

Abstract:

The last two decades have seen significant work undermining the selfishness hypothesis commonly used as a simplifying device for analysis based on the rational actor model. The work traverses a wide range of disciplines, from experimental economics and game theory, to anthropology, social psychology, and evolutionary biology. At the same time, organizational sociology and management science have been chronicling and systematizing observations about the development of collaborative models of production in some of the most innovative, competitive industries in the global economy, while work on the emergence of the networked economy has emphasized the increasing centrality of large scale collaborative models to information and cultural production. These diverse intellectual trends suggest a need to adapt legal and policy design to incorporate the greater diversity of human types and attitudes toward cooperation. The paper synthesizes out of the underlying literature a set of design elements, and suggests how they might be used to systematize our approaches to the design of policy interventions and institutions based on assumptions that differ quite fundamentally from the dominant assumptions of selfishness and solipsism, without requiring an assumption of universal altruism or necessitating a new communitarianism.

Introduction

The marginal productivity of *homo economicus* is declining. The thin view of the rational actor, whose preferences are well ordered and transitive, who readily calculates many steps forward and single-mindedly pursues the optimization of his own welfare, has been a productive model across many fields. It has formed the basis of mainstream economics, much of contemporary political science, as well as law and economics. It has informed business organization and engineering models. But it has always operated under significant pressure. Some criticism was internal to economics,¹ as Jessica Leight discusses in her essay here on the intellectual history of public choice. Mostly, the weight of the other social sciences, literature, critical theory, and philosophy were arrayed against it. The defense of *homo economicus* has usually been some version of Milton Friedman's argument: the model is justified by the quality of the predictions it offers, not the plausibility of its assumptions.²

A large body of empirical work has put *homo economicus* to controlled tests and field studies. It has shown that simplistic heuristic to be less predictive of observable behavior than previously thought. Famous in the legal literature has been observational work on social norms and trust,³ common property regimes,⁴ and, later, experimental behavioral law and economics.⁵ The better known aspect of the experimental work has been divergence from the predictions of rationality itself, as Joseph Stiglitz emphasizes in this volume. Less well known is work that does not take aim at cognitive failures of rationality, but undermines two core simplifications, neither entailed by rationality, that made *homo economicus* user-friendly: (a) individuals are similarly motivated, and (b) they are all selfish. Instead, we find that human beings have diverse motivational-behavioral profiles. In experiments, almost one third indeed behave as predicted by selfish *homo economicus*. But more than half act cooperatively. Many are active reciprocators—respond kindly and cooperatively to cooperating others, and punish, even at a cost to themselves, those who behave uncooperatively. Others cooperate unconditionally, whether because they are true altruists or solidarists, or because they

simply prefer to cooperate and do not measure what others are doing. The overarching finding, however, is clear: In no human society studied under controlled experimental conditions have people on average behaved as predicted by the standard economic model. Given that the assumptions of *homo economicus* are so inconsistent with intuition, experience, and the pervasive cultural practices of inculcating sharing and pro-social behavior in children, the fact that they now have also been seen to be systematically poor predictors of observable human behavior under controlled conditions requires significant attention.

Consider three puzzles for different models of human motivation and interaction.

(a) In firms, some experienced workers have acquired significant know-how. The firm wishes these agents to transfer as much of their know-how as possible to newer employees, to make them more productive. The traditional economic view of human motivation would assume that employees will wish to retain their own value and relative productivity, and so will withhold know-how. That view would therefore recommend that firms create incentives for experienced employees to transfer knowledge by compensating employees for teaching, and then monitoring their performance. It would predict that a relatively unsupervised environment, where employees work in teams without well-monitored interactions and without compensation tied to knowledge transfer, would result in experienced employees keeping as much of their know-how as possible to themselves so as to maximize their nonfungible value to the firm and hence their bargaining power to extract a larger share of the firm's value. Analysis based on a large emerging literature on cooperation, which I will address in this essay, would, to the contrary, predict that at least half of the employees prefer to cooperate. Given the opportunity to engage in more loosely supervised teamwork, many employees would be predicted freely to share information with coworkers through a social system of mutual aid. This view would recommend a radically different organization for the workplace. Empirical and theoretical work suggests that the latter is indeed better as an organizational design to implement knowledge sharing.⁶ Moreover, Toyota's legendary innovations in teamwork, rather than fine-grained incentives pay and monitoring, are usually thought of in organization science precisely in terms of organizing the workplace as a better platform for experimentation and knowledge sharing.⁷

(b) *Homo economicus* predicts that Wikipedia cannot exist. Tens of thousands of volunteers, none paid, acting effectively without crisp lines of authority, would simply be impossible. In a world populated exclusively by self-interested actors, an institutional designer of an online encyclopedia would have focused on ways to monitor and reward contributions, or to identify and sanction deviations. *E pur si muove*. Wikipedia is based on a model of self-selection, communication, human interaction, community norms, and mutual peer-based review and discipline.⁸ Whether it is imperfect, or of lower quality than Britannica, is beside the point. The fact that contemporary debates focus on comparisons to Britannica is the strongest evidence of success. These design characteristics are common in peer-production projects generally, and more closely follow the design characteristics predicted by the cooperation literature than those predicted by selfishness.

(c) Consider two approaches to crime reduction. The first assumes that criminals, as rational actors, are deterred by the size of the penalties discounted by the probability of detection.⁹ This theory calls for tougher penalties and/or more vigorous enforcement efforts. The second approach characterizes high-crime rates as a failure of community—a failure to create a cooperative environment in which most people do not commit crimes, and members actively work together to control, report, and prevent crime when possible. That diagnosis would see high crime rates not in terms of penalties

or probabilities of detection that are too low, but of insufficiently rich cooperation. For proponents of the cooperative perspective, sensible interventions would seek to facilitate a cooperative dynamic in the neighborhood, as exhibited in the dramatic rise and popularity of community policing.¹⁰ It is the foundation of a substantial policy shift toward community policing in many communities in the United States.

The three stories are intended to render intuitive the critical points that will drive my analysis in this essay. People are diversely motivated with regard to cooperation. Policy-makers and organizational entrepreneurs can design institutions and social systems to foster cooperation by shaping social and psychological dynamics, rather than by focusing on individual incentives. The question then becomes, what aspects of the design of an institution or system—be it technical platform, legal rule, business process, or policy intervention—are likely to lead to a stable cooperative social dynamic.

The immediate “policy” or “business” reason to pursue this approach to design is the increasing recognition of the necessity of loosely-coupled systems design. In many different domains, loosely-coupled systems are replacing tightly-bound systems: from Taylorism and Fordism to Toyotism, from the Bell System to the Internet, from IBMs' massive patent portfolio and structured hierarchical model to its new services-based model built around a mixture of proprietary and open source software, from hierarchical command structures in the military to network-centric warfare, and from the Hollywood studio system to do-it-yourself media on the Web.¹¹ As the scale of global interactions increases, as the complexity of operating environments intensifies, and as the rate of economic, social, and cultural change escalates, the large, centralized, monitored, and carefully managed systems that dominated the twentieth century are fraying at the edges. In the face of all these transformations, human creativity, insight, wisdom, and learning capability become core imperatives for all systems design. And as these difficult-to-observe aspects of system performance increase in salience, the capacity of that system fully to characterize, monetize, monitor, and reward all desirable action declines. Instead, what all these systems need is intrinsically motivated human action. The goals of organizational strategies increasingly have turned away from such traditional tasks of organizing action and structuring motivation. The newer objectives call for shaping social and psychological dynamics so that people, acting autonomously in an environment where command is ineffective and pricing inefficient, can assess, experiment, learn, adapt, communicate, and adopt better practices dynamically over time. In short, the operating premises of economic life have become substantially more cooperative than *homo economicus* could have predicted. We now need a modeling framework whose ambition is to become as broadly applicable and flexible, but which will build on the new insights into human cooperation. That is the goal of studying cooperation and human systems design.

A. An emerging literature on cooperation

Deciding which body of literature to use in constructing an approach to design for cooperation is not trivial. From poems, novels, or folk tales, through history, religion, and social theory, to mathematical game theory, we have many ways of talking about foundational questions of human motivations: when do we respect others and care about them? When do we look out for ourselves? Are we by nature generous or selfish? No approach can be fully comprehensive. No approach can be generalizable and usable without being to a significant extent reductionist—taking highly complex problems and reducing them to some manageable number of salient variables, leaving residual uncertainty that is at a minimum comprehensible and sufferable. Using this reducibility to the tractable

as my guide, I focus on six intellectual strains from which to synthesize the new approach to cooperative human systems design. These fields are: (a) experimental economics, with related cooperative game theory; (b) evolutionary biology and a strand of anthropology dealing with gene-culture co-evolution; (c) psychology of motivation and social psychology of solidarity; (d) organizational sociology and management science work focused on cooperative business processes; (e) observational work on successful common property regimes; and (f) the study of online collaboration and social software design. The fields vary in the degree of reductionism they impose and tractability they enable. My hope is that by maintaining all of them within our peripheral vision we can develop a framework both simple enough to be as generalizable and tractable as were mechanism design and law and economics, while providing a richer and more humanly grounded characterization of how human beings are motivated, how they interact with each other, and how they are likely to respond to various system manipulations we can propose—be they in the form of legal-institutional reform, or broader technical, business process, or social interventions.

