

# SOCIAL DILEMMAS

◆324

Robyn M. Dawes<sup>1</sup>

Department of Psychology, University of Oregon, Eugene, Oregon 97403

---

## CONTENTS

INTRODUCTION TO THE LOGIC OF SOCIAL DILEMMAS .....	170
PROPOSALS FOR ELICITING COOPERATIVE BEHAVIOR .....	174
<i>Changing the Payoffs</i> .....	174
<i>From Payoffs to Utilities</i> .....	175
<i>Altruism</i> .....	176
<i>Conscience and norms</i> .....	177
THE MATHEMATICAL STRUCTURE OF DILEMMA GAMES .....	178
<i>The "Take Some" Game</i> .....	179
<i>The "Give Some" Game</i> .....	179
REVIEW OF THE LITERATURE ABOUT EXPERIMENTAL N-PERSON DILEMMA GAMES .....	182
<i>Findings</i> .....	183
<i>Involvement</i> .....	183
<i>Communication</i> .....	185
<i>Group size</i> .....	186
<i>Public disclosure of choice versus anonymity</i> .....	187
<i>Expectations about others' behavior</i> .....	187
<i>Moralizing</i> .....	188
A FINAL HYPOTHESIS ABOUT ELICITING COOPERATIVE BEHAVIOR....	188

Interest in social dilemmas—particularly those resulting from overpopulation, resource depletion, and pollution—has grown dramatically in the past 10 years among humanists, scientists, and philosophers. Such dilemmas are defined by two simple properties: (a) each individual receives a higher payoff for a socially defecting choice (e.g. having additional children, using all the energy available, polluting his or her neighbors) than for a socially cooperative choice, no matter what the other individuals in society do, but (b) all individuals are better off if all cooperate than if all defect. While

<sup>1</sup>This paper was written while I was a James McKeen Cattell Sabbatical Fellow at the Research Center for Group Dynamics at the Institute for Social Research at the University of Michigan and at the psychology department there. I thank these institutions for their assistance and especially all my friends there who helped.

many thinkers have simply pointed out that our most pressing societal problems result from such dilemmas, most have addressed themselves to the question of how to get people to cooperate. Answers have ranged from imposition of a dictatorship (Leviathan) to “mutual coercion mutually agreed upon,” to appeals to conscience.

This paper reviews the structure and ubiquity of social dilemma problems, outlines proposed “solutions,” and then surveys the contributions of psychologists who have studied dilemma behavior in the context of  $N$ -person games ( $N > 2$ ). The hypothesis that follows from this survey and review is that there are two crucial factors that lead people to cooperate in a social dilemma situation. First, people must “think about” and come to understand the nature of the dilemma, so that moral, normative, and altruistic concerns as well as external payoffs can influence behavior. Second, people must have some reason for believing that others will not defect, for while the difference in payoffs may always favor defection no matter what others do, the absolute payoff is higher if others cooperate than if they don’t. The efficacy of both factors—and indeed the possibility of cooperative behavior at all in a dilemma situation—is based upon rejecting the principle of “nonsatiety of economic greed” as an axiom of actual human behavior. And it is rejected.

## INTRODUCTION TO THE LOGIC OF SOCIAL DILEMMAS

Social dilemmas are characterized by two properties: (a) the social payoff to each individual for defecting behavior is higher than the payoff for cooperative behavior, regardless of what the other society members do, yet (b) all individuals in the society receive a lower payoff if all defect than if all cooperate.

Examples abound. People asked to keep their thermostats low to conserve energy are being asked to suffer from the cold without appreciably conserving the fuel supply by their individual sacrifices; yet if all keep their thermostats high, all may run out of fuel and freeze. During pollution alerts in Eugene, Oregon, residents are asked to ride bicycles or walk rather than to drive their cars. But each person is better off driving, because his or her car’s contribution to the pollution problem is negligible, while a choice to bicycle or walk yields the payoff of the drivers’ exhausts. Yet all the residents are worse off driving their cars and maintaining the pollution than they would be if all bicycled or walked. Soldiers who fight in a large battle can reasonably conclude that no matter what their comrades do they personally are better off taking no chances; yet if no one takes chances, the result will be a rout and slaughter worse for all the soldiers than is taking

chances. Or consider the position of a wage earner who is asked to use restraint in his or her salary demands. Doing so will hurt him or her and have a minute effect on the overall rate of inflation; yet if all fail to exercise restraint, the result is runaway inflation from which all will suffer. Women in India will almost certainly outlive their husbands, and for the vast majority who can't work, their only source of support in their old age is their male sons. Thus each individual woman achieves the highest social payoff by having as many children as possible. Yet the resulting overpopulation makes a social security or old-age benefit system impossible, so that all the women are worse off than they would have been if they had all practiced restraint in having children. Untenured assistant professors are best off publishing every article possible, no matter how mediocre or in how obscure a journal. (The deans' committees never actually read articles.) Yet the result is an explosion of dubious information and an expectation that anyone worthwhile will have published 10 or 15 articles within 5 years of obtaining a PhD, a result from which we all suffer (except those of us who own paper pulp mills).

Some of these examples come from the three crucial problems of the modern world: resource depletion, pollution, and overpopulation. In most societies, it is to each individual's advantage to use as much energy, to pollute as much, and to have as many children as possible.<sup>2</sup> (This statement should not be interpreted as meaning that these three phenomena are independent—far from it.) Yet the result is to exceed the "carrying capacity" (Hardin 1976) of "spaceship earth," an excess from which all people suffer, or will suffer eventually. These problems have arisen, of course, because the checks on energy use, pollution, and population that existed until a hundred years or so ago have been all but destroyed by modern technology—mainly industrial and medical. And use of new energy sources or new agricultural techniques for increasing harvests often exacerbate the problems (see Wade 1974a,b). While many societies throughout history have faced their members with social dilemmas, it is these dilemmas that are particularly global and pressing that have attracted the most attention among social thinkers (from an extraordinarily wide variety of fields).

Perhaps the most influential article published recently was Garrett Hardin's "Tragedy of the Commons," which appeared in *Science* in 1968. In it Hardin argued that modern humanity as the result of the ability to overpopulate and overuse resources faces a problem analogous to that faced by herdsmen using a common pasture (1968, p. 1244).

<sup>2</sup>People in affluent or in Communist societies do not contribute to world overpopulation, but in most societies in the world the payoff remains greatest for having as many children as possible.

As a rational being, each herdsman seeks to maximize his gain. Explicitly or implicitly, more or less consciously he asks, "What is the utility to *me* of adding one more animal to my herd?" This utility has one negative and one positive component.

1. The positive component is a function of the increment of one animal. Since the herdsman receives all the proceeds from the sale of the additional animal, the positive utility is nearly +1.

2. The negative component is a function of the additional overgrazing created by one more animal. Since, however, the effects of overgrazing are shared by all the herdsman, the negative utility for any particular decision-making herdsman is only a fraction of -1.

