INDISPENSABLE FRAMEWORK OR JUST ANOTHER IDEOLOGY? THE PRISONER'S DILEMMA AS AN ANTI-HIERARCHICAL GAME

For Rationality and Society

by

Aaron Wildavsky

Professor of Political Science and
Public Policy and
Member, Survey Research Center
University of California, Berkeley

Working Paper 91-11

IGS
INSTITUTE OF GOVERNMENTAL STUDIES

UNIVERSITY OF CALIFORNIA AT BERKELEY

March 1991

INDISPENSABLE FRAMEWORK OR JUST ANOTHER IDEOLOGY? THE PRISONER'S DILEMMA AS AN ANTI-HIERARCHICAL GAME

For Rationality and Society

by

Aaron Wildavsky

Professor of Political Science and Public Policy
and
Member, Survey Research Center
University of California, Berkeley

Working Paper 90-11

INDISPENSABLE FRAMEWORK OR JUST ANOTHER IDEOLOGY?
THE PRISONER'S DILEMMA AS AN ANTI-HIERARCHICAL GAME

On reading the chapter on rationality in Jon Elster's Nuts and Bolts of the Social Sciences, I found a brief discussion of the Prisoner's Dilemma that the author terms "the natural, indispensable framework for understanding human interaction" (p. 28). Wow! This framework is that human beings, acting on their own, do not necessarily, indeed, often do not, achieve as good results for themselves and for others as they could if they acted collectively with a view toward maximizing their joint interests. Is this axiom self-evident, namely, that individual and collective interests often diverge? Or is it a powerful new truth telling human beings that they can damage themselves and destroy their planet by following their selfish interests?

There are, it is true, seemingly endless analyses of what appears to be (almost) every academic's favorite game, The Prisoner's Dilemma (TPD), in which each prisoner could get a lighter sentence if each agreed not to squeal on the other but, since they can't or won't coordinate their decisions, and they cannot make binding commitments, each gets a longer sentence. The most frequent form in which this game appears,

I wish to thank John Harsanyi, Robert Powell, and Paul Tayler for their critical comments.

in my observation, is as Garrett Hardin's Tragedy of the Commons in which common use resources, such as grass, are ruined by overgrazing because it is in the short-term interest of each individual to use as much as he wants.

If choice by individuals alone leads (a) to terrible consequences, such as nuclear war and environmental degradation while (b) being readily fixable through collective arrangements, it is understandable that (c) anger grows at the irrationality of it all. Indeed, anger is mobilized even more at the idea of rationality itself for making it appear rational for each individual to participate in creating a collective catastrophe threatening, some say, the very existence of our common home, planet earth. This combination of mobilizing anger in the service of collective solutions with an apparent scientific (well, at least reasoned) basis has proven exceedingly attractive. What grounds have I to cast suspicion on so elegant and popular a formulation?

The Prisoner's Dilemma As An Anti-Authority Game

The classic account from which most students of the subject take their view of TPD is this:

> Two suspects are taken into custody and separated. The District Attorney is certain that they are guilty of a

> specific crime, but he does not have adequate evidence to convict them at a trial. He points out to each prisoner that each has two alternatives: to confess to the crime the police are sure they have done or not to confess. If they both do not confess, then the District Attorney states he will book them on some very minor punishment; if they both confess, they will be prosecuted, but he will recommend less than the most severe sentence. But if one confesses and the other does not, then the confessor will receive lenient treatment for turning state's evidence, whereas the latter will have "the book" slapped at him.[1]

Presumably the two suspects have committed a crime together. Each is presumably alone when told that if both remain silent they will be convicted and jailed for one year. Each is told that she can go free if she helps convict the one who remains silent. Silent Samantha, as we may call her, will then get ten years while the prisoner who confesses will get none. Should both confess, however, each will get nine years. Further, it appears that each prisoner is quite indifferent to the fate of her colleague. Neither can communicate with the other but both apparently believe what the District Attorney tells them and knows at least that the other prisoner is aware of the rules of this game. What is more, there is no way for

the prisoners to enforce any agreement they might make even if they could figure out how to communicate.

The essence of the game is in its structural payoffs. No matter what the other prisoner does, Silent Samantha (her name is as misleading as her behavior is self-regarding) is better off confessing first. True, Silent Samantha would get at most one year if her partner in crime refuses to confess but not knowing that for sure she minimizes her maximum loss by confessing first. The irrationality allegedly exists in the fact that a cooperative solution would leave them both better off than if each attends to their narrow self-interest of the moment.

My own view is quite different: the rationality of the prisoners, like rationality in general, is context dependent; without a supportive cultural context,[2] no strategy makes sense. The cultural bias of the game players is evident in the fact that I have been unable to find anyone who thinks it would be rational for both to confess on the grounds that admitting the truth would enhance the authority of the criminal justice system thereby improving both the safety and the morality of the collective. Silence makes sense for egalitarians who wish to undermine authority as inegalitarian on its face. But silence would be irrational for those who wish to strengthen hierarchy. Given cultural pluralism, moreover, different strategies will be rational for prisoners

within different cultural contexts. The solution favored by almost everyone, i.e., silence, for instance, would be anathema in a hierarchical culture that seeks to inculcate respect for authority.

Observe that silence against authority in the person of the DA is almost universally accepted as the rational course of action, because it guarantees each prisoner the best result, and that irrationality inheres in the fact that at each step it is rational for each prisoner to confess, i.e., support authority. The elimination of the office of District Attorney and its problems from the analysis (by now, no doubt, the reader has caught my drift) gravely distorts the analysis unless, of course, the analyst assumes that the normative situation in society is one where there is no legitimate authority. Though most analysts of TPD do not go as far as Burns and Buckley--blaming "the state" for creating the dilemma (they deem it part of the imperialist strategy of divide and rule)[3]--none (so far as I know,[4] not a single one) considers the fact, if it is a fact, that the prisoners are actually guilty of a crime.[5] Empirically, there is a chance the police are mistaken but the probability is low.