Contemporary economic studies of cooperation are anchored in experiments.¹² These studies use well-understood games that have predictable theoretical outcomes, and observe people under controlled and manipulated conditions, comparing actual behavior to the behavior predicted by the selfish rational actor model. Most of these are social dilemma games: investigators structure available behaviors and payoffs such that cooperation will lead to higher payoffs for all participants, and lack of cooperation will lead some or all players to have a lower payoff, but may lead to higher payoffs to non-cooperators if others do act cooperatively. The Prisoner's Dilemma is the most famous social dilemma game. Other games more closely approximate contributions to public goods, or elicit the presence or absence of reciprocity under varying conditions. Some games, like the Dictator Game, are designed to detect the presence of “pure altruism” with no gains from cooperation. Social scientists have developed, critiqued, and refined these games for over two decades, observing the behavior of students in many different developed cultures; workers and ethnic groups in various societies;¹³ individuals within several small-scale, relatively isolated, societies,¹⁴ and across industrialized countries with diverse degrees of rule-of-law and social capital indicators.¹⁵ Throughout this period, participants in this field have also developed increasingly sophisticated mathematical models to include other-regarding preferences into the utility function of agents.¹⁶

We have seen a parallel trajectory in the study of cooperation in evolutionary biology. The first rise and fall of sociobiology begins with the 19th century rise of Social Darwinism and eugenics, the emergence of anthropology and Frans Boas's critique, and the resolution of the battle in the revulsion caused by Nazi eugenics and scientific racism.¹⁷ The desire to read the Book of Nature to understand God: to ground our moral self-understandings in our understanding of nature, remained. By the late 1960s to mid-1970s the horrors of the first interaction with biological explanations of human social behavior had subsided, and could be categorized as “bad science.” This created room for a second rise of sociobiology.¹⁸ Though subject of extensive critique,¹⁹ biological inquiry into human sociality and morality has not subsided since. Some of it, particularly evolutionary psychology, followed the heavily criticised adaptationist model, complete with the idea that human moral drives were both instantiated in physically localized modules in the mind and had evolved to stability 50 million years ago.²⁰ Other strands, however, began to focus more heavily on evolutionary dynamics as an alternative approach to the optimality analysis common in economics. Evolutionary stability did not mean optimality; it merely meant feasibility of arising from a random assortment of possible behaviors or proclivities, and relative survivability and stability by comparison to other possible behaviors and proclivities. Work on the evolution of cooperation in particular has changed over time. Its roots in the 1960s and 1970s

focused purely on individual benefits from cooperation, and narrowed the range of possible cooperation to genetic kin²¹ or directly reciprocal exchanges of benefits.²² More recently, both the strict individualism and the need to achieve direct benefits in exchange for “generous” behavior have been superseded. The notion of reciprocal altruism grew, a quarter of a century later, into the concept of indirect reciprocity—that is, agents cooperate in a society when they can process enough information about who did what to whom so as to “reciprocate” indirectly.²³ Indeed, in the last two years the leading scientific journals have published a number of papers suggesting that the need to track cooperation and defection served as the driving force behind the evolution of human intelligence.²⁴ The strict necessity of individualism has been relaxed by the restatement of old views on group selection as a new approach, multi-level selection, which allows for forces to operate as vectors, sometimes in competing directions, at the individual or group level, and thereby create a dynamic that allows for stability even for genuinely self-sacrificial strategies.²⁵ Most significantly from the perspective of contemporary studies of cooperative human systems design, there has been substantial work to mesh evolutionary dynamics with cultural practice. We now see economists, like Bowles and Gintis,²⁶ and anthropologists, like Boyd and Richerson,²⁷ using evolutionary dynamics to explain and explore the stability of cultural practices as salient sources of successful cooperation. Because of its formality, this approach adds theoretical tractability to the analysis of cooperation, applicable at the level of social institutions on historical timeframes, rather than stating biological facts that have to be considered as hard limits on institutional design.

The last of the three lines of literature that seem particularly promising on the tractability side of the cooperation synthesis are lines of work within social psychology that go to human motivation and social solidarity. Much of the experimental work on psychology, certainly work that has made its way into economics and law, has been focused on predictable cognitive failures. The work of Kahneman and Tversky and their followers is particularly influential. But there is substantial work on cognitive failure and its application to law and institutional design beyond the behavioral economics frame.²⁸ While work on designing cooperative human systems will have to consider cognitive failures, it is primarily oriented toward nailing down motivational and social-dynamic effects. To this end, two lines of work within social psychology provide more direct benefits. The largest influence on the framework offered here is work on the difference between intrinsic and extrinsic motivations, and concerns with the risks of crowding out internal reasons for action through imposition of control and explicit incentive systems, in particular the work of Edward Deci and Richard Ryan.²⁹ In much of the observational literature, the emphasis on the importance of intrinsic motivations—reasons and desires for action that come from within—plays a significant role. Deci and Ryan developed over the past two or three decades an approach based on characterizing human needs for competence, autonomy, and relatedness, which offers a convenient framework for mapping design interventions based on the extent to which they go to intrinsic or extrinsic motivations, and offers a reasonably tractable way of considering potential negative effects between the various levers and in the relationship between cooperative and non-cooperative levers. Crowding out, or the non-separability of social preferences from the introduction of explicit extrinsic motivation, poses a systemic challenge to using traditional, incentives-based mechanisms, both private and public, for eliciting desirable behavior.³⁰ Another line of literature within social psychology goes in particular to one important element in cooperation, and that is the construction and functioning of social solidarity.³¹

The remaining three lines of literature are observational. They each take a class of phenomena where cooperation is salient and describe and analyze it in rich detail. In each case, the observations lead to some form of abstraction and systematization. In the approach I outline here, however, the role

of these classes of literature is to ground the more abstract and reductionist approaches in a richer, thicker set of descriptions. It recognizes that practical reason and problem-solving involve application of tacit knowledge as much as explicit knowledge, and tacit knowledge is evoked and applied more through narrative and heuristics than abstraction. Abstraction then allows us to discipline, structure, generalize, and critically examine intuition and experience. The trick, ultimately, of a successful synthesis will be to match up plausible abstractions of the observational work to plausible synthetic generalizations of the experimental and theoretical work to mobilize both tacit and explicit knowledge to the design human systems.

The first of these has emerged within organizational sociology and management science. Growing from the work on post-Fordism, trust, and increasing knowledge intensity in firms, sociologists had observed increasing adoption of networked organization models emerging within firms, and in some cases across firms in supply relationships.³² Globalization and rapid technological change put organizations under increasing pressure to innovate in their processes, adapt to changes, learn about a rapidly changing environment and increasingly complex processes, and implement learning continuously. Under a variety of monikers, such as TQM, team production, quality circles, etc., business processes have emerged that depend heavily on communication, on locating responsibility in the hands of employees, or on the emergence of what Sabel has called new routines for trust-based collaboration, replacing the traditional models of market and hierarchies to govern internal relations within firms and between firms.³³ A second line of literature in the observational vein is the extensive, and in law well-known, political science work on common property regimes, led and epitomized by the work of Elinor Ostrom.³⁴ A third distinct line of work examines online cooperation and peer production generally, and the design of “social software” in particular.³⁵ The creators of social software design platforms with the group in mind – they seek to structure opportunities and constraints that make the group work better, rather than ease the lot of any given individual. Such programs might gradually increase authority and opportunities for individual users as they act over time in ways that show trustworthiness or diligence. The work is largely observational and heuristic at this stage, but because of the explosion of collaboration online as a global practice, there is enormous variety and intellectual effort directed at these design problems, which in turn are also explicitly rendered because they have to be embedded in the software or communicated among far flung communities. This makes online cooperation a particularly rich space for observation and field experimentation.

Several rough regularities emerge from these diverse lines of literature. First, people on average do not behave as predicted by the standard economic model. Second, people appear to have diverse motivational-behavioral profiles. In most experiments about one third behave much as *homo economicus* predicts. About half cooperate.³⁶ Of these cooperators, more than half reciprocate kindness or trust with kindness or trust, and meanness and defection with like meanness and defection, even at material cost to themselves. Others are unconditional cooperators. Not everyone falls neatly, however, into one or the other of these categories. The distribution of behaviors is not smooth, but has modes around what a selfish actor would do and what a cooperator or reciprocator would do. This pattern lends some support to the idea that discrete personality types are deeply ingrained, culturally, biologically, or both, within human societies. Neuroscientific studies support the proposition that different people's brains light up differently in fMRI studies, in similar situations, suggesting a distribution of types whose brain “lights up” differently, and that the differences are consistent with different behavioral patterns.³⁷ On the other hand, cross-country comparisons within industrialized countries show substantial differences in the ways in which people do, or do not, punish defectors in social dilemma games and in how their punishing behavior correlates with their own levels of

contribution. These differences appear to be associated with measures of rule of law in the country, suggesting substantial cultural and learning effects as well.³⁸ Experiments in small-scale societies exhibited even wider variation, which increases the more a society is removed from markets, and to some extent maps on to cultural practices of agonism in giving or the degree of cooperation practiced in hunting and gathering or other day-to-day activities.³⁹ Evolutionary arguments, both genetic and cultural, provide plausible stories about how a population of selfish actors could invade a population of cooperators, while over time a population with a high level of cooperators would drive out groups with too many free riders, stabilizing the average prevalence of each type. In other words, it provides an analytic frame through which to explain the rise of diverse equilibria for any given moment, from cooperative to selfish, in different societies. The stability of types, at least in a known target population in a given timeframe, limits the range of predicted behaviors and responses to design constraints and affordances to a manageable set. Understanding whether that distribution at a given time is itself subject to intervention—that is, whether human beings are perfectible along the dimension of their motivation and proclivity to cooperate—will have a large effect on what the feasible and proper targets for human systems design can be.⁴⁰ But the present literature cannot yet tell us with sufficient confidence whether there is such a thing as an individual proclivity to cooperate independent of the situational condition, whether arising from genetics or education, and if so whether this proclivity is stable over a lifetime or itself susceptible to education and perfection over time.