Adding together the component partial utilities, the rational herdsman concludes that the only sensible course for him to pursue is to add another animal to his herd. And another; and another . . . But this is the conclusion reached by every rational herdsman sharing the commons. Therein is the tragedy. Each man is locked into a system that compels him to increase his herd without limit—in a world that is limited. Ruin is the destination toward which all men rush, each pursuing his own best interest in a society that believes in the freedom of the commons.<sup>3,4</sup>

The gain-to-self harm-spread-out situation does indeed result in a social dilemma, although not all social dilemmas have that precise form (Dawes 1975).

Contrast Hardin's analysis of herdsmen rushing toward their own destruction with Adam Smith's (1776, 1976) analysis of the individual worker's unintended beneficence in a laissez-faire capitalistic society.

It is not from the benevolence of the butcher, the brewer, or the baker, that we expect our dinner, but from their regard to their own interest. We address ourselves, not to their humanity but to their self-love, and never talk to them of our own necessities but of their advantages (Book 1, p. 18).

As every individual, therefore, endeavors as much as he can both to employ his capital in the support of domestic industry, and so to direct that industry that its produce may be of the greatest value; every individual necessarily labors to render the annual revenue of the society as great as he can . . . By preferring the support of domestic to that of foreign industry, he intends his own security; and by directing that industry in such a manner as its produce may be for the greatest value, he intends only his own gain, and he is in this, as in many other cases, led by an invisible hand to promote an end which was no part of his intention (Book 4, p. 477).

<sup>3</sup>Actually, the negative payoff must be more negative than -1 for a true dilemma to exist. Hardin clearly implies a greater value when he discusses the destruction of the commons. If, for example, the commons can maintain 10,000 pounds of cattle when 10 1000-pound bulls are grazed on it, but only 9900 pounds when 11 bulls are grazed, then the herdsman who introduces an additional bull has two 900-pound bulls—a gain of 800 pounds over one 1000-pound one—while the total wealth of the commons has decreased by 100 pounds.

<sup>4</sup>Hardin uses the term "utility" to refer to social economic payoff. As will be emphasized in the next section of this article, there may be other utilities that determine behavior, so it does not follow from his analysis that "freedom in a commons brings ruin to all" (1968, p. 1244).

Hardin and Smith are not social theorists with diametrically opposed views about the effects of self-interested behavior. Rather, they are discussing different situations. Hardin's is a dilemma situation in which the external consequences of each herdsman's trying to maximize his profits are negative, and the negative consequences outweigh the positive ones to him. [Hardin specifically "exorcises" Smith's "invisible hand" in resolving population problems (p. 1244).] Smith's situation is a nondilemma one, in which maximizing individual profit does not hurt others more than it benefits the individual; in fact, it helps them. This difference is captured in the economic concept of an *externality* (Buchanan 1971, p. 7): "we can define an externality as being present whenever the behavior of a person affects the situation of other persons without the explicit agreement of that person or persons." In Hardin's commons the externalities are negative and greater than the individual's payoffs; in Smith's Scotland they are positive.

To define social dilemmas in terms of magnitudes of externalities would, however, involve interpersonal comparisons of payoffs. In most cases such a comparison is simple, but not in all. For example, it is difficult to compare the drivers' positive payoffs for driving during a pollution alert to the bike riders' negative payoffs for breathing polluted air. In contrast, the definition of a social dilemma proposed at the beginning of this paper involves payoff comparison only *within* an individual (who receives a 'higher payoff for defecting but whose payoff for universal defection is lower than that for universal cooperation). It is enough to note that most economic writing about negative externalities that has come to my attention has in fact been about dilemma situations.

Finally, Platt's (1973) concept of *social traps* is closely related to the concept of a dilemma. He defines a social trap as occurring when a behavior that results in immediate reward leads to long-term punishment. For example, many observers have noted that many modern technological advances may be traps; e.g. the good effects of DDT usage were immediately evident, while the disastrous effects took years to ascertain. Moreover, even when the long-term ill effects are known at the beginning, they may be "time discounted." ("If we're still around, we'll jump off that bridge when we come to it.") On an individual level, cigarette smoking, overeating, and excessive alcohol ingestion are traps. On the social level, most social dilemmas are social traps. But again not all—for dilemmas exist in which even defecting behavior is punished (because enough other people are bound to defect)—although not as badly as cooperative behavior would be. Further, not all social dilemmas involve a time lag.

We return then to the original definition of a social dilemma. Each individual receives a higher payoff for a socially defecting choice than for

a socially cooperative one, yet all individuals have a higher payoff if all cooperate than if all defect. All the examples discussed earlier meet these two conditions.

Given the ubiquity of social dilemmas—and the global importance of some of them—the question arises of how individuals and societies can deal with them. One answer is that they can't. The role of the social theorist is to point out where dilemmas exist and then to watch everyone defect—verifying the hypothesis that a social dilemma indeed is there. A far more common answer has been to propose mechanisms by which cooperation may be engendered in people facing social dilemmas.

## PROPOSALS FOR ELICITING COOPERATIVE BEHAVIOR

### *Changing the Payoffs*

Social dilemmas are defined in terms of the social payoff structure. The simplest proposal for eliciting cooperative behavior is to change that structure. That is, when analysis reveals that a social dilemma exists, an effort can be made to obliterate it by appropriate choices of rewards and punishments for cooperative and defecting behavior respectively. Then it is no longer a social dilemma.

The simplicity of this approach is appealing until we ask who will change the payoffs and how. The almost universal answer to the first question is government, and—somewhat surprisingly given the cultural background of the writers—the most common answer to the second question is: through coercion. Thus, for example, Hardin (1968, p. 1247) advocates “mutual coercion mutually agreed upon,” and Ophuls (1977) and Heilbroner (1974) advocate coercion from an authoritarian government in order to avoid the most pressing social dilemmas. These solutions are essentially the same as Hobbes's (1651) Leviathan, constructed to avoid the social dilemma of the “warre of all against all.” But there is *empirical* evidence that those societies where people are best off—currently at any rate—are those whose governments correspond least to Hobbes's authoritarian Leviathan (Orbell & Rutherford 1973). The counterargument (Robertson 1974) is that these societies are those that have been fortunate enough to have ample natural resources, or to have evolved from a more authoritarian state originating at a time when pressing social dilemmas did in fact exist. And if new dilemmas—in the form of overpopulation, pollution, and energy depletion—come as expected, Leviathan will again be necessary.

Most of us would prefer reward to coercion, although there are those who are willing to pay complex and expensive governmental bureaucracies to make sure that only the “deserving” achieve governmental rewards, rather

than to allow “giveaways.” The problem with both reward and coercion, however, is that they are very costly. The society faced with the potential dilemma must deplete its resources either to reward those tempted to defect, or to establish a policing authority that is sufficiently effective that those tempted will not dare do so. This depletion is paid by some or all society members. In effect, the dilemma has been turned into a new situation where everyone must cooperate but where the payoffs to everyone are less than they would be if everyone were to cooperate freely in the original situation. Sometimes, in fact, it is not even possible to avoid a dilemma by reward or coercion, because the costs of rewarding people for cooperating or effectively coercing them to do so exceed the gain the society derives from having everyone cooperate rather than defect.

