactions (Frey and Jegen 2008; Bowles and Gintis 2001). For regulatory systems, crowding out presents a particularly vexing problem, because the body that aims to introduce and foster cooperation is of necessity a body that possesses enormous power, and is typically seen as able to bring coercive power to bear in a relationship. State-based cooperation systems always necessarily involve the risk of some intersystem crowding out; if the state is involved, then citizens may feel that they do not need to contribute, because the state “will take care of it all.” On the other hand, the state can structure its role in the cooperative systems it builds as a potentially neutral third-party referee that may moderate the negative effects of punishment where it is needed. This does indeed seem to be precisely the role that impartial justice is supposed to introduce into societies that otherwise might remain plagued by vendetta-like forms of private mutual monitoring, discipline, and punishment.

In addition to punishment and reward, which operate primarily on individuals who otherwise might not cooperate because of their intrinsic drives, it is important to remember that the claims of prosocial motivations do not exclude considerations of personal costs and benefits. The essential conclusion of the literature on cooperation is not that large numbers of us are altruists regardless of cost. Rather, it is that large numbers of us have prosocial motivations—regard for others because of our empathy and solidarity, or regard for the normative implications of what we do, in addition to our other cost-benefit considerations. It is not surprising, therefore, that the cost of cooperation affects its levels and the number of people who cooperate. In experiments, subjects will cooperate more when the cost of doing so is lower, such as when the opportunity cost of cooperating in a prisoner’s dilemma is lower because of the payoff structure (Camerer and Fehr 2004). In real life we see peer production online improved when a task has been chunked into sufficiently small modules to make the cost of individual contribution correspondingly small (Benkler 2002).

Social Dynamics

An increasing amount of work being done today on social networks' effects on behavior. It turns out that, for example, our own obesity is affected by whether our friends and relatives have recently become obese (Christakis and Fowler 2007). While the mechanism is not entirely clear, it appears that there is at least some role played by benchmarking and imitation: that is, we judge our own behavior and outcomes by comparing ourselves to others in our social neighborhood (Hanaki *et al* 2007). Allowing participants to observe each other (transparency) and to form and break attachments selectively with people who are more or less cooperative, so as to increase the number of interactions they have with cooperators as opposed to defectors, is therefore also a valuable feature in human systems design: in this way, groups of cooperators may stabilize and provide mutual support.

One important aspect of social dynamics is leadership. An emphasis on leadership does not emerge from experimental research, which does not examine the phenomenon, but from organizational sociology, where it is a consistent feature (Maccoby and Heckscher 2007). Leadership is emphasized in the study of open-source software (Weber 2005), and repeatedly crops up in field studies of online cooperation as well. It is important to recognize, however, that *leadership* does not necessarily imply *hierarchy*. Rather, what we see in observational research is that people contribute to a given system at many levels. Thus systems need to be designed to accommodate and recognize people's varying patterns of contribution, especially by offering them a voice in the collective governance of the enterprise, or through symbolic means of expressing honor and respect. Moreover, for at least some people, it is precisely the desire to seek positions of power, of leadership, that drives generous, prosocial behavior. The role of gift giving as a modality of asserting dominance, so-called agonistic giving, is widely recognized in the anthropology of the gift,³ and in fundraising situations in which public exhibition of gifts is a form of asserting status; we also see it in some, but by no means all, online cooperation sites.

The following, then, is a summary of the discussion above as a list of design levers, or design considerations:

- Communication
- Situational framing
- Expanding the utility function:
 - Empathy
 - Solidarity

- Normativity
 - Fairness
 - Moral commitment
 - Norm compliance or conformism
- Trust
 - Trust and authenticity
 - Transparency and reputation
- Autonomy/efficacy
- Calculation
 - Punishment and reward
 - Crowding out
 - Cost
- Social dynamics
 - Social network effects
 - Leadership and asymmetry

No list of fifteen potential design levers can hope to provide the deterministic outcomes implied by simpler models of human motivation and system intervention. For regulators who seek the comfort of a *If you do X, you will increase rewards through action I, and therefore increase the likelihood of outcome O* type of analysis, the work on cooperation will appear too early in its development to provide guidance. However, regulators who understand that human motivation and social, psychological, and cultural interactions are extremely complex phenomena, which cannot be reduced to a simple *If you do X then Y will follow* without enormous loss of information, may be more patient as we try to work out how the design levers nonetheless provide substantial advantages over a mere recognition of the complexity of human action.