Whether we ultimately come to believe that differences among human beings in terms of proclivity to cooperate are innate, acquired, or mixed, the experimental and observational work certainly supports the proposition that these behavioral patterns are also situational. That is to say, certain characteristics of the situation or context, of the system of interaction, will more likely lead people to cooperate stably, and others will likely lead to deterioration into behavior more consistent with the Hobbesian view of the state of nature. This structured context should preoccupy architects of institutions or social systems aiming for productive cooperation. The following discussion therefore assumes that human beings are diverse; that the diversity of human types is not infinite, that types are stable in the short term, and that each type can be predicted to have a typical behavior under specified conditions of interaction, which will change the social conditions and hence payoffs to all of them.

To render the broad and diverse kinds of relevant scholarly literature usable for policy-makers and institutional entrepreneurs alike, I have identified an initial set of thirteen considerations, or focal points, that designers of such systems should consider. I call these “design levers,” because they are elements or focal points in the system design intended to, and capable of, affecting the dynamics of the social system to which they apply. Addressing these areas of concern properly increases the likelihood that participants will be motivated to behave cooperatively, by setting the social context so that it facilitates cooperation among cooperators and allows selfish actors to be policed and incentivized in ways that do not undermine the intrinsic motivation of the more cooperative participants. These design levers are intended to be a mid-level abstraction: abstract enough to capture a wide range of experimental and field observations, and concrete enough to be tied to particular types of feasible interventions that could be causally tied to achieving one or another of the design levers.

By “cooperation” in a given interaction or system I mean behavior that contributes to the attainment of goals by others in that system. The behavior can reflect one of a range of attitudinal and intentional states: (a) *altruism*: action aimed to contribute to success of an other, irrespective of success of self; (b) *committed mutualism*: action aimed to contribute to success of an other, consistent with success of self; (c) *collective efficacy*: action aimed toward the success of a common goal that

transcends the agent's specifiable individual success; (d) *heuristic reciprocity*: action aimed at benefiting one's own goal pursuit, guided by an implicit sense that “what goes around comes around” (less sensitive to explicit payoff structures than to social and psychological structures that trigger a judgment that the context is one where reciprocal benefits may arise); (e) *strategic mutualism*: action aimed at attaining one's own goal, which advances the goal of another as a byproduct. This latter is the behavior that game theoretical mechanism design seeks to elicit, and in animal studies would be thought of as byproduct mutualism.⁴¹ To dispel misinterpretations: “cooperation” does not mean “behaving nicely.” It means acting in ways that advance the goals of others, including in contexts where even a selfish and solipsistic person would do so, but most importantly beyond what a selfish and self-centered person would be predicted to do. Gang members are often highly cooperative. Suicide bombers exhibit high degrees of self-sacrifice for collective efficacy. Learning the dynamics of cooperation can help disrupt successful cooperation that we judge normatively harmful no less than allowing us to construct successful cooperation whose outcomes and processes we normatively affirm.

The design considerations, or levers, are not strictly necessary or sufficient to achieve cooperation. Some, particularly those that aim to manage selfish actors, may present tradeoffs with other levers—as we will see with discipline and trust. But they characterize loci and types of intervention, or questions presented for the designer of a system, which would predictably affect the likelihood that participants will cooperate.

B. Design Levers for Cooperative Dynamics

Communication. Communication plays a robust role in the scholarly literature on social cooperation. It has a large effect in experimental work,⁴² and its routinization is one of the core design principles of the organizational shift to collaborative models.⁴³ The salience of communication partly locates the work on cooperation in the tradition of dialogic theories of the self: the self comes to know its interests, desires, and meaning through communication with others, rather than through solipsistic or egocentric reflection. This makes the relationship stand apart as an object of design, rather than making the individual, fully formed before the interaction, the object of monitoring, reward, punishment, and similar manipulation to achieve the desired behavior. It also suggests the possibility of perfectibility: that the distribution of cooperative proclivities is stable and given, and cannot itself be the object of longer-term design.

Factors affecting intrinsic motivation

A significant focus of organizational sociology, a major line of the work in psychology, and a heavy focus of the management studies, is on intrinsic motivation—that is, motivation to act that comes from within, rather than in response to external efforts to affect the behavior through reward and punishment. The definition of cooperation as I use it here emphasizes the degree to which an agent whose actions we are assessing intends to do something beyond merely advance his or her own goals—be they a commitment to advance the success of a partner to a mutual exchange, the success of a group in its collective effort, or the flourishing of an other in a context that suggests pure altruism. The next four factors, or design levers, focus on the degree to which large portions of the population adopt cooperation as an internally desired activity.

Solidarity and Humanization/Empathy. Two important levers that have been experimentally

and observationally shown to improve the degree of cooperation in a group are solidarity and empathy, or humanization. The fact that people are more generous towards, and cooperate more with, others who are perceived to be within even a very minimally-specified group is long established experimentally,⁴⁴ and team-building activities and solidarity constructing rituals are widely reported in observational work.⁴⁵ A rich literature in psychology has worked to define social identity and the relationship between the formation of social identity and the quality of collaboration in a team.⁴⁶ Psychologists tend to view social identity as a condition through which an individual develops a sense of self, a “knowledge that he belongs to certain groups together with some emotional and value significance to him of the group membership.”⁴⁷ Individuals who so define themselves generally work especially hard to create and sustain positive distinctions between their group and other groups.⁴⁸ Empirical work over the last two decades has tied this change in identity or self-perception to organizational citizenship, the willingness to contribute to collective goals and behave cooperatively in collective action settings, and increasing group productivity.⁴⁹ This work is consistent with work in evolutionary biology on the relative success of groups. Boyd and Richerson make the evolved psychological proclivity to identify with large, symbolically-marked groups as calling for cooperation beyond kin the core of their explanation of the successful rise of human societies of the scope and scale we see, in contradistinction to other primates.⁵⁰ Bowles and Gintis, while recognizing the power of group identity, emphasize its double-edge. The identification of “in” the group is usually associated with a characterization of other, and can (as it so often has in history) breed xenophobia, or simply moral indifference.⁵¹ In work I have done with collaborators, we indeed shown that identifying individuals who respond with high-activation to a strong solidarity signal (in our case, knowing that they are in a team of Democrats or Republicans) and putting them in groups of similar individuals allows them to sustain very high levels of cooperation in public goods games over long periods, without need for reputation, communication, or punishment mechanisms. And yet, in a study about behavior of supporters of the competing contenders in the Democratic primary in 2008, we found that this effect had particularly large effects in mobilizing men to be more “generous” to other members of their group during perceived conflict with a distinct outgroup—consistent with the double-edge warning.⁵²

Tribalism is not, however, the only way we have of avoiding solipsism. Empathy toward another human being *qua* human being is another. In several experimental contexts, humanization—mechanisms to assure that participants know and recognize the humanity of their counterpart—improves the number of cooperators and the degree of “generosity” they are willing to show others.⁵³ Simply seeing the face of the human being involved in the experiment, without any change in the game design or possibility of communication or reputation effect, significantly affects the degree of cooperation and generosity that both average and even mostly selfish players exhibit. Adding more personal information, such as hobby or undergraduate major, further improves cooperation and generosity. The distinct effects of interacting with another human being as such, by comparison to a computer, for example, have also been documented in neuroscience studies.⁵⁴

Trust and Fairness. The importance of trust is central to organizational sociology, but is also strongly present in experimental work. “Trust” has attracted its own immense literature, with a variety of purposes and implications. Often scholars deploy it to characterize the success of a system that removes the possibility of human defection or error. When used in this sense, “trust” does not act as a design lever at all, but rather as a description of the outcome that signifies confidence in the *system's* performance, not the other person. To characterize trust as an element subject to design intervention, I use it to refer to an attitude that agents in the interaction possess toward each other: it is a belief that

others in a given system will act in ways that are cooperative toward the trusting agent or the common goal when they do in fact have an effective choice to act in ways harmful or helpful to the trusting agent. Risk is therefore a precondition to trust. To facilitate trust, the architects of social systems should break down cooperative actions into observable chunks, where participants can lower their exposure to each other while observing the proclivities of others to cooperate or defect, for example. I treat trust in this sense as one of the mechanisms of intrinsic motivation, because it is an internal belief about others, by definition not under conditions where the other's action can be controlled (including, presumably, under conditions where threatened retaliation would under-enforce cooperative behavior), and because this internal belief triggers for many an intrinsic will to reciprocate the anticipated cooperative behavior with like cooperation. Trust as intrinsic motivation plays the role of anticipatory cooperation, in anticipation of which others can open an interaction cooperatively even before actual cooperative action has been taken by the other toward them.

Another consistent finding of the experimental literature is that perceptions of fairness are endogenous to the cooperative dynamic. Selfish rationality puts fairness of outcomes aside, focusing on whether the individual is made better or worse off by an interaction as the sole predictor of that agent's behavior. Consistent with this approach, scholars engaging in policy analysis often separate considerations of fairness from claims about predicted efficiency, turning to the former only after egocentrically-defined incentives have induced the desired level of activity. The experimental literature consistently shows that this approach fails to take account of the extent to which people care about the fair distribution of outcomes, the perceived fairness of the intentions of others in the interactions, and probably the fairness of the process of the interaction.⁵⁵ What is socially and psychologically experienced as fair is, however, cross-culturally contingent and diverse, subject to framing, ideology, and manipulation. The goal of pursuing fairness as a design principle, then, is not to create one universal norm of fairness (e.g., equal division of the gains from interaction); instead, systems designers should seek to build in mechanisms to achieve widely held perceptions of fairness with regard to outcomes and intentions. Fairness in the particular interaction will often be a central target of norm development for the community.