Moreover, societal change in the payoffs by introducing rewards and punishments can be terribly inefficient. Consider, for example, the worker on a collective farm whose productivity is used in part to pay for a police agent whose job is to make sure that that worker does not sell the farm produce privately. Not only does that result in wasted productivity of the worker, but this police agent himself could instead be doing something productive for the society—such as working on the farm. Finally, coercive systems—and some governmental reward systems—apparently create, or at least exacerbate, a motivation to get around the rules.

### *From Payoffs to Utilities*

Many of us would not rob a bank, even if we knew that we could get away with it, and even if we could be assured that none of our friends or neighbors would know. Many of us give money to public television or to the United Fund, even though we know that our paltry contribution will make no difference in terms of the services rendered. Most of us take the trouble to vote, even though we know that the probability that an election will be decided by a single ballot is effectively zero. And some couples desiring a large family do in fact limit its size not out of desire but out of a belief that it is not moral to have too many children.

All these behaviors involve rejecting a payoff that is larger for one that is smaller. The potential bank robbers could be wealthy, the contributors could save their money, the voters could save themselves inconvenience, and the couples who want children could have them. The point is that the people making these decisions have *utilities* that determine their behavior, utilities associated with aspects of their behavior other than the external payoffs they would receive. The question of whether all behavior is “ultimately selfish” because it reflects *some* utilities is beside the point, just as the question of whether such selfishness is a primary human motivator is irrelevant to the question of whether society members facing a dilemma are

doomed to defect. The point is that if a person chooses action A over action B, then A must (by definition) have greater utility; if simultaneously action B provides a higher social payoff in terms of economic benefits or security, then (again by definition) other utilities must be guiding the individual's choice. The problem is to assess what these utilities are and to study their role in encouraging cooperative behavior.

Thus it is possible to have a social dilemma represented by a payoff structure and yet have people cooperate. The reason would be that the individuals' utilities do not present them with a dilemma. The utilities most important in eliciting cooperation are those associated with altruism, following social norms, and obeying dictates of conscience. These will be considered in turn.

**ALTRUISM** It is a demonstrable fact that people take account of others' payoffs as well as of their own in reaching decisions. Good Samaritans exist. (Whether this behavior is "ultimately selfish" in light of some hope of Heaven is again irrelevant.) Few of us would accept \$500 with nothing for our friend in lieu of \$498 for each of us. The importance of payoffs to others has been demonstrated experimentally by Messick and McClintock (Messick & McClintock 1968, Messick 1969, McClintock et al 1973)—albeit in some competitive experimental contexts where subjects apparently wish to minimize the payoffs to others, or at least to maximize the discrepancy between own and others' payoffs (Messick & Thorngate 1967).

The question is whether altruism can lead to cooperative behavior in the face of a social dilemma. If concern for others' payoffs is merely a *tactical* consideration for obtaining future rewards from that other, then utility for behaving altruistically cannot be counted upon as a factor that could outweigh external social payoffs. In most social dilemmas, individuals must behave privately, and the problem occurs because the social outcome results from the aggregate social behavior across a large number of people who do *not* interact. Thus, few people would be motivated to cooperate by tactical altruism.

Does altruism exist other than as a tactic? That question is difficult to answer experimentally, or on the basis of naturalistic observation, but it has been addressed recently by sociobiologists and others interested in the implications of evolutionary theory for modern human behavior. They do not agree about altruism. On the one hand, some see it as occurring in the face of natural genetic selection toward pure selfishness, because societies support the long-term reproductive success of altruists, even though altruistic behavior itself would be deleterious in a context outside the society. Thus, Campbell (1975), for example, believes in a "social evolution" toward altruistic and cooperative norms and morals, one that must be carefully



guarded against rampant individualism and the consequent genetic success of those most selfish. Blany (1976) notes that for whatever reason (the selfish interests of those shaping a society's beliefs?) women in all societies prefer men who are altruistic and brave to those who are self-centered and cowardly. So socially trained sexual preference may involve a social breeding of altruistic traits, again those that might not fare well in a "warre of all against all." Finally, Trivers (1971) proposes that altruism is a tactical advantage due to socially imposed norms of reciprocity.

In contrast, other sociobiologists hypothesize mechanisms by which altruism in and of itself may result in genetic propagation, even if not through direct propagation. Those proposing that such survival works through "group selection" ultimately benefiting the individual currently (1980) have few adherents. Many others (e.g. Alexander 1980) have proposed "kin altruism" as a plausible genetic link to all altruism. People share genes with their close relatives, and to the degree to which they—even in the celibate roles of priest and maiden aunt—help relatives survive, they enhance the probability that their own genes are propagated. Evidence for such kin altruism is most easily found in a mother's sacrifice for her children. Hence, to the degree to which altruistic concern is focused primarily on close kin ("nepotistic") and partially genetically based, it would be expected to increase through genetic selection. Whether such kin altruism would lead to a general altruistic concern for surrounding people, or for a whole tribe or society, is a moot question.

This literature does not provide a clear indication of whether altruism is purely tactical—nor does any other literature to my knowledge. Nevertheless, it may not be limited to tactical concerns, in which case it could be an important factor in leading people to cooperate in a social dilemma situation. There is one important proviso: people have to *know* about the payoffs to others if altruistic utilities are to be effective. This proviso is not trivial.

**CONSCIENCE AND NORMS** Even though conscience may often be only "the inner voice which warns us that someone may be looking" (Mencken, quoted in Cooke 1955), it has been a powerful force throughout history in motivating human behavior. People die for it. Tyrants use it to demand behavior of people that other people believe unconscionable. Desperate appeals are made to it—sometimes successfully—by potential victims of aggression.

Hardin (1968, pp. 1246–47) specifically dismisses appeals to conscience as a means of eliciting cooperative behaviors in social dilemmas. He first hypothesizes that such an appeal is a "double bind," because the person making the appeal may regard the person swayed as a "simpleton." Not necessarily so. For if the person making the appeal also has a regard for his

own “clear conscience” (perhaps as his “only sure reward”), then he is equally bound. A second argument of Hardin’s confuses morality with neurotic guilt and concludes that appeals to conscience are “psychologically pathogenic” (and may, like everything else, be misused by unscrupulous individuals).

But Hardin’s is the main discussion of appeals to conscience in the literature—or at least in the literature which has come to my attention. Psychologists, economists, political scientists, and sociobiologists do not tend to use “conscience” as an explanatory construct, perhaps because it is often considered secondary to other factors. But secondary or not, it does appear to have an important place in determining everyday behavior, and as one paper to be reviewed in the fourth section of this article suggests, it may be efficacious in eliciting cooperation.

Norms are somewhere between conscience and coercion. Most norms that exist may elicit punishment if violated. But norms have the ability to motivate people in the absence of any threat of censure. If we fight bravely because we are in Caesar’s Legions, it is true that we may be decimated if we do not. But it is not the fear of decimation that leads most of us to fight bravely. We fight because of what we are. Similarly, people may cooperate in social dilemmas because of what they are; they are not “the kind of people” who profit at others’ expense, or who contribute to a holocaust.