The ideological bias in the way the Prisoner's Dilemma game has been treated is apparent in the explicit rejection of individualism as selfish and hierarchy as coercive. Morton Deutsch is merely more open than others when he asserts that

"The PD game demonstrates that the unorganized social Darwinism which rationalized capitalism in terms of the competitive survival of the fittest has self-defeating consequences when applied to situations that are not purely competitive."[6] Yet, given the enormous achievements of competition in politics, economics, and science and the fact that most people in recorded history have lived under hierarchical arrangements, which means they cannot be good for nothing, something would seem to be missing. That something is cultural context.

Roger Hurwitz's "Strategic and Social Fictions in the Prisoner's Dilemma" is the work most sensitive to social context. He sees both the variety of possibilities in telling different stories around the same events that are called TPD and that the way in which the dilemma is resolved depends crucially on "preexisting social relations":

> The variety of meanings that the same choice requires over these stories indicates that the options are evaluated for their propriety according to preexisting social relations among the characters and the interest with which the storyteller or reader identifies. This criterion reflects a different social ontology than that of the game theory, where agents are isolated individuals with external and accidental relationships.

> . . . By specifying whether one is to help or hurt, trust or mistrust other members of the relationship, these norms stabilize and reproduce the particular social relationships. Yet compliance with them is vulnerable to the pressures of personal interests and to other norms.[7]

Like the fictions of "the original position," or "the social contract" that it resembles, rational choice analyses of TPD treat individuals as if they had no preexisting values or relations to defend, a condition that might be called anomic if it is at all recognizably human.

In explaining social and political phenomena, there is no such thing as an "objective" situation. That would remove from the situation the knowing perceiver. Without a cultural bias, a frame for perception, there would be not objectivity but chaos. All perception is selective. Actors understand an "objective" situation in the context of their cultures and act accordingly. One therefore could not hope to explain their actions without reference to their cultures. Even if the "objective" circumstances measured in terms of prison sentences is held constant, different actors may understand this "objective" prisoner's dilemma differently. In a very hierarchical society in which the actors have been socialized to value obedience highly, for instance, there may be no

dilemma at all if the confession is more highly prized than escaping with a lighter sentence by remaining silent. Or to use a different example, the threat to impose a certain "objective" amount of military punishment may suffice to deter some, but it is almost sure to fail to deter those who see martyrdom as the surest way to heaven.

Utilities or payoffs in rational choice models are meant to summarize the actors' preferences. To the extent that individuals in different cultures value the same outcomes differently, there will be a different set of payoffs and, consequently, a different model of this situation for each culture. As a very crude first approximation, one might think of culture as being implicit in a model's payoffs.

What a modeler should be saying is, in effect, "Tell me how individuals understand their situation in terms of their preferences, and I will then model the strategic aspects of this situation and try to provide some insight into its strategic dynamics." As cultures change and actors understand their situation differently, new models will be needed. The question of whether or not the Prisoner's Dilemma provides an "indispensable framework" then becomes an empirical question. Can many of the most profound issues individuals construct in a given culture or, more generally and heroically, in any culture be seen as a Prisoner's Dilemma? I argue that they cannot.

The Prisoner's Dilemma As An Ideological Construct

In order to clear up a possible misconception, I should say in advance that an ideology or cultural bias is not necessarily false; the predictions it generates could be true. Like other theories, of course, ideologies present only slanted views of how the world works. To make sure that the partial character of ideologies is understood, cultural theorists, following Mary Douglas, call them "cultural biases."[8] At the same time, perceptions of the same objects or situations may and do vary widely. But that is not to say that those who hold them can make them come true.

In order to make my point about the ideological character of The Prisoner's Dilemma, it is useful to consider its main competitor as a basis for public policy, so far as I can see, the unintended consequences, to use Karl Popper's formulation, of purposeful social action. Nowadays it is used almost always to warn against the unwanted consequences of proposed or existing governmental policies.

Here we have it: rival assumptions about what the world is really like--The Prisoner's Dilemma in which individuals could so easily better their lot and protect nature if only they adopted collective regulations, and its mirror image, the Entrepreneur's Paradise Lost (EPL), endless damage done by

governmental regulators who make things worse when all they need do is allow the magic of the market to work its wonders by liberating individual initiative. The difference is that TPD converts what ought to be a positive-sum collective game, in which everyone benefits, into a negative-sum encounter, in which all the players lose due to individual self-interest, while EPL converts the magic-market into a negative-sum game because of collective regulation that undermines the beneficent effects of competition among self-interested actors.

Just as deconstruction of texts, by undermining the words of authors, subverts their authority, so these rival metaphors-as-models (TPD versus EPL) constitute part of the struggle for power between collectivists and individualists. Which is not to say these model metaphors are unreal or any more unreal than other ways of organizing social life. These models are both problems and solutions, each the solution to a problem generated by another mode of organizing and each in turn the creator of problems it cannot solve. Were this not so, the world could only be one way with only a single optimal mode of social organization. Were this so, the one superior way would drive out the inferior ones as a sort of sameness settled over society. Since we do not find convergence but instead observe conflict, we will see that TPD, like EPL, not only criticizes the other cultures but provides solutions to

the defects created by them. To what solution, we may ask, is TPD the accompanying problem?

Russell Hardin argues that TPD is a general solution of another and simpler problem of collective action, the free-rider.[9] It will facilitate understanding, I think, to begin with the more immediately understandable problem--small members exploit large ones by failing to pay their fair share (free-riding) because they know that the large have so great a stake they will contribute the whole amount. Comparability to TPD comes from the symmetry of their situations: a good result for all could be achieved except for the unfortunate circumstance that some people have an incentive to shirk their responsibilities to the collective by free-riding. The highest payoff to each individual comes when he doesn't pay but everyone else does. Unfortunately, if too many attempt to ride free, the collective enterprise will collapse so everyone loses. As in TPD, free-riding subordinates the general interest of the collective to the private interests of individuals to their common disadvantage.