Below I will suggest at least an initial set of principles of regulation that would take advantage of the insights of the literature on cooperation, even if we cannot yet provide a deterministic analysis of the desired structure of cooperation-eliciting regulations.

Principles

The design levers provide a framework for thinking about how most effectively to structure cooperative models of solving problems that are the subject of regulation or characterize the regulatory process. To conclude this chapter, I offer five principles of regulation through which regulators may implement the lessons of cooperative dynamics.

1. Develop a capacity for relying on social mechanisms equivalent to the focus in the past quarter century on market mechanisms

One of the dogmas in regulation from the past quarter century has been that government agencies should seek market-based mechanisms to achieve goals previously fulfilled by government. The cooperation literature argues instead for a new focus on cooperative solutions to the problem of providing public goods. This does not mean that cooperative mechanisms will always be best; it does suggest that they are available and sustainable, and that they have different benefits and potential pitfalls from those of either market mechanisms or government-provided services, and should be considered in any given context for their feasibility and relative desirability.

Take for example the peer-to-patent system. Developed by Beth Noveck in collaboration with IBM and the Patents and Trademark Office (PTO), this program attempts to resolve the relative lack of expertise in the patent office in software by opening up software patent applications to a process of peer review. The model relies upon the culture of substantial contribution to the public good that has grown up among software developers, owing to the rise of free and open-source software. Noveck's system invites patent applicants to submit their applications to an expedited review process through which the members of the open community of software developers can look at the application, research potential reasons to preclude patentability, and submit their observations to the patent examiner appointed by the PTO. The point here is not to displace the PTO, but to harness the knowledge of the community and to take advantage of the general willingness to contribute among developers, who are motivated to some degree, to be sure, by competitiveness, but who also are eager to make a difference in the field they care about. Contributors are allowed to submit potential "prior art"—that is, earlier publications or patents that would make the current claim insufficiently novel, or too obvious, to be granted a patent. Others may then annotate and comment on whether the proposed prior art in fact provides good grounds for objecting to the patent, and vote on which are the most likely instances of prior art to effectively limit or entirely invalidate the patent application. The examiner then reviews the top proposed pieces of prior art. My point is not to assert that this system is the best example of harnessing all the design levers to optimize levels of cooperation. Rather, the PTO site provides an example of isolating a particularly sticky regulatory problem and restructuring it in a way that takes advantage of volunteers from the most relevantly affected community to work, together with the government, on implementing a solution.

The more general point is that government agencies must begin to consider engaging citizens in cooperative systems as one potential way to improve aspects of the regulatory process, and also to improve the delivery of services. Today, government lags behind in its capacity to harness social volunteerism to achieve public goods. Compared to the free and open-source software community, which is harnessing volunteer efforts to produce some of the world's most important software, or to civil-society organizations, such as the Sunlight Foundation, which is harnessing volunteer efforts to sift through government accountability-related data to analyze it for evidence of corruption of public officials, government processes have not focused on using their professional staff and public funding to support volunteer efforts that could improve their processes. Agencies need to develop the capabilities, and to offer funding models, that will facilitate the provision of public goods—from skills training in schools and adult educational facilities to the analysis of government corruption—by partnering with market and nonmarket organizations that specialize in harnessing and structuring volunteer efforts, both online and offline.

2. Include in the evaluation of proposed market or command mechanisms expected impact on social provisioning that exists, or that could be constructed as an alternative

It is an unusual characteristic of American society, compared to other major industrial social democracies, that we continue to rely to a much extent greater than they do on volunteer organizations, both religious and secular, to provide social services. In the past, when the market and the state were considered the only possible providers of public goods, our reliance on volunteerism was criticized by some as constituting a partial abdication by the modern social-democratic state of its responsibilities. The contemporary emergence of social production in the networked economy, and the increasing understanding of the possible benefits of sustained social action to provision public goods, places the United States, perhaps paradoxically, at a surprising advantage in making the transition to a new system that would permit greater leeway for social institutions to play a substantial role in serving governmental agencies' purposes.