Norms: intrinsic and extrinsic. Another relevant and substantial literature explores social norms, probing the workings of long-standing, usually tightly-knit communities.⁵⁶ These communities have typically integrated many of the design levers I try to separate into discrete building blocks, thereby fashioning ongoing, stable, social relations. When thinking of design for systems that may be as new as a collaborative wiki launched yesterday, or a new system for getting local citizens to offer solutions to local problems, social norms must refer to something more primitive than long-standing internalized norms. Norms in this minimal sense must be instances of more-or-less clearly understood behavioral expectations about what counts as “cooperative” in a given system. Once participants know what counts as cooperation, and what is defection, they can adjust their own actions, as well as judge the actions of others. At the simplest levels, these could be Schelling coordination norms: that is, norms that have no claim to apply other than that they are convenient focal points for coordination (e.g., if two New Yorkers decide to meet at 1 O'Clock PM, they will disproportionately choose to meet next to the clock in Grand Central). Beyond that, they can be explicitly stated expectations about behavior, like those that anchored Wikipedia and made it unique among cooperation models in its early days on being purely norms-based. Finally, they can be evocations of existing background norms that are internalized and long-standing in a given applicable community, applied metaphorically or by analogy to the new context.

Norms can be either extrinsically or intrinsically motivating, depending on the history of their adoption and the degree to which they reflect internalized discipline. Some work suggests that norms that are explicitly adopted by a group after deliberation achieve high adherence.⁵⁷ This is consistent with the idea, central to the psychology of crowding out, that what makes for intrinsic motivation is self-determination. Here, at the collective or “public” level, norms chosen through a process of self-governance can be internalized as “belonging” to the agent who lives by them (an actual, psychologically-felt consent of the governed, rather than the notional one of social contract). These norms, while “extrinsic” in their form as external constraints on behavior, really are a species of internal reasons for action, what Deci and Ryan called “integrated regulation.”⁵⁸ This is the kind of effect observed by Edward Balleisen, in this volume, as a common argument in favor of self-determination in the construction of norms for self regulation. On the other hand, norms can be externally given, or largely random in their selection (such as driving on the right hand or left hand side), in which case they will improve cooperation, but through external processes of compliance and comprehension of what is cooperative, not through internal adaptation of desired behavior.

Efficacy. Coming out of the psychological work on motivation, people seem to work on projects that make them feel effective, or competent. A cooperation system is improved to the extent that people see their actions individually, as well as their collective effort, as being effective. We see this in studies of fundraising, where reports of success and exhortations of how close we are to the goal are common, as well as in the shape of online peer production efforts, whose adoption and take off suggests that as the prospect of successful cooperation becomes clearer over time, it draws increasing numbers of contributors.

Factors affecting extrinsic motivation

Punishment and reward. Of all the findings in the experimental literature, most consistent with the selfish rational actor model is the importance of punishment and reward. In many studies, the presence of mechanisms for punishing defectors and/or rewarding cooperators improve cooperation. Given that roughly one-third of agents generally behave selfishly, and that in many contexts enduring cooperation depends on effective deterrence of their free riding, this result should come as no surprise. The “surprising” or “altruistic” aspect of this research rather involves the existence of a sizable segment of participants in experiments who will undertake actions to punish defectors, and thereby police them, even when these actions have a net-negative private return to them personally in the game design. This punishment represents a second-order public goods problem, because those who mete out punishment bear the full cost, while the benefits are shared by all participants, in some game designs, or reaped by entirely other participants, in other game designs. The experimental literature finds, however, that with the right design, reciprocators can solve the second-order public goods problem of punishment without intervention from an external body, such as the state or management.⁵⁹

This finding is complemented by many field studies, which show that mutual monitoring and graduated sanctions on defectors play an important role in sustaining, for example, commons-based resource systems. Ostrom's studies of irrigation districts, and the role played by local “riders” checking that participants do not take too much water from the common irrigation district accompanied by a system of sanctions enforced locally offer the classic example.⁶⁰ Experimental work suggests, however, that poorly designed punishment mechanisms can also backfire, leading to punitive rounds of misfired retaliation and deterioration of cooperation. Fehr and Rockenbach, for example, ran a trust game with punishment. In a trust game, experimenters give “principals” a sum. Principals can transfer as much or

as little of the endowment to a “trustee” agent. The experimenter then multiplies whatever was transferred to the trustee, and gives the new large endowment to the trustee. The trustee can then go home with the entire multiplied endowment, or transfer as much or as little of it as she wishes to back to the principal. Selfish rational actor model predicts that trustee will transfer back nothing, and therefore principal will entrust nothing and go home with the original endowment. Many experiments have shown that trustees do transfer back substantial amounts, and that these amounts increase in the degree of trust exhibited—that is, the proportion of the initial endowment transferred. In this particular series of experiments, however, principals were allowed to specify in advance that they would impose a penalty on trustees who did not transfer back, or were given the option affirmatively to abjure the power to punish. Inverting the prediction of the selfish rational actor model, trustees transfer back the smallest portion of the entrusted amount when punishment is threatened, a middling amount when punishment is not possible, and the most when principals, though they have the option to punish, disclaim it in advance.⁶¹ This is an example of another design consideration that I discuss later: crowding out. Here, it suggests that punishment, introduced as a threat before action in a two-person game, has a very different valence than punishment available as a background power, without affirmative threat, and in group settings where the imposition of punishment itself is an act of cooperation in the second-order public goods game. Calibrating the negative and positive effects of punishment in repeat games is likely to continue to be a substantial focus of research. One study suggested that negative reciprocity can be triggered even in parties who “deserved” the punishment because they were first to defect,⁶² and another, that the extent to which retaliation and spiteful punishments devolve into feuds is cross-culturally diverse.⁶³ On the other hand, a recent study suggests that the potential negative effects of punishment disappear in longer games.⁶⁴

Transparency/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 depend on participants knowing who did 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 did not call it that) – such accounting, they suggest, was necessary to sustain indirect reciprocity, which in turn may have been the driving force behind the evolution of human intelligence.⁶⁵ Whether or not they are correct, studies in experimental economics typically show that reputation-rich games lead to cooperation more quickly and robustly than anonymous games.⁶⁶ Similarly, reputation systems play a significant role in social software platforms, ranging from commercial systems like eBay and Amazon, to the wide range of commons-based peer production projects that deploy the possibility of creating a stable locus for reputation, and observable behavior and opinion, as a major design element.

Transparency can affect both intrinsic and extrinsic motivations. Intrinsically, participants can observe that others are cooperating, and be driven to reciprocate. Participants can also gauge the payoffs and processes, and determine the fairness of the interaction. Extrinsically, it facilitates monitoring, punishment, and reward. Transparency requires a system that truthfully reflects actions, outcomes, and intentions, and hence must be designed to weed out deception. This is an observation that is, of course, quite central to regulation. In the public regulatory framework, Joseph Stiglitz in his essay here identifies disclosure as the first modality of formal public regulation, while Balleisen's essay emphasizes the centrality of transparency to the design of self-regulation as well.

Cost. Cost is an additional consideration. People 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 payoff structure,⁶⁷ or because cooperative tasks are structured in fine-grained modules so that the cost of useful incremental contribution becomes smaller for any given individual participating in peer production.⁶⁸

Crowding Out. An important design constraint already introduced in the discussion of punishment and norms is crowding out. Intra-system crowding out refers to situations when use of one design lever would reduce the efficacy of another. The effect of threatened punishment in a two-person trust game I described above is one example, but the crowding out of trust by punishment has older roots in the literature.⁶⁹ Inter-system crowding out can occur when one tries to mix-and-match elements from cooperative systems with elements from other systems, such as market mechanisms or bureaucratic control systems. In economics, the literature is anchored in the old Titmuss-Arrow debate over blood donations and sales. Since then, a large literature has demonstrated that introducing money into a relationship can crowd out non-monetary motivations and undermine otherwise cooperation-based interactions.⁷⁰ The most widely used psychological explanation for crowding out is self-determination theory. Arising from the work of Edward Deci, beginning around the same time, but in apparant mutual ignorance of, the Titmuss-Arrow debate, self-determination theory posits that individuals have innate needs for competence, autonomy, and relatedness.⁷¹ Certain kinds of extrinsic rewards and controls tend to make individuals feel that all or some aspects of these needs are rejected by the person offering the reward or punishment, leading to a decline in motivation. Other possible explanations may have more to do with the social dynamic more consistent with a reciprocity model of cooperation—the fact that social capital, for example, can only be built within activities not fully fungible with monetized relationships.⁷² Given that we observe many mixed systems, such as open source software innovation and certainly the introduction of cooperative models into firms, mixing is not impossible. Indeed, the psychological literature itself sees the tenor and framing of rewards as important, as well as the type of tasks involved. But introduction of money in particular, and similar tangible rewards, as well as formal bureaucratic control, requires attention to the interactions between the motivational and organizational forms, rather than a simple assumption of additive effect.⁷³ This creates a particular problem for policy, where a policy maker will, by definition, be external to the social interaction and acting upon it from the outside. Monica Prasad's discussion of carbon taxes in this volume is an excellent example of the tension. Managing crowding out then becomes a central problem that others, for example technical platform designers, have less of a need to address as long as use of their platform is not mandatory or users are not conscious of the constraints the system imposes on their interaction.

Exit and Entry. First, the ease of exit and entry into a cooperation platform is an important design element. At baseline, where exit is not itself a form of defection, and where participation does not easily permit expropriation of collectively created value, easy exit and entry will usually draw cooperators and repel selfish actors. Cooperators will tend to select into a cooperative framework, and would perceive its cooperative characteristics and absence of defection opportunities as affirmative benefits. Selfish actors, by contrast, will select themselves out unless there are opportunities for gainful abuse.⁷⁴ In other contexts, such as employment, easy exit tends to leave firms with easier recourse to maket-based mechanisms to structure their relationship, which undermines trust.⁷⁵ Here, longer term, enforceable commitments, and even asymmetric ease of entry and exit, may be preferable. Ironically, in many states, at-will employment combined with enforceable non-compete agreements achieves exactly the inverse of the desired asymmetry.⁷⁶ More generally, expensive entry and exit will deter defectors, who will only undertake entry when the potential gains from defection outweigh the cost. As another of the design levers suggests, however, high cost may also decrease the demand for

participation among cooperators.