Given this general structure--the problem of reconciling individual and collective, public and private, interests--the basic modes of social organization may be conceived of either as embodiments of this dilemma or as efforts to guard against it. Thus, in a functioning hierarchy, whose adherents believe in its principles, free-ridership cannot occur because the

individual components are required by the system to pay their share. They believe they ought to fulfill these commands and, if they don't, there is plenty of coercive authority to compel them to do so. (Of course, people in hierarchical positions may be tempted to use their places for personal advantage; this ordinary occurrence illustrates how solving one problem creates others.) Competitive individualists overcome free-riding because each investor is rewarded according to her contribution; less input receives lower reward. Without suggesting that this equivalence is entirely achieved, a culture whose adherents believe that there ought not to be a free lunch should not have many free riders. (A belief in competition, I should add, is opposed to winning by rigging the game.) Fatalists would love to free ride for they will not risk anything by cooperating with others. By the same token, however, their reluctance to contribute or to cooperate at all provides them with far fewer opportunities. You have to walk, as the saying goes, before you can ride.

Only egalitarianism contains the conditions for creating and then sustaining free ridership. These conditions are three: (1) one person (or group), one vote, together with (2) grossly unequal stakes in the outcome sought after, followed by (3) no normative requirement of members to adhere to majority decisions.[10] If there is coercion, as in hierarchy, or voting in proportion to contribution (the stake in society

theory), as in individualism, there cannot be free riding. Do the same sort of considerations fit The Prisoner's Dilemma?

TPD in Cultural Context

TPD requires that there be two equal participants who could gain if they make a collective choice but are prohibited from communicating with each other or altering their preferences on the basis of new information. The flaw here lies in rooting rationality in individual behavior as if these individuals were social isolates. But they are not. The moment the fog created by the fiction of social isolation is pierced, the world looks quite different because the individual actor faces quite different social contexts. By exploring these contexts, I will show that the dilemma exists but only in a well-defined social context.

Suppose TPD is played in an individualistic culture. There could be no limit on communication without violating the precept that allows any deals to which there is genuine consent. Each prisoner would then have the ability as well as the incentive to seek better information. Individualists, then, would not meet the conditions required to experience TPD.

In a formal game-theoretic analysis, to be sure, it is not the inability of the prisoners to communicate or seek new

information that is most debilitating; the problem is that they cannot make binding contracts or promises to remain silent. Even if they have agreed to remain silent, one will see this as cheap talk when the other is taken off to the interrogation room with the D.A. It is equally obvious that inability to make binding agreements between consenting adults violates the norms of individualism.

Nor could hierarchists be afflicted with TPD: should the prisoners identify with the same hierarchy as the DA, they might put the good of their group above their own welfare by confessing. Only if the prisoners were in a rival culture might silence appear to them an unalloyed good. If the prisoners were of equal status, moreover, each, being hierarchical, would sacrifice for the other, who stood for the collective. If their statuses were unequal, the lower status prisoner would try to help the higher and the higher status prisoner would follow the norm of the sacrifice of the parts for the whole. The dilemma does not exist unless both prisoners feel it is right as well as efficacious to gain by informing on the other.

Egalitarians would also insist on talking with other egalitarians because, for them, deliberation, followed by mutual consent, is the only way of legitimating decisions. Worse still, that someone else, a hierarchical authority, like the DA, should decide for them would be anathema, adding

inequality to injustice. Best is that both confound their oppressors by staying silent. Better both should be hung than one should be treated differently than the other.

Where, if not in these forms of social organization, is TPD located? Among fatalists. Consider their views of human and physical nature. Because fatalists believe that nature operates in a random fashion, they see no point in intervention. Because they believe that human nature is capricious (one cannot tell what people will do), they are unwilling to cooperate with others. Banfield called this view "amoral familism" in that, at most, cooperation was limited to the family.[11] Thus fatalism is the culture that inculcates extremely short-time horizons--do whatever is necessary to avoid harm at the moment--and downgrades knowledge because there is no pattern to be discovered. It is this behavior that TPD captures. The prohibition of contact among the prisoners merely formalizes what fatalists would not do, i.e, seek knowledge or social support or sign any contract.

TPD does not describe just any social context but a particular type. The kinds of relationships envisaged under TPD are succinctly described by Eleanor Ostrom:

> There is no more irrational way to structure any enduring situation than that represented by the PD game: no communication among the participants, no previous

ties among them, no anticipation of future interactions, and no capacity to promise, threaten, or cajole.

Would reasonable humans, trying to order their own long-term relationships in a productive manner, structure a situation in such a perverse way?[12]

Yes, I would answer, they would, if, like fatalists, they believe human nature is capricious and physical nature is random. Only fatalists, the point is, would wish normatively to behave as Ostrom indicates. Though they might not wish communication among them to be prohibited, they act as if such a rule were in force because they do not trust anyone else enough to believe what they say or to cooperate with them. Unless fatalists were dealing with members of their immediate family, past ties would not matter to them; nor would it seem advisable for them to consider future consequences because one could never tell what those would be. Neither promising nor threatening would be desirable since neither would be effective. TPD does not describe just any social context but a very special one.

Ostrom does not wish to accept the view that "we are doomed to accept the imposition of structure by external authorities as the only way out of perverse situations such as the Commons Dilemma."[13] The usual solution is to recommend some version of Hobbes' Leviathan, i.e., the prisoners give

their power of attorney to a third party who controls their responses so they will do better. Carried over to the tragedy of the Commons, the same sort of reasoning would lead to national government control of resource decisions.[14] Ostrom, by contrast, would like a private property system to be considered as well as voluntary associations.[15] I agree that there is always a plurality of institutional alternatives. If that pluralism is accepted as a basic condition of life, however, then the conditions underlying TPD cannot be accepted as given but are seen for what they are, namely, one but far from the only one of a number of alternative modes of social organization. Thus what Ostrom calls "analysis of rules for making rules" or James Buchanan and Gordon Tulloch in their Calculus of Consent call constitutional choice is always pluralistic because the rules that make the rules for making rules are the ways of life these rules are designed to favor. Culture is bedrock.