## THE MATHEMATICAL STRUCTURE OF DILEMMA GAMES

A game is simply a system of payoffs depending on the combination of choices made by the players. (An additional “choice” may be made by a random element that receives no payoff.) In dilemma games, each player makes one of two choices: D (for defecting) or C (for cooperating). The payoff to each player depends wholly on his or her choice of D or C and on the number of other players who choose C or D.

Let  $D(m)$  be the payoff to the defectors in an  $N$ -person game where  $m$  players cooperate, and let  $C(m)$  be the payoff to the cooperators when  $m$  players (including themselves) cooperate. A social dilemma game is characterized by two simple inequalities.

1.  $D(m) > C(m + 1)$

That is, the payoff when  $m$  other people cooperate is always higher for an individual who remains a defector than for one who becomes the  $m$  plus first cooperator ( $m$  goes from 0 to  $N - 1$ ).

2.  $D(0) < C(N)$

That is, universal cooperation among the  $N$  players leads to a greater payoff than does universal defection.

The statement of condition No. 1 in game theory language is that defection *is a dominating strategy*. But if everyone chooses that dominating strategy, the outcome that results is one that is less preferred by all players to at least one other (e.g. that resulting from universal cooperation). Since according to game theory all players should choose a dominating strategy, the result is termed an *equilibrium*. (No player would want to switch his or her choice.) Because the outcome dictated by the dominating strategy is less preferred by all players to the outcome of unanimous cooperation, this outcome is termed *deficient*. Hence, *a dilemma game is one in which all players have dominating strategies that result in a deficient equilibrium*. Two games developed for experimental research are illustrative.

### *The "Take Some" Game*

Each of three players simultaneously holds up a red or blue poker chip. Each player who holds up a red chip receives \$3.00 in payoff, but each of the three players *including that player* is fined \$1.00 for that choice. This is the negative externality. Each player who holds up a blue chip receives \$1.00 with no resultant fine. Three blue chips being held up provides a \$1.00 payoff to all players (and a social product of \$3.00) while three red chips being held up provides a zero payoff for all (and a zero social product). At the same time, however, each player reasons that he or she is best off holding up a red chip, because that increases the fines he or she must pay by only \$1.00 while increasing the immediate amount received by \$2.00 (\$3.00 – \$1.00). In effect, the player gets \$2.00 from the other two players' \$1.00 fines. In this game, one can *take some* from others. Such a choice is analogous to that involved in the decision to pollute (Dawes, Delay & Chaplin 1974).

### *The "Give Some" Game*

Each of five players may keep \$8.00 from the experimenter for himself or herself, or give \$3.00 from the experimenter to each of the other players. Again it is a dilemma because if all *give*, all get \$12.00 ( $4 \times \$3.00$ ) while if all *keep*, all get \$8.00; yet it is clearly in each player's individual interest to keep. In fact, each player is getting \$8.00 more by keeping than by giving. This game is based on the research of Bonacich (1972). The *give some* game presents the subjects with a choice analogous to that of deciding whether to contribute to a public good (Olsen 1965). (Each of us can reap the benefit of others' contributions while withholding ours.)

The "take some" and "give some" games can be presented in matrix form displaying the payoffs to defectors and cooperators as a function of the number of cooperators (Table 1).

**Table 1** Payoffs for the two games

The "Take Some" Game			The "Give Some" Game		
Number of cooperators	Payoffs to defectors	Payoffs to cooperators	Number of cooperators	Payoffs to defectors	Payoffs to cooperators
3	--	\$1.00	5	--	\$12.00
2	\$2.00	0	4	\$20.00	\$ 9.00
1	\$1.00	-\$1.00	3	\$17.00	\$ 6.00
0	0	--	2	\$14.00	\$ 3.00
			1	\$11.00	0
			0	\$ 8.00	--

In addition to properties 1 and 2 (above), the "take some" and the "give some" games have three further properties:

A.  $D(m + 1) - D(m) = c_1 > 0$

B.  $C(m + 1) - C(m) = c_2 > 0$

C.  $D(m) - C(m + 1) = c_3 > 0$

In the "take some" game,  $c_1 = \$1.00$ ,  $c_2 = \$1.00$ , and  $c_3 = \$1.00$ . In the "give some" game,  $c_1 = \$3.00$ ,  $c_2 = \$3.00$ , and  $c_3 = \$8.00$ .

If we were to plot the payoffs for defection and cooperation as a function of the number of cooperators, properties A and B state that both functions are straight lines with positive slopes (see Schelling 1973, Hamburger 1973). Property C states that these slopes are equal. Condition No. 1 (that an additional cooperator makes less than had he or she remained a defector) follows directly from property C, and condition No. 2 states that the right hand extreme of the cooperating function is above the left hand extreme of the defecting function.<sup>5</sup> Graphically, a social dilemma exists when the D payoff function is above the C function for its entire length and the right extremity of the C function is higher than the left extremity of the D function. It is apparent that a very wide range of configurations will meet this specification. Schelling (1973) has discussed many such configurations and has given a host of imaginative examples.

Hamburger (1973) has shown that dilemma games having properties A through C are equivalent to games in which each participant simultaneously

<sup>5</sup>Properties A and B do not imply property C unless  $c_1 = c_2$ , because it is possible that payoffs for cooperation and defection are linear but do not have equal slopes. On the other hand, properties A and C not only imply property B, but that  $c_1$  is equal to  $c_2$  as well. Properties B and C yield the same implication. Property C by itself has no implication other than condition No. 1, because it does not specify that the payoff functions need be straight lines.

plays identical two-person prisoners' dilemma games having property C (termed "separable" in the literature) against each of the remaining  $N - 1$  participants. Dawes (1975) has shown that they are also equivalent to the algebraic expression of the "commons dilemma" described by Hardin (1968). Figure 1 plots the payoffs for the "take some" game and the "give some" game respectively.

In the literature to be described here, most of the dilemma games have properties A-C. We shall term these *uniform games*, following Kahan (1973) and Goehring & Kahan (1976). One group of experimenters, working primarily at Arizona State University in the 1970s, uses much different games—those in which subjects may draw points from a pool that can "replenish itself" (i.e. be increased by the experimenter) at varying intervals in amounts depending upon the subjects' behavior (e.g. restraint or self-sacrifice). This paradigm, which defies a simple mathematical description, is similar to a card game devised by Rubenstein in his doctoral dissertation (cf Rubenstein et al 1975). Such games will be referred to as *variable*.