There is a substantial literature suggesting that government services can sometimes crowd out private volunteerism. The mechanism is not as yet entirely clear: it may be that when a government agency takes over a particular social service, people cease to see providing it as a shared social responsibility—it becomes something we “get from the government”; it may be that there is a loss of the “social-capital value” of contributing, once the contribution is no longer strictly necessary, and volunteering ceases to enhance people's social status; or it

may be that when a government agency takes over some social service, volunteers cease to feel motivated, independent, and needed. Whatever the reasons, government action can sometimes undermine social provisioning. Government actions aimed at providing public goods should therefore take into consideration the existing system of providing services, and assess: (a) the presence or absence of a well-functioning social provisioning system for the desired good; (b) the likely effect of the proposed action on that social provisioning; and (c) potential adjustments to the proposed policy that would mitigate the negative effects, if any, on social provisioning. Preferably, government should partner with and support existing social providers.

This principle is, of course, not a hard constraint. There may be other considerations, such as the value placed on the availability of public, nonsectarian sources of provisioning a public good where the only source is sectarian—say, special education or soup kitchens. But the broad principle is clear. Just as in the past proposed government regulations were analyzed to anticipate their interactions with existing and potential market mechanisms, and redesigned to minimize any negative regulatory impact and to take advantage of market mechanisms, so too now, with the newly emerging recognition of the role of social action, government must also consider its effects on the social provisioning of desired public goods, and redesign its programs accordingly.

3. Use network technologies to reconstruct government decision processes to enable effective participation by citizens and affected populations on a continuous basis, including implementation

People tend to follow regulations that they choose themselves more willingly and more observantly than rules set down by a remote other. Improving the actual and perceived level of participation in government regulatory processes will encourage citizens to view new regulations as legitimate, and should also make citizens more likely to view regulations as intrinsically binding. That is, when people participate in making the rules, they are more likely to think that obeying them is the right thing to do, and not something they do merely to avoid censure. A substantial literature already exists on opening up regulatory processes to citizen participation, including sensitive studies of stakeholder participation, on which regulators can now draw. Efforts to reach beyond stakeholders to the citizenry at large have, in the past, been largely treated under the rubric of “eGovernment” and have relied on relatively passive Web 1.0 technologies that allow people with Internet access to read and add comments to Web pages. Such an approach lacks structure, and offers little potential for aggregation of comments and debate among citizens in ways that could render their contributions more salient and politically meaningful.

The next generation of regulators needs to learn from the work on participatory government generally, and adopt processes that will offer the broadest range of citizens a much greater degree of freedom to participate effectively, by submitting comments and commenting on the comments of others, and by collecting and analyzing citizens' preferences to determine their preferences. The Environmental Protection Agency has begun to take steps in this direction; Change.gov began, during the Obama transition, to experiment with ways of increasing citizen's direct participation, such as its house-party discussions of healthcare reform. Beyond these examples, this approach needs to be understood as a new principle of regulation: government must provide the means, both online and offline, for effective, widespread participation by citizens in the regulatory process, from its inception to its conclusion and subsequent oversight.

4. Implement transparency through highly accessible visual interfaces, with capabilities for social observation and commentary on official behavior

Transparency, and the widespread perception that processes are transparent, is essential to cooperation. In earlier times, transparency depended largely on professional auditors and observers—internal government auditors, such as comptrollers, or external independent actors, such as an independent press. Freedom of Information legislation supported this transparency by requiring government agencies to provide information upon request to individuals and nongovernmental organizations. While these elements of transparency are critical, and need to be preserved, current regulatory systems must embrace more radical transparency and more open data communications processes to harness peer production and online social action to the task of assuring that government always operates under the public gaze, that regulatory processes are not hidden from view, and that all the materials involved are available for everyone to examine and analyze. The data that government agencies collect, and records of what they do and how they spend their allocated funds, must be collected and kept up to date; government databases need to be made available for public search and comment; and data needs to be made available in standard formats that will allow civil-society organizations and online unstructured collaborative groups to develop their own open interfaces to it. Only in this way can we harness innovation throughout the network to design ever better platforms for monitoring regulators and holding them accountable.