Like other important social phenomena, TPD cannot occur just anywhere but only under facilitating cultural conditions. The fascination with this dilemma, its portrayal viz Elster as hardwired into the human condition, I reiterate, is culturally biased. It reflects the largely collectivist orientation of the scholars attracted to this kind of game. TPD is a form of system blame for placing people in a position where they exhibit fatalistic responses.

Hurwitz is concerned that concentration of the technical aspects of TPD obscures its moral relevance. "It is a dubious tribute to U.S. social science," he writes, "that its textual traditions have so easily transformed this inherently political parable into a technical problem." The morals he draws include "(1) the state's interference with individuals' talking helps reproduce state power and (2) silence is resistance to authority."[16] I suspect that it is precisely the ability to intertwine moral with technical concerns, to be good while being scholarly as well, that accounts for TPD's popularity.

A nice illustration of "combining mathematical muscle and redistributional relevance, the quintessentially desirable academic mix," as Robert Klitgaard put it to me, comes from a (if not the) leading student of TPD, Anatol Rapoport, who writes that

> From the moral point of view, therefore, Harsanyi's defense of the non-cooperative "solution" of Prisoner's Dilemma is by no means reprehensible. It need not be construed as an "advice" to players to play noncooperatively, but on the contrary, as an advice to seek ways of effecting enforceable agreements so as to turn the uncooperative game into a cooperative one.[17]

Were papers on TPD flurries of arrows shot by archers of old, one could hardly avoid getting hit by the conviction that cooperation is good while going it alone is not.

It has occurred to observers that while some institutional arrangements work out poorly and thus may be classified as self-stultifying dilemmas, other arrangements appear to work out well. How, if TPD is ubiquitous, is this possible? One line of inquiry suggests that the game is played not only once but many, many times so that in a reiterated TPD people observe that things are going badly and sometimes are able to make positive changes. In his Anarchy and Cooperation, Taylor shows that cooperation is possible unless the time horizons of individuals are exceedingly short.[18] As we have seen, fatalists fit this formula. Arthur Stinchcombe says that people solve problems only when this is necessary. Observing that fidelity in marriage works out better than its opposite or that inculcating norms of personal honesty is better than universal thievery, people do what is necessary to secure such outcomes. There must be more to sociology, Stinchcombe argues, than prisoners' dilemmas, given that a dilemma like shirking at work is routinely solved in capitalist societies.[19]

Stinchcombe is aware that he is using a functional form of explanation. As people observe that certain patterns of behavior make them unhappy, Stinchcombe thinks, they change

their behavior ". . . Just as the seducer often loses out in the long run to the kindly, dull lover because the seduced one realizes kindness as a better solution." He doubts that the solution is a conscious one; rather, he hypothesizes that there "is a contingent conscious motivation," dependent on being activated by the consciousness over time that a problem needs solving.[20] I would add that in a society of some size there are people who have internalized several different forms of social organization that comprise potential solutions to a variety of problems. The Prisoner's Dilemma is not life but a slice of life that can be carved differently depending on the social context of the people involved.

Might fatalism be a rational response to events, or even desired as the best (or, at least, a good way) to live? Is it irrational to wish to live as if the world was populated by random events and capricious human beings? It might be rational if that was how the world was. A lot of social science, as well as common discourse, assumes that cooperation is always desirable. But not, as we shall see, for everyone in every culture.

Cooperation Occurs in Different Cultures

"Collective action problems arise," Elster informs his readers, "because it is difficult to get people to cooperate

for their mutual benefit." This important and apparently unexceptional sentence contains multiple difficulties due to its methodologically individualistic premises. Cultures, as I argued earlier, may be conceived as viable institutional answers to problems of how to retain order or, in this locution, cooperation. It follows that there are different forms of cooperation that support different cultures.

One kind, cooperation via coercion, binds individuals to accept the collective will. Individuals can cooperate by authority (hierarchy) or by consensus (egalitarianism) to decide what to do. But individuals can also cooperate through a series of bilateral bargains, A with B, B with C, A with D, on and on. Or there can be a central figure to which all are connected--the individualist network--though none are directly connected to each other. Thus activity may be directed toward common ends without most individuals being aware of or agreeing upon a common objective. <u>It is the form of cooperation--hierarchical or egalitarian or individualistic--that is at issue, not the fact of cooperation</u>.

The first moral of this story is that not everyone believes in cooperation. The second moral is that not everyone agrees on how cooperation should be organized. Though hierarchists value cooperation, for instance, they prefer that it be organized from top down, while individualists and egalitarians (who dislike intensely others

prescribing for them) want it to grow from the bottom up. And while individualists believe that what is fair is what is freely agreed, egalitarians believe that agreements cannot fairly be reached unless the participants possess equivalent resources.

Elster claims that "Collective action problems arise because it is difficult to get people to cooperate for their mutual benefit."[21] That the existence of any form of organization is a marvel, given the obstacles, I agree; for the most part, however, I believe this widely shared proposition omits essential qualifications. Since people prefer different ways of life, they also differ on what is "mutual" or a "benefit." Thus egalitarians may object to programs that improve life for most people if inequality is thereby increased; hierarchists may object to increased education if it results in attacks on authority; fatalists may object to paying for programs that promise to better their lot because they nevertheless do not believe this will happen. The assumption that all of us use the same evaluative framework, stated thus baldly, is false.

I expect that most game theorists would have some difficulty locating themselves in the preceding account. I shall try to help them but I do not guarantee they will agree that this is to our mutual benefit.

What Game? Whose Theory?