Leadership/asymmetric contribution. Another significant element in the design of cooperation dynamics involves leadership, asymmetric contribution, and influence. This observation is anchored in organizational sociology⁷⁷ and examinations of open source software⁷⁸ and online cooperation. It is important to recognize, however, that “leadership” does not necessarily imply “hierarchy.” What is required of leadership, how asymmetric contribution leads to leadership and motivates it, are important areas of research into cooperation platforms, currently less well worked out than other areas. In the domain of online cooperation and social software, systems frequently allow some participants to take on heavier workloads, thereby earning a claim on their cooperators to exert greater influence over the process. This dynamic certainly characterizes Wikipedia and the larger free and open source software projects. Some scholars have argued that this tendency implies the reemergence of hierarchical organization, but this position overstates the degree of control exerted by “leaders,” and understates the degree of communication, persuasion, and agreement.

[insert Table 1 here]

C. Law and Policy: Examples of Cooperative Systems Design

Law and various other policy mechanisms are systems. They interact with other systems—technical, organizational, and social, most often—to allow people to act in the world: say, drive a car, or make an investment. Like these other systems, which have at times formed various observational grounds for the cooperation literature, law and policy are systems of affordance and constraint that can be (even if the often are not) susceptible to conscious design through purposeful human action. How might legal and policy design look if it were subject to analysis as a cooperation system? While no one has yet begun to apply the full range of the literature on cooperation to questions of law and policy, there are enough efforts to apply the insights of some of the work in experimental economics, and separately the work in organizational sociology, to offer us a glimpse of how applying this approach may facilitate the design of interventions in the future.

The rise of cooperation as an alternative approach to markets and hierarchies has placed the most direct and politically-mobilized pressure on law and policy in the areas of copyrights and patents, particularly, but not solely, as they relate to software and Internet-based cultural production and communication. The rise of peer production, first as free or open source software and then more generally throughout the digitally networked environment, has offered not only strong existence proof of enormously creative and innovative alternative models of production, but also widespread practices that are negatively affected by excessively strong exclusivity regimes.⁷⁹ It joined with the more general critique aimed at the position that strong patents and copyrights are justified by the need for powerful monetary incentives to motivate individuals and firms to undertake the costs and risks associated with research and development or mass-scale cultural production. While some of the arguments made against excessively strong exclusive rights rested on the public goods nature of information goods from the consumption perspective, much of it focused on the role of intrinsically and socially driven innovation and creative expression to defend the importance of the public domain and the commons.⁸⁰ Reflecting a deep recognition of the constitutive role of law in markets, similar to claims in Marc Eisner's essay in this volume, the strong emphasis on enhancing copyrights and patents was seen as putting the thumb on the scales in favor of incumbents, at the expense of new modalities of social

production.

Nowhere have the proponents of this alternative economic culture been more politically effective than in fighting to a standstill the efforts of the copyright industries to push for stronger exclusivity regimes in this decade. The year 1998 was the last great successful legislative year for the copyright industries in the United States: the passage of the Digital Millennium Copyright Act, the No Electronic Theft Act, and the Sonny Bono Copyright Term Extension Act marked what appeared to be an unstoppable coalition for expanding rights. But a loose coalition of free and open source software developers, civil society activists, and a small number of large companies dedicated to building free-software-based services and computing equipment business models, and whose business model did not depend on, and was harmed by, exclusivity rules, successfully fought the Hollywood-based coalition to a standstill around new paracopyright regulations like “trusted systems,” a proposed database protection law, and an exclusivity-expanding reform of the Uniform Commercial Code. In Europe, a similar coalition successfully blocked the adoption of a particularly contentious expansion of software patents by the European Union.

The case of the commons, and peer production, as the basis for resisting regulation of information production through expanded patents or copyrights offers one example of a broader implication of the study of cooperation. It suggests the possibility that, to some extent, the provisioning of public goods can be achieved through voluntary cooperation, rather than through regulation or markets. In this regard, it points in similar directions as does the study of social norms. I certainly have made similar claims about the possible use of user-owned wireless networks to provision last-mile Internet connectivity, in place of spectrum auctions or regulation.⁸¹ Christine Jolls has claimed that the fairness dynamic documented in experimental economics may suggest why minimum wage protection is unnecessary for domestic workers: they function in a work relationship that requires employers to pay “fair” (that is, above market-clearing) wages, without legal constraint.⁸² Robert Scott has made similar claims on behalf of the argument that courts should refuse to enforce incomplete contracts, rather than filling in missing details with commercially reasonable terms, thereby creating a regulation-free zone in which parties can explicitly be vulnerable to each other so as to occupy a trust and reciprocity dynamic, rather than a court-regulated one. Neither argument is clearly supported by the cooperation literature. Jolls's argument does not incorporate the cultural contingency of what is perceived as “fair” and its manipulability. Cultural “out group” status and local economic dynamics can make the baseline conditions for a class of workers—in this case domestic help—so abusive that even otherwise quite shabby treatment (by comparison to wages or terms in other sectors) can seem generous and generate a reciprocity dynamic by comparison to the treatment of others in the “out” group. In that case, application of minimum wage laws could effectively set a minimal level, without impeding the capacity of the fairness dynamic in places where fairness would drive wages above that level. Scott understates the significance of the third-party status of courts, and the trappings of “fairness” attached to their decisions, that actually offer discipline, or punishment and reward, services to the parties to enforce their own understanding of what is cooperative, or at least can do so if done properly. But the point is not to adopt any given particular presently-available interpretation of cooperation to law. The point is to emphasize that a focus on cooperation can have the implication that people can run their own affairs cooperatively without state intervention.

Moreover, the crowding out effect between discipline systems and trust also suggests not only that regulation may be unnecessary, but that it can be counterproductive. Taiwanese irrigation associations are one prominent case. Created and run by farmers, these voluntary associations achieved

considerable success in managing common irrigation systems. For local political reasons, in 1993 the Taiwanese government sought to look “pro farmer” by replacing association fees that farmers paid out of their own pockets with government grants. In later studies, that benign intervention seems to have shifted the associations from highly participatory associations of farmers who knew their system well and volunteered to run the association and to manage it, to a system that “one gets from the government.” The study does not account for the details and framing of displacement, but it is not implausible of how loss of a personal sense of being part of collective responsibility of a self-managed and provisioned system seems to have been associated with the breakdown. The introduction of “the state” *as system* crowded out the community as system.⁸³

While the dynamics of social cooperation, on the one hand, and corrupt political economy, on the other hand, caution us in adopting government-run policy interventions, the study of cooperation would be only mildly interesting if all it did was replace the mantra that “the market will take care of it all” with a similarly lethargic “social cooperation will take care of it all.” In particular, the fact that cooperation includes a significant component of group identity and member/non-member distinction, and harbors the potential for intergroup conflict, and the fact that groups can have all sorts of goals, not all of them laudable or defensible from the perspective of a society at large, suggests that an important role of applying cooperation to institutional design is to achieve greater efficacy in the pursuit of democratically-adopted public goals, as well as offering “services,” like legal dispute resolution, that can be relied upon by parties in a cooperative dynamic as an aspect of, at least, the discipline elements of the social interaction.

One important line of implementation in this vein has been the work of Dan Kahan and Tracy Meares on policing and criminal law enforcement.⁸⁴ At root, their approach has been to emphasize interventions that are aimed at treating the community in which there is a high crime rate, in contradistinction to the Becker-inspired rational maximizer focus on adjusting the payoffs to the individual criminal. Such interventions may prove uncomfortable to a wide range of political views. Liberals would oppose low-threshold stops and temporary arrests intended to disrupt gangs and allow community members to congregate on the streets instead, fostering stronger local community and commitment, or the idea that the police should engage with African American churches to mobilize local communities. Conservatives would likely oppose the argument that the ubiquity and length of incarceration do more harm than good by breaking up communities and in particular excluding families of convicts from the communities.⁸⁵ My point here is not to support these particular interventions, but to show an instance of policy intervention aimed to address a certain public goods problem—local security—through treating the local social dynamics of cooperation in the community, rather than through adjusting the incentives of the individual criminal conceived as a selfish utility maximizer.

Another important area that lends itself to design based on the insights of cooperation is labor and employment law. Much of the sociological literature on cooperation examines the workings of business organizations. Stone's work on the new psychological contract offers an especially fruitful example, drawing on this literature to suggest a wide range of adaptations in labor, employment, trade secret and antidiscrimination law.⁸⁶ Stone argues that business processes have changed, with leading firms seeking a more fluid relationship, aimed at engaging employees' affective commitment but moving away from lifetime employment. This strategy requires a new level of training and general knowledge acquisition by employees, at the employers' expense, which employees are then free to take with them as they move from one job to the next. The relative stability of union defined jobs and job ladders as well as non-union white collar positions that married excellent benefits with an expectation

of career-long service, has thus been replaced by, among other features, flatter structures, and less formal task definitions and structures of promotion. As a result, many older legal arrangements have fallen out of step, suggesting the importance of adjusting them to fit the more cooperative dynamic observed in these kinds of workplaces.