5. Assess fairness as an integral part of effectiveness

A core finding of the study of cooperation is that people care deeply about fairness, and that their perceptions of the fairness of their interactions have a direct,

significant impact on their motivation to participate productively. Fairness should not be mere window dressing for regulations that have efficiency as their main goal. Rather, fairness is essential to stimulate effective, productive engagement by the citizenry in any regulatory process. Fairness is not an ideal, universal state; it is a culturally embedded idea that people in a given society, at a given time, hold about the rewards for their common actions, as well as the intentions and processes involved in determining the distribution of those rewards. Government agencies should develop explicit metrics to assess the impact of their policies. Initially, the focus should be on the distribution of wealth and welfare. The distribution of these social goods should be fully described and measured against a culturally meaningful scale of fair distribution—individually equal (one common measure of fairness in a market-integrated society), progressively redistributive, proportionate to contributions, or based on an articulated theory of desert. These outcomes should be included in the formal analysis of an agency's proposed regulations, and published as part of the regulatory process. Alongside current practices of producing cost-benefit analyses and environmental impact statements, agencies should develop and publicize outcomes-fairness analyses of proposed regulations.

Conclusion

The 1970s, '80s and '90s saw the exquisite refinement of rational-actor theory and used it to justify an ever increasing emphasis on market-based models of regulation and mechanism design. In regulatory practice, this approach became dominant in the Reagan and Thatcher period, and extended its reach with the increasing scope of a European Commission empowered to focus on building an integrated market in what was then called the European Economic Community. The commission therefore emphasized competition and efficiency, and supported institutions of the global trade and monetary system that relied on market-based mechanisms as a matter of principle. Throughout this period a countervailing trend among some scholars advocated more participatory structures, but the global financial crisis of 2008 has brought into the sharpest relief the severe limitations of market-based regulatory approaches. Or at least, the crisis has undermined the nearly absolute dominance that market fundamentalism has had in determining the regulation—or absence of regulation—in financial services.

At the same time, an increasing body of work in the social sciences and particularly in the management and sociology of organizations, as well as on the study of online social practices, has provided substantial new evidence that refines our understanding of the conditions that can sustain forms of social cooperation that are not based on either market mechanisms or command and control. While it is too early to embrace cooperation and collaborative practices

as *the* solution to all our regulatory problems, we certainly have enough information and insight to begin to implement regulatory practices that will likely enhance cooperation and the social provision of public goods.

The principles proposed in this essay are intended to orient regulators toward thinking about whether and how they can harness social cooperation to achieve some of the goals that, in the past, they have sought through command and control or through market mechanisms; how their expected actions will affect existing social cooperation, and how core drivers of cooperation in society can be harnessed by focusing on improved participation in, and transparency and fairness of, regulatory processes and outcomes.

Notes

- 1 See, e.g., Anabel Quan-Haase and Barry Wellman, "Hyperconnected Net Work: Computer Mediated Community in a High-Tech Organization," in Heckscher and Adler 2007 (describing communications flows in collaborative segments of a firm).
- 2 See, e.g., "Wikipedia:Barnstars" in Wikipedia: <http://en.wikipedia.org/wiki/Wikipedia:Barnstars>; or the team-competition contributions in Yochai Benkler, "Sharing Nicely: On Shareable Goods and the Emergence of Sharing as a Modality of Economic Production," *The Yale Law Journal* 114 (2004): 273.
- 3 See, e.g., Maurice Godelier, *The Enigma of the Gift* (Chicago: University of Chicago Press, 1999).