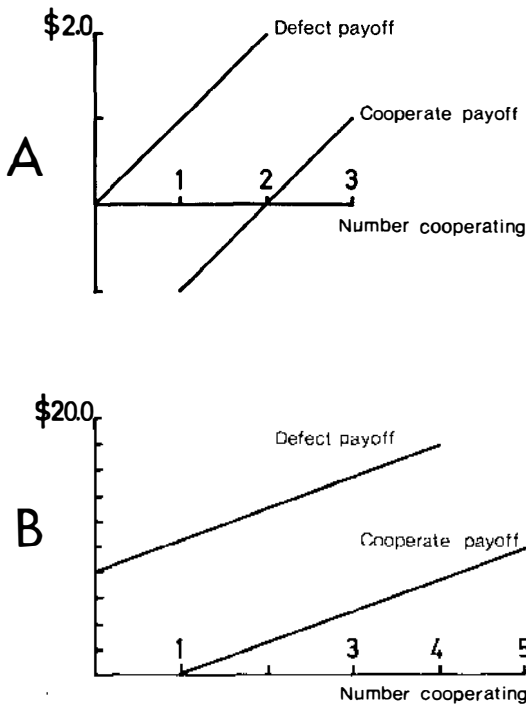


Figure 1 Graphs of payoffs for the two games.

## REVIEW OF THE LITERATURE ABOUT EXPERIMENTAL N-PERSON DILEMMA GAMES

The *prisoner's dilemma* is a two-person dilemma game. The name derives from an anecdote concerning two prisoners who have jointly committed a felony and who have been apprehended by a District Attorney who cannot prove their guilt. The District Attorney holds them incommunicado and offers each the chance to confess. If one confesses and the other doesn't, the one who confesses will go free while the other will receive a maximum sentence. If both confess they will both receive a moderate sentence, while if neither confesses both will receive a minimum sentence. In this situation, confession is a dominant strategy. (If the other confesses, confession leads to a moderate sentence rather than to a maximum one; if the other doesn't, it leads to freedom rather than to a minimum sentence.) But confession leads to a deficient equilibrium, because dual confession results in moderate sentences, whereas a minimum sentence could be achieved by neither confessing. Hence, the dilemma.

In the experimental gaming literature prisoner's dilemmas are often played repeatedly. That leads to an additional constraint on the payoffs so that the players cannot take turns playing the defecting strategy. (The sum of the payoffs for one defecting and one cooperating choice must be less than the sum for two cooperating choices.) Uniform dilemma games satisfy this constraint, but so do many others.

The overwhelming majority of experimental investigations of behavior in social dilemma games have studied subjects' responses in two-person prisoner's dilemmas that are played repeatedly by the same subjects (or by subjects who believe that they are playing against the same other subject—who may be a computer program). Payoffs for these two-person games have usually been in small amounts of money (e.g. cents); in virtually all experiments, subjects have been told that their purpose should be to maximize their own gain—although we suspect that many other motives such as maximizing relative gain (Messick & Thorngate 1967) or minimizing boredom may have been involved. There may well be over 1000 experiments reported in the psychological literature documenting how college students behave in such iterated prisoner's dilemmas.

The two-person iterated prisoner's dilemma has three characteristics, however, that make it unique—and hence unrepresentative of the social dilemmas discussed in this article.

1. In the two-person prisoner dilemma (iterated or not) all harm for defection is visited completely on the other player; harm is focused rather than spread out. In most social dilemmas in contrast, harm for defecting behavior is diffused over a considerable number of players.

2. In most social dilemmas defecting behavior may be anonymous; it is not necessarily so, but the possibility is there. In the two-person iterated game, in contrast, each player knows with certainty how the other has behaved. This necessary knowledge is unique to the two-person situation.

3. Each player has total reinforcement control over the other in the iterated two-person dilemma. That is, each player can “punish” the other for defection or cooperation (behavior that is socially optimal if individually suboptimal) by choosing defection on the subsequent choice, and can “reward” the previous choice of the other by choosing cooperation. Thus, each player can attempt to shape the other’s behavior by choice of defection and cooperation, while partially determining his or her own outcome by that same choice. The situation is very complicated. Each “game” is analogous to a play in chess which has meaning only within the metagame of the entire match. In fact, Amnon Rapoport (1967) has shown that if subjects really can influence each others’ subsequent choices, then the iterated prisoner’s dilemma isn’t a dilemma at all! So if subjects believe that they have such influence it is not a dilemma to them. This characteristic is unique to the two-person iterated dilemma; when there are more people involved it is not possible to attempt to shape a particular other person’s behavior by judicious (or believed to be judicious) choice of one’s own behavior. (There may be some element of such attempted shaping when the number of people involved approaches two—i.e. three or four—but the potential effectiveness of doing so is clearly diluted.)

Due to the specificity of harm, the lack of possible anonymity, and the potential use of one’s own behavior as a strategy to shape the other, two-person iterated prisoner’s dilemmas cannot be considered to be representative social dilemmas *in general*. The review of the literature and its findings that follow will be limited to investigations of dilemmas involving three or more people.

### *Findings*

**INVOLVEMENT** While any correlation between the “ecological validity” of an experiment and the degree of subject involvement is far from perfect, the assessment of such involvement is certainly an important factor in evaluating a domain of studies. When social dilemma games are played for substantial amounts of money, subjects are *extremely* involved. In 1972, Bonacich ran two conditions of 5-person “give some” games; in both conditions  $c_1 = c_2 = \$0.25$ ; in a “low temptation” condition  $c_3$  ranged from \$.01 to \$.20 across five trials, while in a “high temptation” condition it ran from \$.01 to \$.75, with a special trial at the end where subjects could win up to

\$16 by betraying their groups. In both conditions, communication was allowed, and the subjects made ample use of evaluative terms (“cheat,” “screw,” “greed,” “fink” being the four most common). In a later study (1976) Bonacich used larger amounts of money, which resulted in even more striking involvement. All subjects, in 5-person groups, played two games; in the first  $c_1 = c_2 = c_3 = \$0.30$ , while in the second, which was not a uniform game, any defection resulted in no payoff to cooperators and a payoff as high as \$9.00 to a single defector.

Bonacich writes (1976, p. 207):

During the coding of the tapes we noticed occasional joking threats about what the group would do to a noncooperator; he would not leave the place alive, they would push him down the stairs as he left, they would beat him up, they would write a letter to the student newspaper exposing his perfidy, or they would take him to small claims court. These threats could be intimidating and could suggest how angry the group would be toward the noncooperator.

Dawes, McTavish & Shaklee (1977) conducted an experiment involving even larger amounts of money; subjects played just once. Total cooperation resulted in \$2.50 for each member of their 8-person groups, total defection resulted in no payment to anyone,  $c_1 = c_2 = \$1.50$ , and  $c_3 = \$8.00$ , a substantial monetary incentive to defect. Some groups could communicate while others could not. Dawes, McTavish & Shaklee (p. 7) write:

One of the most significant aspects of this study, however, did not show up in the data analysis. It is the extreme seriousness with which the subjects take the problems. Comments such as, “If you defect on the rest of us, you’re going to have to live with it the rest of your life,” were not at all uncommon. Nor was it unusual for people to wish to leave by the back door, to claim that they did not wish to see the “sons of bitches” who double-crossed them, to become extremely angry at other subjects, or to become tearful. . . .