Game theorists, it would be fair to say, do not look upon TPD as ideology but as the way the world is. (To paraphrase John Wayne in one of his Western roles, that ain't brag, that's fact.) Fair enough but not fair under all conditions. If we are concerned with how a game, a set of social relationships really, plays out, then the payoff matrices become facts for those who stipulate them in advance. Here we have the simplest TPD in which P-1 and P-2 are faced with a payoff matrix such that they are always better off seeking immediate benefits, even if a transcendent vision would reveal that cooperation is better for them than immediate self-interest. In diagrammatic form the payoff matrix is protean enough presumably to cover such disparate events as arms races and tragedies of commons, providing only that the prisoners either cannot make agreements, because they cannot talk to each other, or because any agreement they could make would not be enforceable.

Payoff Matrices*#

	P-1	P-2
P-1	3, 3	1, 4
P-2	4, 1	2, 2

*Each player-prisoner is assumed to want a higher rather than a lower number.
#According to Robert Axelrod, the payoff matrix is not quite a Prisoner's Dilemma. The prisoners have a strategy which is as good as mutual cooperation; they can get as good a payoff exploiting one another on alternate turns. This is another kind of cooperation, but in order to get the "pure" PD Axelrod argues that the (4,1) and 1,4) payoffs should be changed to, for example, (5,0) and (0,5). See his The Evolution of Cooperation (New York: Basic Books, 1984).

What could be more factual than the combination of payoffs and stipulations commonly called TPD? Recall here W.V. Quines' famous (albeit insufficiently appreciated) adage to the effect that all facts are theory laden. It is not facts by themselves (no conjectures a la Popper, no refutations, no relevant facts) but facts together with the theories that alone can make sense out of them that are part of science. Thus I am not arguing that ideology determines whether a situation of social interaction modeled by the stipulations just enumerated is, according to convention, a prisoner's dilemma. Of course it is. If we raise not the secondary question of the player-prisoner's preference ranking among the

four possible outcomes, however, but the primary question of what game under which rules, i.e., under which cultural context should it be played, then that choice is ideological.

Ideology comes in in two ways: the constitutional choice, if that is available, of what kind of game to play and, if that is foreclosed, in attempted explanations of the prisoner's motivations in choosing to rank the outcomes the way they do. It is this cultural-ideological-value baggage people bring with them (given that they do not go through life naked of preferences for how human beings should live with each other) that might motivate one prisoner to defect while the other chose to cooperate and, as life would have it, to later feel ashamed of such behavior. Suppose, then, the defector lowers the value of confessing so that it is less than that of cooperating. Voila! There is no more TPD. My question is whether this change is merely a change in preference or a change of game?

The one-shot, single game TPD, I have argued in support of the structural change thesis, fits fatalists (and fatalists alone) because "distrust everybody" is a strategy that supports only fatalistic cultures. When reiterated or tournament TPDs are played, however, with the game going onward into the future, the players never knowing whether or when it will end, lasting social life comes into play, and that, except for fatalists, is something else again. For the

three other active cultures require trust of some kind--hierarchists trust authority in established institutions; individualists trust market exchange made by networks led by successful entrepreneurs; egalitarians trust local, participatory, anti-establishment, voluntary groups. That is why I think that a strategy of tit-for-tat, or trust until proven otherwise, beats unmitigated distrust. The necessity of trust also helps explain why the three active cultures not only oppose their rivals but also need them to prove the necessity of internal trust to counter external opposition.

Thompson and Warburton's analysis of the Tragedy of the Commons illustrates its dependence on different types of trust. The way the dilemma (or tragedy) is posed, solutions must either be egalitarian--keep the commons but change human nature from competitive to cooperative--or individualistic--although human nature can't be changed, the commons can be privatized.[22] Actually, either of these or other solutions might avert the tragedy. Tragedy is not inherent in the commons but in inappropriate institutional arrangements viewed in relation to the cultural context in which the task is performed.

In game theory, as in social science generally, emphasis is on how the participants go about getting what they want from the polity, economy, or society, given that they already know what they want but must act amidst relative prices, i.e.,

what others want. Cultural theory, by contrast, begins not by taking preferences for granted but by trying to explain why it is that people who adhere to different ways of life want what they want in the first place. Their actions in the second place are derived from an understanding of why they want what they want in the first place and how these preferences are, over time, reinforced or revised or rejected. By getting rid of the uncaused cause (interests or preferences explain everything but nothing explains them), the cultural context through which preferences are formed and reformed becomes central. Thus it is not rationality per se, as if one cultural form of rationality was reason itself, but plural rationalities, as Michael Thompson calls them, rationalities that vary with the preferences or objectives to be realized, that should be the mainstay of game theory.

NOTES

1. Duncan Luce and Howard Raiffa, Games and Decisions (New York: John Wiley, 1957), p. 95.

2. For a definition of cultures as composed of people who share values and beliefs (their biases) justifying their preferred pattern of social relations, see Michael Thompson, Richard Ellis, and Aaron Wildavsky, "Political Cultures," forthcoming in Routledge Encyclopaedia of Government and Politics; Thompson, Ellis, and Wildavsky, Cultural Theory (Boulder, Colo.: Westview Press, 1990); and Mary Douglas, Natural Symbols, Explorations in Cosmology (Harmondsworth: Penguin, 1970).