References

- Adler, Paul S., Barbara Goldoftas, and David Levine. "Flexibility Versus Efficiency? A Case Study of Model Changeovers in the Toyota Production System." *Organization Science* 10, no. 1 (1999): 43–68.
- Benkler, Yochai. "Coase's Penguin, or, Linux and the Nature of the Firm." *The Yale Law Journal* 112, no. 3 (2002): 369–446.
- Bohnet, Iris, and Bruno Frey. "The Sound of Silence in Prisoner's Dilemma and Dictator Games." *Journal of Economic Behavior & Organization* 38, no. 1 (1999): 43–57.
- Bowles, Samuel. "Policies Designed for Self-Interested Citizens May Undermine 'The Moral Sentiments': Evidence from Economic Experiments." *Science* 320 (2008): 1605–09.
- Bowles, Samuel, and Herbert Gintis. "Homo Reciprocans: Altruistic Punishment of Free Riders." *Nature* 415 (2002): 125.
- Boyd, Robert, and Peter J. Richerson. *Culture and the Evolutionary Process*. Chicago: Chicago University Press, 1986.

- . *Not by Genes Alone: How Culture Transformed Human Evolution*. Chicago: Chicago University Press, 2004.
- Camerer, Colin F., and Ernst Fehr. “Measuring Social Norms and Preferences Using Experimental Games: A Guide for Social Scientists.” In Joseph Henrich, *et al*, eds. *Foundations of Human Sociality: Economic Experiments and Ethnographic Evidence from Fifteen Small-Scale Societies*. New York: Oxford University Press, 2004.
- Christakis, Nicholas A., and James H. Fowler. “The Spread of Obesity in a Large Social Network over 32 Years.” *New England Journal of Medicine* 357 (2007): 37–79.
- Dreber, Anna, *et al*, “Winners Don’t Punish.” *Nature* 452 (2008): 348–51.
- Ellickson, Robert. *Order Without Law: How Neighbors Settle Disputes*. Cambridge: Harvard University Press, 1991.
- Falk, Armin, *et al*. “Driving Forces Behind Informal Sanctions.” *Econometrica* 73 (2005): 2017.
- Fehr, Ernst, and Simon Gächter. “Cooperation and Punishment in Public Goods Experiments.” *American Economic Review* 90, no. 4 (2000): 980–94.
- . “Altruistic Punishment in Humans.” *Nature* 415 (2002): 137.
- Fehr, Ernst, and Bettina Rockenbach. “Detrimental Effects of Sanctions on Human Altruism,” *Nature* 422 (2002): 137.
- Fehr, Ernst, and Klaus M. Schmidt. “Theories of Fairness and Reciprocity—Evidence and Economic Applications. Paper prepared for the invited session on Behavioral Economics of the 8th World Congress of the Econometric Society in Seattle, 2000. In Dewatripont, M., *et al*, eds. *Advances in Economics and Econometrics*, Eighth World Congress of the Econometric Society, vol. 1. Cambridge: Cambridge University Press, 2003.
- Frey, Bruno S., and Reto Jegen. “Motivation Crowding Theory: A Survey of Empirical Evidence.” *Journal of Economic Surveys* 15, no. 5 (2001): 589.
- Goffman, Erving. *Frame Analysis: An Essay on the Organization of Experience*. New York: Harper & Row, 1974.
- Hanaki Nobuyuki, Alexander Peterhansl, Peter S. Dodds, and Duncan J. Watts. “Cooperation in Evolving Social Networks.” *Management Science* 53 (2007): 1036–50.
- Haslam, S. Alex. *Psychology in Organizations: The Social Identity Approach*. Thousand Oaks, CA: Sage Publications, 2004.
- Hauser, Marc. *Moral Minds: How Nature Designed Our Universal Sense of Right and Wrong*. New York: Ecco, 2006.