More generally, the scholarly emphasis on social cooperation has prompted numerous proposals for reform of tax policy and tax enforcement.⁸⁷ Furthermore, Bowles and Gintis have written about how our understandings of the demands of reciprocity should shape the design of our welfare system. In particular, they emphasized that the demands of reciprocity justify reducing or limiting welfare payments made to those whom the state can properly identify as free riders, making those payments available only to reciprocators and cooperators.⁸⁸ Amy Wax later developed a more worked out version of this argument, applied specifically to the 1996 Personal Responsibility and Work Opportunity Act.⁸⁹

One of the more detailed studies of interventions is Ann Carlson's review of literature on recycling norms, suggesting which interventions seem to have resulted in higher utilization and compliance rates.⁹⁰ Consistent with some of the design levers outlined here, lowering costs of recycling (by curb pickup) was an important element, but, consistent with a crowding-out dynamic, high prices had a more ambiguous impact. Face-to-face group meetings and feedback, and appeals from boy/girl scout visitors were effective, consistent with the importance of humanization, communication, and norm-setting. Feedback on household performance tended to be more effective in increasing low contributors' contributions when given in a group meeting than in writing individually, but slightly lowered participation by those households contributing more than their fair share—consistent with the fairness lever. Group feedback, delivered in face-to-face meetings, resulted in higher compliance level after the intervention ended than observed in groups with individual written feedback. Whether this more persistent efficacy resulted from the creation of a mechanism for social sanctions in the community, from a sense of humanization of the participants and local solidarity, or from an adaptation of personal principles and goals through communication is not clear from any of these studies. The efficacy of meeting face-to-face, talking, and providing feedback in the context of a community, which clearly enables any and all of these forms, was clear.

Conclusion

Application of the extensive work on cooperation to law and policy interventions remains in its infancy. Present translational work of the kind I described on minimum wage law, contract law, or recycling, has mostly focused on one or two of the design levers, using them as a springboard from some particular reform proposal, or some defense of an existing doctrine in the face of critique. An important part of the reason to outline the design levers or focal points and map their interaction is to begin to develop a framework that would allow us to organize and discipline our use of this literature as we come to apply it to questions of law and policy design.

Design for cooperation begins with a different model of human beings than the selfish rational actor model. It emphasizes the diversity of human motivational profiles, and the importance of the interaction to determining actual behavior. To the extent that the literature probing cooperation better predicts human behavior under differently designed systems, it holds the promise of improvement in the design of systems for human action. Just as theorists and policy-makers applied the selfish rational actor model in very different contexts, so too can scholars and officials apply cooperation to very different systems. Technical systems, such as online collaboration forums; business processes and

organizational strategies; legal and regulatory regimes; and constructed social contexts are all systems of affordance and constraint for human action. They can all utilize cooperation-based design approaches.

One may wish to analyze whether a GPL license is better than a BSD license for free or open source software development models, or whether trade secret law should, or should not, include the inevitable disclosure doctrine given a shift to knowledge-based, more cooperative business processes. The questions one would ask would be different under the cooperative or the selfish model. Similarly, if one is trying to decide whether it is important to include a profile page on a collaborative wiki site, or whether to allow anonymity, introduce tiered privileges based on length of time that a user has been part of a cooperative effort, or introduce explicit pricing into one's technical platform, these are all amenable to cooperation-based analysis. My point is merely to emphasize that refining design based on the large and diverse empirical and theoretical literature on cooperation holds the promise of significant improvement in our ability to design human systems for cooperation.

We have seen a long term trend in “scientific policymaking” to push back on foundational cultural norms and educational practices that are, if not universal, certainly broadly shared across cultures. The hours spent by parents teaching their children to share toys, to play with a child who is perhaps unpopular, to be polite, to tell the truth because it is the truth, not because it is expedient, etc.; the religious teachings of “love thy neighbor;” and the long tradition of argument from principle and the demands of ethical behavior all fall by the wayside when “scientific policy makers” brush off their game theoretic models and translate the myriad emotions, relations, commitments, and beliefs that make up human action into a question of mechanism design. In the past twenty years or so in academic literature, in the past decade and a half in some of the business literature, and in the past decade in explicit writing about our core technical systems of communication and computation, we have seen a series of efforts to reassert the human and the social. The drive in some cases has been intellectual—in the case of evolutionary theory, for example, too many observations refused to fit pure selfishness, and were better explained by expanded models of reciprocity or group selection. In other cases, the drive has been practical—in the case of management studies and the businesses studied in organizational sociology, rapid rates of change and the imperative to optimize learning and adaptation, rather than necessarily efficiency for more-or-less known conditions. In all events, the inputs into scientific policy making—theoretical, experimental, and observational—are increasingly pushing against holding on to universal selfishness as a core design assumption, and towards learning how to improve the cooperative social dynamics of human beings who will interact through the system under contemplation. We need to develop a field of cooperative human systems design to fill that need.

Design Element	Description
communication	Allowing participants to communicate, even without any enforceable commitment, increases cooperation ⁹¹
empathy/ humanization	Participants who identify with the counter-party to a game increase cooperation in social dilemma games and altruistic giving in Dictator Games ⁹²
solidarity	Participants who see themselves as part of a common identity group increase cooperation ⁹³
fairness	Participants consistently appear to care about the fairness of the outcomes, the intentions of other participants, and the processes. ⁹⁴
norms	The presence of even minimal coordination focal points, or Schelling norms, can improve cooperation by clarifying what is expected from whom and what counts as defection or cooperation. Self-chosen norms appear to improve cooperation, as will, likely, background norms already encoded by participants as “values.” ⁹⁵
trust	Trust as a design lever refers to an attitudinal stance participants can have toward each other. As a design lever it is narrower than the term is usually applied, and characterizes a belief agents have about the likely actions of others when unconstrained by other system elements. When it functions, it acts as a form of anticipatory cooperation, which agents can “reciprocate” by themselves cooperating in their first move. ⁹⁶
efficacy	Individuals are internally driven to act with competence and efficacy. Providing a sense of efficacy in the cooperation likely improves intrinsic commitment to the cooperative project. ⁹⁷
punishment/ reward	There is consistent evidence that introduction of a possibility of punishment into social dilemma games increases cooperation by keeping selfish actors in check. Punishment is a complex social and psychological phenomenon, is cross culturally contingent, may decrease overall value of the cooperative activity, and may crowd out trust and fairness. ⁹⁸
crowding out	Systems can crowd each other out, and elements within a given system can crowd each other out. The introduction of money into an interaction can limit participation motivated by intrinsic motivations. There is also evidence that introducing punishment can crowd out trust. Crowding out complicates cooperative systems design: not all potential interventions interact positively. ⁹⁹
transparency (reputation)	Transparency of cooperation platforms enables agents to observe what others are doing, characterize actions, intentions, and outcomes, and identify cooperation for positive reciprocity and defection for negative reciprocity. Reputation is a core instance of transparency. ¹⁰⁰
cost	The level of cooperation is sensitive to, but not dominated by, the cost of cooperation. ¹⁰¹
exit/entry	Whether a system is easy or hard to enter will affect the mix of types that participate and the level of trust participants will have. The direction of the effect will depend on whether exit itself is a form of defection and whether there are opportunities for appropriation within the interaction. ¹⁰²
leadership/asymmetric contribution	Leadership is important in creating and sustaining cooperation. ¹⁰³ Leadership is not hierarchy, but the flexibility of a system to allow asymmetric contributions and levels of capabilities and powers within a system without upsetting the cooperative dynamic.