The affect level was so high that we are unwilling to run intact groups because of the effect the game might have on the members’ feelings about each other. The affect level also mitigates against examining choice visibility [NB in experiments involving high stakes]. In pretesting we did run one group in which choices were made public. The three defectors were the target of a great deal of hostility (“You have no idea how much you alienate me!” one cooperator shouted before storming out of the room); they remained after the experiment until all the cooperators were presumably long gone.

Experimenters whose payoffs consist of points to be converted to trivial amounts of cash or course credits do not report the affect level of their subjects. It may also be high, but I suspect that if it were it would be mentioned.

Whether or not high stakes and affect are necessary to reach valid conclusions about behavior in social dilemmas is a question that cannot be answered a priori, but depends in part upon a general finding of congruent or



disparate results across high involvement and low involvement studies. As yet there are not enough investigations in the field to know.

Certainly most of the dilemma situations in which we are interested involve high affect—e.g. that experienced by the author during the 1973 gasoline crisis as friend and neighbor after friend and neighbor finked out to become a “regular customer” of some service station.

**COMMUNICATION** The salutary effects of communication on cooperation are ubiquitous. In the first experiment by Bonacich reported above, communication was allowed in all groups, and 93% of the choices were cooperative. In the second experiment, there was a 94% cooperation rate. Bonacich did not run a no-communication control group (because he was not studying the effects of communication per se), but Dawes, McTavish, and Shaklee did. They found 72% cooperation in their communicating groups (which consisted of two different types to be described shortly) as opposed to 31% in their no-communication groups (which also consisted of two types).

Using points as payoffs, Rapoport et al (1962) and Bixenstine et al (1966) found that communicating groups cooperated more.<sup>6</sup> Using variable games with points taken from a replenishing pool, Brechner (1977), Edney & Harper (1978, 1979), and Harper (1977) all found that groups able to communicate cooperated more, with the result that more points were “harvested” from the pool.

Using a hypothetical uniform business game (in which manufacturers could cooperate against consumers), Jerdee & Rosen (1974) found that communication enhanced cooperation, but in a uniform game in which subjects “should act as if each point were worth \$1,” Caldwell (1976) did not. Caldwell did find, however, that a communication condition in which subjects could sanction defectors resulted in greater cooperation. Moreover, he found that communication per se did yield higher cooperation, although not significantly so, and as he wrote (p. 279), “Perhaps with real money subjects would be less inclined to treat the experiment as a competitive game.”

What is it about communication that leads to more cooperation? While most of the studies mentioned above simply pitted communication against

<sup>6</sup>These results require qualification. The communication that was effective in the Rapoport et al study was unintended; it occurred during a break between two 3–4 hour sessions, and because the experimenters’ (p. 40) “main interest was in the distribution of choices in the absence of communication;” the results after the break were ignored except for noting the high degree of cooperation. The game in the Bixenstine et al study was not strictly a dilemma, because there were some points at which defection did not dominate cooperation.

no communication, Dawes, McTavish, and Shaklee attempted to study the effects of various aspects of communication. They argued that there is a hierarchy of at least three aspects involved in any face-to-face communication about dilemma problems. First, subjects get to know each other as human beings (humanization); second, they get to discuss the dilemma with which they are faced (discussion); third, they have the opportunity to make commitments about their own behavior, and to attempt to elicit such commitments from others (commitment). Commitment entails discussion, and discussion in turn entails humanization. What Dawes, McTavish, and Shaklee did was to run four types of groups: those that couldn't communicate at all, those that communicated for 10 minutes about an irrelevant topic (they were asked to estimate the proportion of people at various income levels in Eugene, Oregon), those that could discuss the problem but couldn't ask for public commitments, and those that were required to "go around the table" and make public commitments after discussion. The first two types yielded cooperation rates of 30% and 32% respectively, while the last two had rates of 72% and 71%. Thus, humanization made no difference—at least not personal acquaintance based on a 10 minute discussion (the average amount of time that the discussion and commitment groups spent on the problem). Surprisingly, commitment made no difference, but it must be remembered that this commitment was one forced by the experimenters rather than one arising spontaneously from the group process. (Moreover, *every* subject promised to cooperate, which is the only reasonable statement to make no matter what one's intentions.)

**GROUP SIZE** All experimenters who have made explicit or implicit comparisons of dilemma games with varying number of players have concluded that subjects cooperate less in larger groups than in smaller ones. Rapoport et al (1962) and Bixenstine et al (1966) simply noted the low degree of cooperation in their three- and six-person games and stated that it is less than in comparable two-person prisoner's dilemmas. But they had no strict criterion of comparability. Marwell & Schmidt (1972) studied two- and three-person uniform games with  $c_3$  equal in each and found less cooperation in the three-person game. Unfortunately,  $c_1$  and  $c_2$  were not equated, being twice as large in the two-person as in the three-person game (which resulted in the "expected values" of cooperation and defection being identical *if* the other players were to respond in a 50–50 random manner). Harper et al (unpublished) compared one-, three-, and six-person groups in the variable dilemma involving pool replacement; they found cooperation decreased with group size, but it is not clear what the results were for a "one-person group" test—other than the intellectual ability of a single

individual to solve the replenishment problem in an optimal manner given the experimenter's replenishment rule.<sup>7</sup>

The problem is, of course, how to "equate"  $N$  and  $N'$  person dilemma games, or even whether such an equating is desirable (from the standpoint of "ecological validity"). Could it not be argued, for example, that the motive to defect (e.g.  $c_3$ ) should "naturally" increase with more players because the harm from defection—i.e. negative externality—should be diffused among more people?

The most careful job of equating we have found is in one game from a larger study by Bonacich et al (1976). These investigators set  $c_1$ ,  $c_2$ , and  $c_3$  equal in three-, six-, and nine-person games, and they discovered that cooperation decreased with increasing size (contrary to their theoretical expectations, which was that these parameters alone would determine rate of defection).

**PUBLIC DISCLOSURE OF CHOICE VERSUS ANONYMITY** Three studies have compared private with public choice (Bixenstine et al 1966, Jerdee & Rosen 1974, Fox & Guyer 1978); all found higher rates of cooperation when choice was public. While the difference between anonymity and public disclosure in these studies is not striking, they used minimal payoffs—and given the involvement obtained with significant amounts of money, we suspect that the difference would be much greater were the payoffs more significant.

**EXPECTATIONS ABOUT OTHERS' BEHAVIOR** There are three studies that collected subjects' expectations about whether others playing the games would cooperate or defect (Tyszka & Grzelak 1976, Dawes, McTavish & Shaklee 1977, Marwell & Ames 1979). There are two possible predictions. To the degree to which a subject believes others *won't* defect, he or she may feel it is possible to obtain a big payoff without hurting others too much. This desire to be a "free rider" [or "greed" as Coombs (1973) terms it] could result in a negative correlation between the propensity to cooperate and beliefs that others will. To the degree to which a subject believes that others *will* defect, he or she may feel that it is necessary to avoid a big loss by defecting himself or herself. The desire to "avoid being a sucker" [or "fear" as Coombs (1973) terms it] could result in a positive correlation

<sup>7</sup>Interestingly, there *is* an optimal solution for harvesting animals in their natural environment. Determine the maximal population size where there is no harvesting, and then keep the population at precisely half that size. See Dawes, Delay & Chaplin (1974) and Anderson (1974).

between the propensity to cooperate and beliefs that others will. In fact, all three studies report strong *positive* correlations. This finding is compatible with those reviewed by Pruitt & Kimmel (1977) in the area of iterated games.