3. T. Burns and W. Buckley, "The Prisoner's Dilemma Game as a System of Social Domination," Journal of Peace Research, Vol. 11 (1974), pp. 221-28

4. More than 2000 papers have been written about it by 1975 and many thousands more since, while I have read only a few dozen. See Bernard Grofman and Jonathan Poole, "Bayesian Models for Iterated Prisoner's Dilemma Games," General System, Vol 20 (1975), pp 185-94

5. To his credit, Roger Hurwitz does note that "In contrast,

some American grade B movies of the 1930s and 1940s (for example, Boy's Town) celebrate a prisoner's confession as a victory for society and as the prisoner's repudiation of a corrupt influence and the start of a spiritual redemption to dispel the prisoner's or audience's fears that the confession was a betrayal, the accomplice would usually forgive the confessor and even praise him for 'taking the chance to go straight.' Nevertheless . . . the confession transforms the confessor's primary social bond from a partnership of equals to probationary status with society or subservience to the state as represented by the district attorney. The prisoner is let off, but not empowered." (Roger Hurwitz, "Strategic and Social Fictions in the Prisoner's Dilemma," in James Der Derian and Michael J. Shapiro, eds, International/Intertextual Relations: Post-Modern Readings of World Politics (Lexington Books, ____). Despite the author's obvious egalitarian preference, he at least lets the reader know that other people with other values might take a different view.

6. Quoted by Hurwitz, "Strategic and Social Fictions," p. 122, from Morton Deutsch, The Resolution of Conflict: Constructive and Destructive Processes (Yale University Press, 1973).

7. Hurwitz, p. 120.

8. Douglas, "Cultural Bias."

9. Russell Hardin, Collective Action (Baltimore: The Johns Hopkins University Press, for the Resources for the Future, 1982). Hardin writes that "The appeal of Prisoner's Dilemma, as with the logic of collective action, has been its generality and its apparent power in representing manifold social interactions. Indeed, the problem of collective action and the Prisoner's Dilemma are essentially the same . . .. [A] result that tells us that individual effort to achieve individual interests will preclude their achievement, because if the collective good is not provided, the individual member fails to receive a benefit that would have exceeded the individual's cost in helping purchase that good for the whole group" (p. 25).

10. See Betty H. Zisk, "Movement Politics, Voter Education, and the Electoral Process: Referenda Campaigns of Peace and Environmental Activists," Prepared for presentation at the Annual Meeting of the American Political Science Association, Washington, D.C., Aug. 31 to Sept. 4, 1988.

11. Edward Banfield, The Moral Basis of a Backward Society (New York: Free Press, 1958).

12. Eleanor Ostrom, "Institutional Arrangements and the Commons Dilemma," in Vincent Ostrom, David Feeny, and Hartmut Picht, eds, Rethinking Institutional Analysis and Development (International Center for Economic Growth, Institute for Contemporary Studies, 1988), p. 107.

13. Ibid.

14. Ibid., p. 106.

15. Ibid., pp. 106-108.

16. Hurwitz, "Strategic and Social Fictions," pp. 130-31.

17. Anatol Rapoport, "Prisoner's Dilemma--Recollections and Observations," in Brian Barry and Russell Hardin, eds, Rational Man and Irrational Society? An Introduction and Sourcebook (Beverly Hills: Sage Publications, 1982), pp. 71-84; quote on p. 81.

18. M. Taylor, Anarchy and Cooperation (London: John Wiley, 1976).

19. Arthur L. Stinchcombe, "Is the Prisoner's Dilemma All of Sociology?" in Inquiry, Symposium: Jon Elster, Logic and

Society, Vol. 23 (____), pp 187-92. Stinchcombe is commenting on Jon Elster's Logic and Society: Contradictions and Possible Worlds (New York: John Wiley, 1978).

20. Ibid, pp. 191-92.

21. Elster, Nuts and Bolts, p. 131.

22. Michael Thompson and Michael Warburton, "Uncertainty on a Himalayan Scale," Mountain Research and Development, Vol. 5, No. 3 (1985), pp. 75-76, 84-87.

IGS Working Papers

$3.50 each plus 20% for shipping and handling, tax where applicable

1991

91-18 **Exit, Voice and Loyalty: Bureaucratic Responses to Presidential Control During the Reagan Administration** Marissa Martino Golden

91-17 **Democracy and Self-Organization: The Systemic Foundation for the Democratic Party** Gus diZerega

91-16 **The Myth of the Independent Voter** Raymond E. Wolfinger

91-15 **Bureaucratic Responses to the Administrative Presidency: The Civil Rights Division Under Reagan** Marissa Martino Golden

91-14 **The Emergence of Strong Leadership in the 1980's House of Representatives** Barbara Sinclair

91-13 **Social Ecology, Deep Ecology, and Liberalism: A Constructive Explorationof Murray Bookchin's Confusions** Gus diZerega

91-12 **Elites and Democratic Theory: Insights From the Self-Organizing Model** Gus diZerega

91-11 **Indispensable Framework or Just Another Ideology? The Prisoner's Dilemma as an Anti-Hierarchical Game** Aaron Wildavsky

91-10 **Intra-Party Preferences, Heterogeneity, and the Origins of the Modern Congress: Progressive Reformers in the House and Senate, 1890-1920** David W. Brady

91-9 **The Information-Seeking Behavior of Local Government Officials** Marc A. Levin

91-8 **Working in Practice But Not in Theory: Theoretical Challenges of "High-Reliability Organizations"** Todd R. LaPorte, and Paula M. Consolini

91-7 **Pork and Votes: The Effect of Military Base Closings on the Vote in Ensuing Congressional Elections** David Hadwiger

91-6 **Designing an Interactive, Intelligent, Spatial Information System for International Disaster Assistance** Louise K. Comfort

91-5 **Constitutional Mischief: What's Wrong with Term Limitations** Nelson W. Polsby

91-4 **Thermidor in Land Use Control?** Paul van Seters

91-3 **Parchment Barriers and the Politics of Rights** Jack N. Rakove

91-2 **The Study of American Politics in Japan: Its State, Trends, and Problems** Ichiro Sunada

91-1 **The Revision of California's Constitution: A Brief Summary** Eugene C. Lee

1990

90-34 **Recent Developments in Disease Prevention/Health Promotion in the Federal Republic of Germany,** Rolf Rosenbrock

90-33 **Recent Developments and Reform Proposals in the Politics of Pharmaceutical Supply in the Federal Republic of Germany,** Rolf Rosenbrock