Table 1: Design Considerations for Cooperation

- 1 Herbert Simon, *A Behavioral Model of Rational Choice*, in *MODELS OF MAN, SOCIAL AND RATIONAL: MATHEMATICAL ESSAYS ON RATIONAL HUMAN BEHAVIOR IN A SOCIAL SETTING* (1957); Amartya Sen, *Rational Fools: A Critique of the Behavioral Foundations of Economic Theory*, *PHIL & PUB AFF.* (1977).
- 2 Milton Friedman, *ESSAYS IN POSITIVE ECONOMICS* (1953).
- 3 Robert Ellickson, *ORDER WITHOUT LAW: HOW NEIGHBORS SETTLE DISPUTES* (1991).
- 4 Ostrom *GOVERNING THE COMMONS* (1991).
- 5 Christine Jolls, Cass R. Sunstein, and Richard Thaler, *A Behavioral Approach to Law and Economics*, 50 *STAN L. REV.* 1551 (1998).
- 6 Bruno Frey and Margit Osterloh, *MANAGEMENT BY MOTIVATION: BALANCING INTRINSIC AND EXTRINSIC MOTIVATIONS* (2002); Charles F. Sabel, *Ungoverned Production*, in *CONVERGENCE AND PERSISTENCE IN CORPORATE GOVERNANCE*, Jeffrey N. Gordon and Mark J. Roe, eds. 310-327 (2004); *THE FIRM AS A COLLABORATIVE COMMUNITY*, Charles Heckscher and Paul Adler, eds. (2006); John Hagel and John Seeley Brown, *THE ONLY SUSTAINABLE EDGE: WHY BUSINESS STRATEGY DEPENDS ON PRODUCTIVE FRICTION AND DYNAMIC SPECIALIZATION* (2005).
- 7 Sabel, *Ungoverned Production.*; John Seeley Brown and Paul Duguid, *Knowledge and Organization A Social-Practice Perspective*.
- 8 Yochai Benkler, *Coase's Penguin or Linux and the Nature of the Firm*, 112 *YALE LJ* 369 (2002); Andrea Forte and Amy Bruckman, *Scaling Consensus: Increasing Decentralization in Wikipedia Governance*, *PROC. 41ST HAWAII INT'L CONFERENCE ON SYSTEMS SCIENCES* (2008).
- 9 Gary S. Becker, *Crime and Punishment, an Economic Approach*, 76 *J. POL. ECON.* 169-217 (1968).
- 10 Dan M. Kahan and Tracey L. Meares, *The Coming Crisis of Criminal Procedure*, 86 *GEO. L.J.* 1153 (1998); for the current state of federal and state investment in community policing see <http://www.cops.usdoj.gov/Default.asp?Item=34>.
- 11 See Sabel, *supra*; Heckscher and Paul, *supra*; David Isenberg, *Rise of the Stupid Network*, 1997, <http://isen.com/stupid.html>; Jerome Saltzer, David P. Reed, and David Clark, *End-to-End Arguments in System Design*, 2 *ACM Transactions in Computer Systems* 277 (1984); Henry Jenkins, *Convergence Culture: Where Old and New Media Collide* (2006) David S. Alberts, John J. Garstka, and Fredrick Stein, *Network Centric Warfare: Developing and Leveraging Information Superiority* (2d ed. 1999).
- 12 Three excellent reviews are Ernst Fehr and Herbert Gintis, *Human Motivation and Social Cooperation: Experimental and Analytical Foundations*, 33 *ANNU. REV. SOCIOLOG.* 43, 50 (2007); Colin F. Camerer & Ernst Fehr, *Measuring Social Norms and Preferences Using Experimental Games: A Guide for Social Scientists*, in *FOUNDATIONS OF HUMAN SOCIALITY: ECONOMIC EXPERIMENTS AND ETHNOGRAPHIC EVIDENCE FROM FIFTEEN SMALL-SCALE SOCIETIES*, Joseph Henrich *et. al.*, eds, 2004) 55-95; Elinor Ostrom, *A Behavioral Approach to the Rational Choice Theory of Collective Action*, 92 *AM. POL. SCI. REV.* 1, 1-22 (1998).
- 13 Camerer & Fehr.
- 14 Henrich *et. al* *FOUNDATIONS*; The findings were introduced and subject to insightful critique in Henrich *et. al.*, *"Economic Man" in Cross-Cultural Perspective: Behavioral Experiments in 15 Small-Scale Societies*, 28 *BEHAVIORAL AND BRAIN SCIENCES* 795-855 (2005).
- 15 Benedikt Herrman, Christian Thonni, and Simon Gaechter, *Antisocial Punishment Across Societies*, 316 *SCIENCE* 1362 (2008).
- 16 Matthew Rabin, *Incorporating Fairness into Game Theory and Economics*, 83 *AM. ECON. REV.* 1281 (1993).
- 17 Elazar Barkan, *THE RETREAT OF SCIENTIFIC RACISM: CHANGING CONCEPTS OF RACE IN BRITAIN AND THE UNITED STATES BETWEEN THE WORLD WARS* (1993).
- 18 E.O. Wilson, *SOCIOBIOLOGY* (1976); Richard Dawkins, *THE SELFISH GENE* (1976).
- 19 Stephen J. Gould and Richard C. Lewontin, *The Spandrels of San Marco and the Panglossian Paradigm: A Critique of the Adaptationist Model*, 205 *PROC. ROYAL SOC.* 581-598 (1979).
- 20 See. *e.g.*, David Buss, *EVOLUTIONARY PSYCHOLOGY: THE NEW SCIENCE OF THE MIND* (3rd ed. 2008).
- 21 W.D. Hamilton, *The Evolution of Altruistic Behavior*, 97 *AMERICAN NATURALIST* 354, 354-356 (1963); John Maynard Smith, *Group Selection and Kin Selection*, 201 *NATURE* 1145, 1145-46 (1964).
- 22 Robert L. Trivers, *The Evolution of Reciprocal Altruism*, 46 *Q. REV. OF BIOLOGY* 35, 35-57 (1971).
- 23 Richard Alexander, *THE BIOLOGY OF MORAL SYSTEMS* (1987); Martin Nowak and Karl Sigmund, *Evolution of indirect reciprocity* 437 *NATURE* 1291-1298 (2005).
- 24 Martin Nowak, *Five Rules for the Evolution of Cooperation* 314 *SCIENCE* 1560-1563 (2006).
- 25 Elliot Sober and David Sloan Wilson, *UNTO OTHERS: THE EVOLUTION AND PSYCHOLOGY OF UNSELFISH BEHAVIOR* (1998).
- 26 Samuel Bowles and Herbert Gintis, *The Evolution of Strong Reciprocity: Cooperation in Heterogeneous Populations*, 65 *THEORETICAL POPULATION BIOLOGY* 17, 17-28 (2004); Sober & Wilson, *supra*.

- 27 See Robert Boyd and Peter J. Richerson, *CULTURE AND THE EVOLUTIONARY PROCESS* (1986); Peter J. Richerson and Robert Boyd, *NOT BY GENES ALONE: HOW CULTURE TRANSFORMED HUMAN EVOLUTION* (2004).
- 28 Dan Simon, *A Third View of the Black Box: Cognitive Coherence in Legal Decision Making*, 71 *CHICAGO LAW REVIEW* 511 (2004); Jon Hanson & David Yosifon, *The Situational Character: A Critical Realist Perspective on the Human Animal*, 93 *Geo. L.J.* 33 (2004).
- 29 Edward L. Deci and Richard M. Ryan, *The "What" and "Why" of Goal Pursuits: Human Needs and Self-Determination of Behavior* 11 *PSYCH. INQUIRY* 227 (2000).
- 30 Samuel Bowles, *Policies Designed for Self-Interested Citizens May Undermine "The Moral Sentiments": Evidence from Economic Experiments* 320 *SCIENCE* 1605-1609 (20 June 2008).
- 31 Toshio Yamagishi, Nobuhito Jin, and Toko Kiyonari, *Bounded Generalized Reciprocity, Ingroup Boasting and Ingroup Favoritism*, 16 *ADVANCES IN GROUP PROCESSES* 161, 161-197 (1999).
- 32 Charles Heckscher and Paul Adler, *THE FIRM AS A COLLABORATIVE COMMUNITY: RECONSTRUCTING TRUST IN THE KNOWLEDGE ECONOMY* (2006); Walter W. Powell, *Neither Market Nor Hierarchy: Network Forms of Organization*, 12 *RESEARCH IN ORGANIZATIONAL BEHAV.* 295 (1990).
- 33 Charles F. Sabel, *A Real Time Revolution in Routines*, in Heckescher and Adler.
- 34 Elinor Ostrom *GOVERNING THE COMMONS: THE EVOLUTION OF INSTITUTIONS FOR COLLECTIVE ACTION* (1990).
- 35 Benkler, *Coase's Penguin*; Clay Shirky, *Social Software and the Politics of Groups*, March 9 2003 (published online, http://shirky.com/writings/group_enemy.html).
- 36 Ernst Fehr and Herbert Gintis, *Human Motivation and Social Cooperation: Experimental and Analytical Foundations*, 33 *ANNU. REV. SOCIOL.* 43, 50 (2007).
- 37 See Fehr et al, *Neuroeconomic Foundations*.
- 38 Herrman et al, *supra*.
- 39 For variations in quality and focus of punishing responses between Western European and some post-Communist societies see Simon Gaechter, Benedikt Herrman, and Christian Thoni, *Cross Cultural Differences in Norm Enforcement*, 28 *BEHAVIORAL AND BRAIN SCIENCES* 822 (2005). The big difference was represented in the studies of small-scale societies, where deviations were much larger and correlated to cultural mapping of the games onto group-specific practices, some of which called for super-generosity or none at all where market-integrated societies would have had a focal point of equal division of windfall gains. See Henrich et al, *FOUNDATIONS*.
- 40 Yochai Benkler & Helen Nissenbaum, *Commons-based Peer Production and Virtue*, 14 *J. POL. PHIL.* 394 (2006).
- 41 Lee A. Dugatkin, *COOPERATION AMONG ANIMALS: AN EVOLUTIONARY PERSPECTIVE* (1997).
- 42 David Sally, *Conversation and Cooperation in Social Dilemmas VII RATIONALITY AND SOCIETY* 57, 57-92 (1995).
- 43 Sabel, *Ungoverned Production*, *supra*.
- 44 Henri Tajfel & John Turner, *An Integrative Theory of Intergroup Conflict*, in *THE SOCIAL PSYCHOLOGY OF INTERGROUP RELATIONS*, (W.G Austin and S. Worchel eds. 1979); Toshio Yamagishi, Nobuhito Jin, and Toko Kiyonari, *Bounded Generalized Reciprocity, Ingroup Boasting and Ingroup Favoritism*, 16 *ADVANCES IN GROUP PROCESSES* 161, 161-197 (1999).
- 45 *FIRM AS COLLABORATIVE COMMUNITY*, *supra*.
- 46 See S. Alexander Haslam, *PSYCHOLOGY IN ORGANIZATIONS: THE SOCIAL IDENTITY APPROACH* (2001);
- 47 Henri Tajfel, quoted in S. Alexander Haslam, Rachel A. Eggin, and Katherine J. Reynolds, *The ASPIRe Model: Actualizing Social and Personal Identity Resources to Enhance Organizational Outcomes*, 76 *J. OCCUPATIONAL AND ORGANIZATIONAL PSYCH.* 83, 84 (2003).
- 48 Tajfel & Turner, *supra*.
- 49 Haslam, *et. al.* at 84-86.
- 50 Boyd and Richerson, *NOT BY GENES ALONE*, *supra*, chapter 6.
- 51 Samuel Bowles and Herbert Gintis, *Persistent Parochialism: Trust and Exclusion in Ethnic Networks* 55 *J. ECON BEHAVIOR & ORG.* 1 (2004); Samuel Bowles and Herbert Gintis, *Social Capital and Community Governance* 112 *ECON J.* F419-436 (2002).
- 52 D. Rand *et. al.*, (in progress); D. Rand *et al.* *From Friend to Foe* (submitted).
- 53 Bohnet & Frey, *supra*.
- 54 Rilling et al (2002), *supra*; Rilling et al, (2004) *supra*. See also Tania Singer *et al*, *Empathetic neural responses are modulated by perceived fairness of others*, 439 *NATURE* 466 (2006) (describing changes in empathy level based on counterparty's play in a prisoner's dilemma game).
- 55 Fehr and Schmidt, *Fairness and Reciprocity*, *supra*.
- 56 Robert Ellickson, *ORDER WITHOUT LAW: HOW NEIGHBORS SETTLE DISPUTES* (1991).
- 57 Ostrom, Walker, Gardner, 1994, *supra*.