There is one other interesting finding in the Dawes et al and the Tyszka and Grzelak studies. Defectors are more accurate at predicting cooperation rates than are cooperators. But Dawes, McTavish, and Shaklee found that they were *not* more accurate at predicting specifically who would cooperate and who would not. This apparent discrepancy between base rate accuracy and specific accuracy can be best understood by considering the predictions of the outcome of coin tosses. A person who predicts heads 50% of the time will be correct only 50% of the time despite a perfect base rate accuracy; a person who predicts heads 100% of the time will also be correct 50% of the time despite making the worst possible base rate prediction. In fact, in the Dawes, McTavish, and Shaklee study subjects were *very* poor at predicting who would and who would not cooperate.

**MORALIZING** Noting that the subjects in the Dawes, McTavish, and Shaklee study often raised moral issues in the discussion and commitment groups Dawes et al (unpublished) ran two experiments in which the experimenters themselves moralized at the subjects. These two studies, one conducted at Santa Barbara, California, and one conducted at Eugene, Oregon, contrasted a no-communication condition with a no-communication condition in which the experimenter delivered a 938 word sermon about group benefit, exploitation, whales, ethics, and so on. At both locations, the sermon worked—yielding rates of cooperation comparable to those found in the discussion and commitment groups of the earlier experiments. Of course, these sermons confounded logic, social pressure, experimental demand, emotional appeal, and so on.

## A FINAL HYPOTHESIS ABOUT ELICITING COOPERATIVE BEHAVIOR

The experiments reviewed in this article are lousy *simulations* of the social dilemmas with which most of us are concerned. In our current overpopulated world, the dilemmas of greatest import involve thousands to millions of people, large-scale communication or public disclosure is impossible, and most of the people choosing do not share the cultural background of American high school or college students. Findings about how small groups of such students behave in contrived situations cannot be generalized to statements about how to save the world (even though as part of our

own research-finding dilemma game, we often pretend that they can, thereby leading granting agencies to expect such statements).

What must be assumed is that the psychological and social factors that lead to defection or cooperation in small-scale dilemmas are roughly the same as those that influence behavior in large dilemmas. (Of course, small  $N$  dilemmas may be studied in their own right, in which case no such assumption is necessary.) This assumption cannot be based purely on the formal (i.e. mathematical) identity of small and large dilemmas. Rather, such an assumption must be based on broader theoretical ideas about human behavior—ideas that imply what might lead people to cooperate or defect in general, and which may then be tested in the small dilemma situation. Most of the studies reviewed in this article are based on such ideas. (Those, for example, that merely examine the effect of changing mathematical parameters in the experimental situation have been omitted.)

This distinction between experimental dilemmas as simulations and as hypotheses-testing devices is not just one of regard. For example, most simulation studies vary parameters of the dilemma itself (following the precedent of numerous iterated prisoner's dilemma studies); such studies are based on the assumption that these parameters (e.g. a mathematically defined "degree of conflict") have counterparts in the "real world," although it is difficult if not impossible to identify them with any precision. In contrast, those studies that investigate variables outside the structure of the game—e.g. communication, public disclosure, moralizing—vary these; such studies are based on the assumption that the experimental dilemma is (just) another "real" dilemma to the subjects, and that their behavior will be affected by these variables in the same way (more or less) as it would be affected in other dilemma situations.<sup>8</sup> And the expectation that these variables will affect behavior must always be based on some theoretical orientation or belief.

The analysis and literature reported thus far support a very simple theoretical proposition, one derived from extensive literature documenting that people have very limited abilities to process information on a conscious level, particularly social information. This ability is "limited" relative to what we naively believe; that is, study after study has shown a surprising inability to process information correctly on what appear to be the simplest tasks, provided they are not overlearned or automatic. The literature supporting this limited processing phenomenon is too vast to be referenced here without doubling the bibliography, but see Dawes (1976).

<sup>8</sup>I grant that it is always possible to attempt to construct a meta-game incorporating such variables, although their exact role and parameterization is extremely difficult to determine.

Such cognitive limitation may often result in an inability to understand or fully grasp the utilities in a social dilemma situation other than those that are most obvious, i.e. those connected with the payoffs. But it is precisely the payoff utilities that lead the players to defect, while the other utilities—e.g. those connected with altruisms, norms, and conscience—lead the players to cooperate. It follows that manipulations that enhance the salience and understanding of these utilities should increase cooperation. Communication (with or without commitment), public disclosure, and moralizing are precisely such manipulations.

Moreover, there are two additional studies—one mentioned briefly and one involving an iterated game—that support this hypothesis that greater knowledge yields greater cooperation. Marwell & Ames (1979) contacted high school students both by telephone and mail and asked them to invest a number of “tokens” supplied by the experimenter in either a “private” or “public” stock. The tokens invested in the private stock resulted in a fixed monetary yield per token. Those invested in the public stock resulted in a payoff to *all* members of the subject’s group (of 4 or 80 members whom the subject didn’t know); this payoff was an accelerating function of the number of people who invested their tokens in this public stock. The dilemma occurred because subjects received money from the public investment whether or not they personally contributed tokens to it. (It was not, however, strictly a dilemma situation, because if enough other group members invested in the public stock a “provision point” was reached, beyond which the public stock was also personally more rewarding than was the private stock.)

Marwell and Ames obtained a much higher rate of cooperation (public investment) than would be predicted from economic theory; their subjects were as much concerned with “fairness” as with monetary return. Why? The hypothesis proposed here suggests that the concern with the internal utility of fairness could have been brought about by the *length of time* the subjects had to consider their choice. They had a minimum of 3 days. (The time in the typical no-communication experiment is 10 minutes.) It follows that they had time to think about factors other than the external payoffs—e.g. to think about “fairness.” Note that this study was done under a condition of total anonymity, a factor most common in large-scale social dilemmas.

The other study supporting the general hypothesis presented here is that of Kelley & Grzelak (1972). When interviewing subjects who had played an iterated social dilemma game in groups of 13 subjects, these investigators found that subjects who had made a (relatively) high proportion of cooperative responses were better able to identify the response best for the group than were those who made a low proportion of cooperative responses. While

the hypothesis stated here is the converse of that finding, direction (not magnitude) of statistical association is symmetric.

Is knowledge *all* that is necessary? No, for while utilities associated with altruism, norms, and conscience may be made salient by knowledge, they do not necessarily overwhelm those associated with the payoffs. Repugnant as it may be from a normative point of view, moral and monetary (or survival) utilities combine in a compensatory fashion for most people.

He: Lady, would you sleep with me for 100,000 pounds?

She: Why, yes. Of course.

He: Would you sleep with me for 10 shillings?