90-32 **Speech Before the Meeting of Texas Public Interest Organizations,** David Cohen

90-31 **A Curious Life—The Pursuit of an Understanding of Public Administration,** James W. Fesler

90-30 **The Cultural Conquest of the Presidency: Incorporation and the Transformation of American Political Life, 1890-1916,** Peter Schwartz

90-29 **The Fat Lady Has Not Yet Sung: Is the Tax Revolt Over?** Randy H. Hamilton

90-28 **A Tightrope Walk Between Two Spheres of Logic: Observations—and Self-Observations—of a Social Scientist in Parliamentary Politics,** Rolf Rosenbrock

90-27 **Iran Air Flight 655 and the USS *Vincennes*: Complex, Large-Scale Military Systems and the Failure of Control,** Gene I. Rochlin

90-26 **Political Leadership and Value Change: Reagan, Thatcher and the Conservative Revolution?** Pippa Norris

90-25 **Party Rules and Women's Representation: Reforming the British Labour Party,** Joni Lovenduski and Pippa Norris

90-24 **Political Cultures,** Michael Thompson, Richard Ellis, and Aaron Wildavsky

90-23 **Ecological Regression and Voting Rights,** D. A. Freedman, S. Klein, J. Jacks, C. Smyth, and C. Everett

90-22 **The San Jose Metropolitan Area: A Region in Transition,** Donald N. Rothblatt

90-21 **The Demand for Referendums in West Germany "Bringing The People Back In?"** Wolfgang Luthardt

90-20 **Sunset As Oversight: Establishing Realistic Objectives,** Cynthia Opheim, Landon Curry, and Pat Shields

90-19 **Government Expenditure Levels: Alternative Procedures for Computing Measures,** Brian Stipak

90-18 **Transformation of American Liberalism, 1940s-1980s: An Analysis of Liberal Policy Change and the ADA,** Ichiro Sunada

90-17 **The Politics of Policy: "Political Think Tanks" and Their Makers in the U.S.-Institutional Environment,** Winand Gellner

90-16 **CAUTION: Excessive Use of Government Statistics May be Injurious to the Health of the Body Politic,** Randy H. Hamilton

90-15 **Thermidor In Land Use Control?** Paul van Seters

90-14 **Taxation For a Strong and Virtuous Republic: A Bicentennial Retrospective,** W. Elliot Brownlee

90-13 **How the Cases You Choose Affect the Answers You Get: Selection Bias in Comparative Politics,** Barbara Geddes

90-12 **Counterfactuals and Hypothesis Testing in Political Science,** James D. Fearon

90-11 **Pat Crashes The Party: Reform, Republicans, and Robertson,** Duane M. Oldfield

90-10 **The Acquisition of Partisanship by Latinos and Asian-Americans: Immigrants and Native-Born Citizens,** Bruce E. Cain, D. Roderick Kiewiet, and Carole J. Uhlaner

90-9 **New Perspective on the Comparative Method,** David Collier

90-8 **California Agency Reconnaissance Project: Teaching Public Administration Through Field Research,** Todd R. La Porte and David Hadwiger

90-7 **Earthquake Safety For New Structures: A Comprehensive Approach,** Stanley Scott

90-6 **Government Policies And Higher Education: a Comparison of Britain and the United States 1630 to 1860,** Sheldon Rothblatt and Martin Trow

90-5 **Dominance and Attention: Images of Leaders in German, French, and American TV News,** Roger D. Masters, Siegfried Frey, and Gary Bente

90-4 **Nonverbal Behavior and Leadership: Emotion and Cognition in Political Information Processing,** Roger D. Masters and Denis G. Sullivan

90-3 **The Dredging Dilemma: How Not to Balance Economic Development and Environmental Protection,** Robert A. Kagan

90-2 **Turning Conflict Into Cooperation: Organizational Designs for Community Response in Disaster,** Louise K. Comfort

90-1 **The Effect of Campaign Spending, Turnout, and Dropoff on Local Ballot Measure Outcomes** and **The Initiative and California's Slow Growth Movement,** David Hadwiger

1989

89-27 **On Campaign Finance Reform: The Root of All Evil is Deeply Rooted,** Daniel Hays Lowenstein

89-26 **Toward A Dispersed Electrical System: Challenges to the Grid,** Jane Summerton and Ted K. Bradshaw

89-25 **Top Bureaucrats and the Distribution of Influence in Reagan's Executive Branch,** Steven D. Stehr

89-24 **Designing an Emergency Information System: The Pittsburgh Experience,** Louise K. Comfort

89-23 **Learning From Risk: Organizational Interaction Following the Armenian Earthquakes,** Louise K. Comfort

89-22 **The Elusiveness of Rural Development Theory and Practice: Domestic and Third World Perspectives Joined,** Ted K. Bradshaw

89-21 **Saints and Cardinals in Appropriations Subcommittees: Academic Pork Barreling and Distributive Politics in an Era of Reistributive Budgeting,** James D. Savage

89-20 **The Case for Experiential Knowledge**, Gene I. Rochlin

89-19 **Choice vs. Control: Increasing Organizational Effectiveness in Interdependent Environments**, Louise K. Comfort and Keun Namkoong

89-18 **Bingo! An Untapped Revenue for California Cities**, William B. Rumford, Jr. and Randy H. Hamilton

89-17 **The Research on Higher Education Program: An Appreciation of Eskil Bjorklund**, Martin Trow

89-16 **Santa Cruz County Planning Issues: Papers on Planning, Housing And Forestry**, Edward J. Blakely and Ted K. Bradshaw

89-15 **Why Pretend One Size Fits All: An Examination of Management Issues That Concern Small Federal Agencies**, Randy H. Hamilton

89-14 **Music of the Squares, a Lifetime of Study of Public Administration**, Herbert Kaufman

89-13 **Technology and the Future: Issues Before the Bush Administration**, Edward Wenk, Jr.