- 58 Deci and Ryan, *supra*.
- 59 Bowles and Gintis, *Homo Reciprocans*; Fehr & Gächter, *Altruistic Punishment*.
- 60 Ostrom GOVERNING THE COMMONS; Ostrom *Behavioral Approaches*; see also Henry E. Smith, *Semicommon Property Rights and Scattering of in the Open Fields*, 29 J. LEGAL STUD 131 (2000).
- 61 Fehr and Rockenbach, *Detrimental Effects*; Falk et. al, *Driving Forces*.
- 62 Dreber et al, *supra*.
- 63 Hermann et al, *supra*.
- 64 Simon Gaechter, Elke Renner, and Martin Sefton, *The Long-run Benefits of Punishment* (unpublished).
- 65 Nowak and Sigmund, *supra*.
- 66 Ernst Fehr and Simon Gaechter, *Cooperation and Punishment in Public Goods Experiments*, 90 AM. ECON. REV. 980, 980-94 (2000).
- 67 Fehr and Camerer, in FOUNDATIONS.
- 68 Benkler, *Coase's Penguin*.
- 69 Toshio Yamagishi, *The Provision of a Sanctioning System as a Public Good*, 51 J. PERSONALITY SOCIAL PSYCHOLOGY 110-16 (1986).
- 70 Bruno S. Frey and Reto Jege, *Motivation Crowding Theory: A Survey of Empirical Evidence*, 15(5) J. ECON. SURVEYS 589 (2001).
- 71 Deci and Ryan, *supra*.
- 72 Yochai Benkler, *Sharing Nicely: On Shareable Goods and the Emergence of Sharing as a Modality of Economic Production*, 114 YALE LJ. (2004).
- 73 Samuel Bowles, *Policies Designed for Self-Interested Citizens May Undermine "The Moral Sentiments": Evidence from Economic Experiments* 320 SCIENCE 1605-1609 (20 June 2008).
- 74 Ostrom, *Behavioral Approaches*.
- 75 John Paul MacDuffie and Susan Helper, *Collaboration in Supply Chains With and Without Trust*, in FIRM AS COLLABORATIVE COMMUNITY.
- 76 Matt Marx, Deborah Strumsky, and Lee Fleming, *Mobility, Skills, and the Michigan Noncompete Experiment* (working paper 2008).
- 77 Michael Maccoby and Charles Heckscher, *A Note on Leadership for Collaborative Community*, in FIRM AS COLLABORATIVE COMMUNITY, *supra*, at 469-478.
- 78 Steve Weber, *THE SUCCESS OF OPEN SOURCE* (2005).
- 79 Benkler, *Coase's Penguin*; Benkler, *Sharing Nicely*.
- 80 See, among many other works, Jessica Litman, *DIGITAL COPYRIGHT* (2001); Siva Vaidhyanathan, *COPYRIGHTS AND COPYWRONGS: THE RISE OF INTELLECTUAL PROPERTY AND HOW IT THREATENS CREATIVITY* (2001); Lawrence Lessig, *FREE CULTURE: HOW BIG MEDIA USES TECHNOLOGY AND THE LAW TO LOCK DOWN CULTURE AND CONTROL CREATIVITY* (2004); Eric Von Hippel, *DEMOCRATIZING INNOVATION* (2006); Yochai Benkler, *THE WEALTH OF NETWORKS, HOW SOCIAL PRODUCTION TRANSFORMS MARKETS AND FREEDOM* (2006).
- 81 Yochai Benkler, *Overcoming Agoraphobia: Building the Commons of the Digitally Networked Environment* 11 HARV. J. L. & TECH 287 (1998); Yochai Benkler, *Some Economics of Wireless Communications*, 16 HARV. J. L. & TECH. 25 (2002).
- 82 Christine Jolls, *Fairness, Minimum Wage Law, and Employee Benefits*, 77 NYU L. REV. 47 (2002).
- 83 See Elinor Ostrom, *Policies that Crowd Out Reciprocity and Collective Action*, in *MORAL SENTIMENTS AND MATERIAL INTERESTS: THE FOUNDATIONS OF COOPERATION IN ECONOMIC LIFE*, (Herbert Gintis, Samuel Bowles, Robert Boyd, and Ernst Fehr, eds., 2005) 253, 261-63.
- 84 Dan M. Kahan and Tracey L. Meares, *The Coming Crisis of Criminal Procedure*, 86 GEO. L.J. 1153 (1998).
- 85 Donald Braman, *DOING TIME ON THE OUTSIDE: INCARCERATION AND FAMILY LIFE IN URBAN AMERICA* (2004).
- 86 Katherine Stone, *The New Psychological Contract: Implications of the Changing Workplace for Labor and Employment Law*, 48 UCLA L. REV. 519 (2001).
- 87 Dan Kahan, *The Logic of Reciprocity: Trust, Collective Action, and Law*, 102 MICH L. REV 71 (2003); Leandra Lederman, *The Interplay Between Norms and Enforcement in Tax Compliance*, 64 OHIO L.J. 1453 (2003).
- 88 Samuel Bowles and Herbert Gintis, *Is Equality Pasa?* 22 BOSTON REVIEW. no 3-4 (Summer 1997).
- 89 Amy Wax, *A Reciprocal Welfare Program*, 8 VA. J. SOC. POL'Y 477 (2001).
- 90 Ann Carlson, *Recycling Norms* 89 CAL. L. REV 1231 (2001).
- 91 Sally, *supra*.
- 92 Iris Bohnet & Bruno Frey, *The sound of silence in prisoner's dilemma and dictator games*, 38 J OF ECON. BEHAVIOR & ORG. 43-57 (1999); James K. Rilling et al, *A Neural Basis for Social Cooperation*, 35(2) *Neuron* 395-405 (2002); James

- K. Rilling et al, Opposing Bold Responses to Reciprocated and Unreciprocated Altruism in Putative Reward Pathways, 15(6) *Neuroreport*, 2539-43 (2004).
- 93 Toshio Yamagishi, Nobuhito Jin, and Toko Kiyonari, *Bounded Generalized Reciprocity, Ingroup Boasting and Ingroup Favoritism*, 16 *ADVANCES IN GROUP PROCESSES* 161, 161-197 (1999); L. Zucker, *Production of Trust: institutional sources of economic structure, 1840-1920* 8 *RESEARCH IN ORGANIZATION BEHAVIOR* 53 (1986); Miller McPherson, Lynn Smith-Lovin, and James M. Cook, Birds of A Feather: Homophily in Social Networks, 27 *Annual Review of Sociology* 415 (August 2001); S. Alexander Haslam, *Psychology in Organizations: The Social Identity Approach*. (Sage Publications Ltd. 2001).
- 94 Ernst Fehr and K. Schmidt, *Theories of Fairness and Reciprocity, Evidence and Economic Applications, in Advances in Economics and Econometrics -- 8th World Congress*, Econometric Society (M Dwatripont et al, eds 2001).
- 95 Elinor Ostrom, Roy Gardner, and James Walker, *RULES, GAMES, AND COMMON POOL RESOURCES* (1994).
- 96 Diego Gambetta (ed.). *TRUST, MAKING AND BREAKING COOPERATIVE RELATIONS* (Blackwell, Oxford. 1988).
- 97 Deci and Ryan, *supra*.
- 98 Samuel Bowles and Herbert Gintis, *Homo Reciprocans: Altruistic Punishment of Free Riders*, 415 *NATURE* 125 (2002); Ernst Fehr & Simon Gächter, *Altruistic Punishment in Humans*, 415 *NATURE* 137 (2002); Ernst Fehr and Bettina Rockenbach, *Detrimental effects of sanctions on human altruism*, 422 *NATURE* 137 (2002); Armin Falk et al., *Driving Forces Behind Informal Sanctions*, 73 *ECONOMETRICA* 2017 (2005). Punishment is a double edged sword and may undermine cooperation, see Anna Dreber et. al. *Winners Don't Punish*, 452 *NATURE* 348, 348-51 (2008); effects are cross-culturally varied Benedikt Herrman et. al, *Antisocial Punishment Across Societies*, 319 *SCIENCE* 1362, 1362-67 (2008).
- 99 Bruno S. Frey and Reto Jeje, *Motivation Crowding Theory: A Survey of Empirical Evidence*, 15(5) *J. ECONOMIC SURVEYS* 589 (2001); Fehr and Rockenbach, *supra.*; Bowles (2008) *supra*.
- 100 Ernst Fehr and Simon Gächter, *Cooperation and Punishment in Public Goods Experiments*, 90 *AM. ECON. REV.* 980-94 (2000) (explaining that partner treatment with reputation yields higher levels of cooperation).
- 101 Camerer & Fehr, *supra*; Benkler, *Coase's Penguin, supra*.
- 102 Elinor Ostrom, *A Behavioral Approach to the Rational Choice Theory of Collective Action*, 92 *AM. POL. SCI. REV.* 1, 1-22 (1998)
- 103 Michael Maccoby and Charles Heckscher, *A Note on Leadership for Collaborative Community, in FIRM AS COLLABORATIVE COMMUNITY, supra*, at 469-478; Weber, *supra*.