- Heckscher, Charles, and Paul Adler, eds. *The Firm as a Collaborative Community: Reconstructing Trust in the Knowledge Economy*. New York: Oxford University Press, 2007.
- Herrman, Benedikt, Christian Thonni, and Simon Gächter. "Antisocial Punishment Across Societies." *Science* (2009) 316: 1362.
- Lieberman, Varda, Steven M. Samuels, and Lee Ross. "The Name of the Game: Predictive Power of Reputations versus Situational Labels in Determining Prisoner's Dilemma Game Moves." *Personality and Social Psychology Bulletin* 30, no. 9 (2004): 1175–85.
- McAdams, Richard. "The Origin, Development, and Regulation of Norms." *Michigan Law Review* 96, no. 338 (1997): 338–433.
- McPherson, Miller, Lynn Smith-Lovin, and James Cook. "Birds of a Feather: Homophily in Social Networks." *Annual Review of Sociology* 27 (2001): 415–44.
- Maccoby, Michael, and Charles Heckscher. "A Note on Leadership for Collaborative Community." In Heckscher 2007 (see above).
- Krueger, Joachim I., and Theresa E. DiDonato. "Social Categorization and the Perception of Groups and Group Differences." *Compass: Social and Personal Psychology* 2 (forthcoming 2008): 1–18.
- Nowak, Martin, and Karl Sigmund. "Evolution of Indirect Reciprocity." *Nature* 437 (2005): 1291–98.
- Ostrom, Elinor, Roy Gardner, and James Walker. *Rules, Games, and Common Pool Resources*. Ann Arbor: University of Michigan Press, 1994.
- Resnick, Paul, and Richard Zeckhauser. "Trust Among Strangers in Internet Transactions: Empirical Analysis of eBay's Reputation System." In *The Economics of the Internet and E-Commerce*. Michael R. Baye, ed. Advances in Applied Microeconomics, vol. 11. Amsterdam: Elsevier Science, 2002.
- Rilling, James K., et al. "A Neural Basis for Social Cooperation." *Neuron* 35, no. 2 (2002): 395–405.
- . "Opposing Bold Responses to Reciprocated and Unreciprocated Altruism in Putative Reward Pathways." *Neuroreport* 15, no. 6 (2004): 2539–543.
- Ryan, Richard M., and Edward L. Deci. "Self-determination Theory and the Facilitation of Intrinsic Motivation, Social Development, and Well-Being." *American Psychologist* 55 (2000): 68–78.
- Sally, David. "Conversation and Cooperation in Social Dilemmas." *Rationality and Society* 7, no. 1 (1995): 57–92.
- Sen, Amartya. "Rational Fools: A Critique of the Behavioral Foundations of Economic Theory." *Philosophy and Public Affairs* 6, no. 4 (1977): 317–44.

- Tajfel, Henri, and Turner, John. C. "An Integrative Theory of Intergroup Conflict." In W. G. Austin and S. Worchel, eds. *The Social Psychology of Intergroup Relations*. Monterey, CA: Brooks-Cole, 1979.
- Yamagishi Toshio. "The Provision of a Sanctioning System as a Public Good." *Journal of Personality Social Psychology* 51 (1986): 110–16.
- . "Generalized Reciprocity and Culture of Collectivism." *Organizational Science* 33, no. 1 (1999): 24–34. (In Japanese.)
- Yamagishi Toshio and Mifune Nobuhiro. "Does Shared Group Membership Promote Altruism? Fear, Greed, and Reputation." *Rationality and Society* 20 (2008): 5–30.
- Weber, Steve. *The Success of Open Source*. Cambridge: Harvard University Press, 2005.
- Zucker, Lynn G. "The Production of Trust: Institutional Sources of Economic Structure, 1840–1920." In Barry M. Staw and L. L. Cummings, eds. *Research in Organizational Behavior*, vol. 8. Greenwich, CT: JAI Press, 1986.

This work is distributed under a Creative Commons Attribution–Noncommercial–No Derivative Works 3.0 license. Readers are free to share, copy, distribute, and transmit the work under the following conditions: All excerpts must be attributed to: Moss, David, and John Cisternino, eds. *New Perspectives on Regulation*. Cambridge, MA; The Tobin Project, 2009. The authors and individual chapter titles for all excerpts must also be credited. This work may not be used for commercial purposes, nor may it be altered, transformed, or built upon without the express written consent of the Tobin Project, Inc. For any reuse or distribution, the license terms of this work must always be made clear to others: the license terms are available at <http://creativecommons.org/licenses/by-nc-nd/3.0/us/>.

Copyright © 2009 The Tobin Project, Inc.

All rights reserved. For information address

The Tobin Project, One Mifflin Place, Cambridge, MA 02138.