89-12 **How Much Does Law Matter? Labor Relations in Rotterdam and U.S. Ports**, Robert A. Kagan

89-11 **Party, State, and Ideology in the U.S. House of Representatives, 1967-76**, K. G. Armstrong

89-10 **Ernst Fraenkel Lecture, Free University of Berlin the American Election of 1988: Outcome, Process and Aftermath**, Nelson W. Polsby

89-9 **The Ambiguous Status of Science and Technology in Australia**, Anthony Pecotich and Kelvin Willoughby

89-8 **1992, European Integration and *The Times***, David Morgan

89-7 **American Higher Education: "Exceptional" or Just Different?** Martin Trow

89-6 **American Higher Education: Past, Present and Future**, Martin Trow

89-5 **The Distribution of Academic Earmarks in the Federal Government's Appropriations Bills, FY 1980-1989**, James Savage

89-4 **London 2001**, Peter Hall

89-3 **The University of London: An American Perspective**, Eugene C. Lee, Frank M. Bowen

89-2 **Ukiah, 1904: A Modest Footnote to the History of the Council-Manager Form of Municipal Government in the United States**, Randy H. Hamilton

89-1 **American Identity and the Politics of Ethnic Change**, Jack Citrin, Beth Reingold, Donald P. Green

1988

88-27 **Locality and Custom: Non-Aboriginal Claims to Customary Usufructuary Rights as a Source of Rural Protest**, Louise Fortmann

88-26 **Bruce Keith's Almanac: Patterns of Voting in California**, Bruce Keith

88-25 **Cold Turkeys and Task Force: Pursuing High Reliability in California's Central Valley**, Todd R. La Porte and Ted Lasher

88-24 **Environmental Ethics in California**, Carolyn Merchant

88-23 **Crisis as Opportunity: Designing Networks of Organizational Action in Disaster Environments**, Louise K. Comfort

88-22 **The Logic of Uncertainty: Interorganizational Coordination in International Disaster Assistance**, Louise K. Comfort

88-21 **Information Networks in International Disaster Assistance**, Louise K. Comfort

88-20 **The Decay of Federal Theory**, S. Rufus Davis

88-19 **Inside Japan's Leviathan Decision-Making in the Government Bureaucracy**, Brian Woodall and Nobuhiro Hiwatari

88-18 **Technology and Adaptive Hierarchy: Formal and Informal Organization for Flight Operations in the U.S. Navy**, Gene I. Rochlin and Energy Resources Group

88-17 **From Crisis to Community: The 1988 Oil Spill in the Pittsburgh Metropolitan Region**, Louise Comfort, Joel Abrams, John Camillus, and Edmund Ricci et al.

88-16 **The Arrogance of Optimism**, Martin Landau, Donald Chisholm

88-15 **American Democracy in World Perspective and What to Do About It**, Nelson W. Polsby

88-14 **Modernization of the U.S. Senate,** Nelson W. Polsby

88-13 **The Iowa Caucuses in a Front-Loaded System: A Few Historical Lessons,** Nelson W. Polsby

88-12 **The Reagan Presidency After Seven Years,** Eugene C. Lee (moderator)

88-11 **The United States Air Traffic System: Increasing Reliability in the Midst of Rapid Growth,** Todd La Porte

88-10 **Issues in Rural and Small Development, Case Study: Watsonville, Santa Cruz County California,** Trish Ramos, Lakshmi Srinivas, Miriam Chion, Ana Lopez, Harry Hecht, Chris Broughton, Robert Murray

88-9 **White Reactions to Black Candidates: When Does Race Matter?** Jack Citrin, Donald Philip Green, David O. Sears

88-8 **Are Chicanos Assimilating?** Jorge Chapa

88-7 **California Agency Reconnaissance Project Reports,** Todd R. La Porte, David Hadwiger, Steven Stehr

88-6 **Do You Have To Be Crazy To Do This Job? Causes and Consequences of Job Satisfaction Among Local Legislators,** Edward L. Lascher, Jr.

88-5 **American All-Mail Balloting: A Summation of a Decade's Experience,** Randy H. Hamilton

88-4 **Corporate Campaign Spending and Initiative Outcomes in California,** Tom E. Thomas

88-3 **Research Applications: Perspectives on the California Seismic Safety Commission,** Stanley Scott

88-2 **Earthquake Engineering and Public Policy: Key Strategies for Seismic Policy,** Stanley Scott

88-1 **What Do Decision Models Tell Us About Information Use?** Evert A. Lindquist

1987

87-7 **The Politics of the AIDS Vaccine or How the California Legislature Searched for the Magic Bullet—And Wound Up Squabbling With the Trial Lawyers, the Budget-Cutters, and the Alzheimer's Establishment,** David L. Kirp and Hugh Maher

87-6 **The Reagan Presidency After Six Years,** Eugene C. Lee (moderator)

87-5 **A Critical Theory of Community,** Dennis J. Coyle

87-4 **The Reluctant Revival of Landowner Rights,** Dennis J. Coyle

87-3 **Informal Pluralism and LDP Guidance—Examination of Japan's Protectionism of Raw Silk Importation,** John Q. Zhao

87-2 **Towards a Typology of New Subnational Governmental Actors in International Relations,** Ivo D. Duchacek

87-1 **The Rocky Road to Privatization,** Lyle C. Fitch

SEND ORDER TO:
Institute of Governmental Studies
102 Moses Hall
University of California
Berkeley, CA 94720
(415) 642-5537

PLEASE PRE-PAY ALL ORDERS
UNDER $30: checks payable to The Regents of the University of California.
SALES TAX: California residents add sales tax.
HANDLING AND SHIPPING: add 20% of sales price. Allow 4 weeks for delivery.

Please add my name to the *PAR* mailing list (free). ____

TITLE QUANTITY/COST

SUBTOTAL
SALES TAX
HANDLING (20%)
TOTAL

NAME
ADDRESS

CITY STATE ZIP