She: (angrily) What do you think I am, a prostitute?

He: We have already established that fact, madam. What we are hag-  
gling about is the price.

Everyone may not have his or her price, but it does not require a systematic survey to establish that most people in the world will compromise his or her altruistic or ethical values for money or survival. Thus, the negative payoffs for cooperative behavior must not be too severe if people are to cooperate. It may be for precisely this reason that the expectation that others will cooperate is so highly correlated with cooperation itself. If others cooperate, then the expected payoff for cooperation is not too low, even though—in a uniform game, for example—the *difference* between the payoff for cooperation and that for defection is still quite large. People may be greedy, may prefer more to less, but their greed is not “insatiable” when other utilities are involved.

Thus, three important ingredients for enhancing cooperation in social dilemma situations may be: knowledge, morality, and trust. These ancient virtues were not discovered by the author—or by the United States Government, which invested millions of dollars in research grants over the years to have subjects play experimental games. But the above analysis indicates that they may be the particular virtues relevant to the noncoercive (and hence efficient) resolution of the social dilemmas we face.

## Literature Cited

- Alexander, R. D. 1980. *Darwinism and Human Affairs*. Seattle: Univ. Washington Press
- Anderson, J. M. 1974. A model for the "tragedy of the commons." *IEEE Trans. Syst. Man Cybern.* pp. 103-5
- Bixenstine, V. E., Levitt, C. A., Wilson, K. R. 1966. Collaboration among six persons in a prisoner's dilemma game. *J. Conflict Resolut.* 10:488-96
- Blany, P. H. 1976. Genetic basis of behavior—especially of altruism. *Am. Psychol.* 31:358
- Bonacich, P. 1972. Norms and cohesion as adaptive responses to political conflict: an experimental study. *Sociometry* 35:57-75
- Bonacich, P. 1976. Secrecy and solidarity. *Sociometry* 39:200-8
- Bonacich, P., Shure, G. H., Kahan, J. P., Meeker, R. J. 1976. Cooperation and group size in the N-person prisoner's dilemma. *J. Conflict Resolut.* 20:687-705
- Brechner, K. C. 1977. An experimental analysis of social traps. *J. Exp. Soc. Psychol.* 13:552-64
- Buchanan, J. M. 1971. *The Bases for Collective Action*. New York: Gen. Learn. Press
- Caldwell, M. D. 1976. Communication and sex effects in a five-person prisoner's dilemma. *J. Pers. Soc. Psychol.* 33: 273-81
- Campbell, D. T. 1975. On the conflict between biological and social evolution and between psychology and the moral tradition. *Am. Psychol.* 30:1103-26
- Cooke, A. 1955. *The Vintage Mencken*, p. 231. New York: Vantage
- Coombs, C. A. 1973. A reparameterization of the prisoner's dilemma game. *Behav. Sci.* 18:424-28
- Dawes, R. M. 1975. Formal models of dilemmas in social decision-making. In *Human Judgment and Decision Processes*, ed. M. F. Kaplan, S. Schwartz, pp. 88-107. New York: Academic
- Dawes, R. M. 1976. Shallow psychology. In *Cognitions and Social Behavior*, ed. J. Carroll, J. Payne, pp. 3-12. Hillsdale NJ: Erlbaum
- Dawes, R. M., Delay, J., Chaplin, W. 1974. The decision to pollute. *Environment and Planning*, pp. 2-10
- Dawes, R. M., McTavish, J., Shaklee, H. 1977. Behavior, communication and assumptions about other people's behavior in a commons dilemma situation. *J. Pers. Soc. Psychol.* 35:1-11
- Dawes, R. M., Shaklee, H., Talarowski, F. On getting people to cooperate when facing a social dilemma: moralizing helps. Unpublished manuscript
- Edney, J. J., Harper, C. S. 1978. The effects of information in a resource management problem: A social trap analog. *Hum. Ecol.* 6:387-95
- Edney, J. J., Harper, C. S. 1979. Heroism in a resource crisis: a simulation study. *Environmental Management*. In press
- Fox, J., Guyer, M. 1978. "Public" choice and cooperation in n-person prisoner's dilemma. *J. Conflict Resolut.* 22:468-81
- Goehring, D. J., Kahan, J. P. 1976. The uniform n-person prisoner's dilemma game. *J. Conflict Resolut.* 20:111-28
- Hamburger, H. 1973. N-person prisoners dilemmas. *J. Math. Sociol.* 3:27-48
- Hardin, G. R. 1968. The tragedy of the commons. *Science* 162:1243-48
- Hardin, G. R. 1976. Carrying capacity as an ethical concept. *Soundings: Interdiscip. J.* 59:121-37
- Harper, C. S. 1977. Competition and cooperation in a resource management task: a social trap analogue. In *Priorities for Environmental Design Research*, ed S. Weidman, J. R. Anderson, pp. 305-12. Washington DC: Environ. Res. Assoc.
- Harper, C. S., Gregory, W. L., Edney, J. J., Lindner, D. Group size effects in a simulated commons dilemma. Unpublished manuscript
- Heilbroner, R. 1974. *An Inquiry into the Human Prospect*. New York: Norton
- Hobbes, T. 1651, 1947. *Leviathan*. London: Dent
- Jerdee, T. H., Rosen, B. 1974. Effects of opportunity to communicate and visibility of individual decisions on behavior in the common interest. *J. Appl. Psychol.* 59:712-16
- Kahan, J. P. 1973. Noninteraction in an anonymous three-person prisoner's dilemma game. *Behav. Sci.* 18:124-27
- Kelley, H. H., Grzelak, J. 1972. Conflict between individual and common interest in an n-person relationship. *J. Pers. Soc. Psychol.* 21:190-97
- Marwell, G., Ames, R. E. 1979. Experiments on the provision of public goods I: resources, interest, group size, and the free rider problem. *Am. J. Sociol.* 84: 1335-60
- Marwell, G., Schmidt, D. R. 1972. Cooperation in a three-person prisoner's dilemma. *J. Pers. Soc. Psychol.* 31: 376-83
- McClintock, C. G., Messick, D. M., Kuhlman, D. M., Campos, F. T. 1973. Moti-



- vational bases of choice in three-choice decomposed games. *J. Exp. Soc. Psychol.* 9:572-90
- Messick, D. M. 1969-1970. Some thoughts on the nature of human competition. *Hypothese: Tijdschr. Psychol. Opvoedkunde.* 14:38-52
- Messick, D. M., McClintock, C. G. 1968. Motivational bases of choice in experimental games. *J. Exp. Soc. Psychol.* 4:1-25
- Messick, D. M., Thorngate, W. B. 1967. Relative gain maximization in experimental games. *J. Exp. Soc. Psychol.* 3: 85-101
- Olsen, M. 1965. *The Logic of Collective Action.* Cambridge, Mass: Harvard Press
- Ophuls, W. 1977. *Ecology and the Politics of Scarcity.* San Francisco: Freeman
- Orbell, J. M., Rutherford, B. 1973. Can Leviathan make the life of man less solitary, poor, nasty, brutish, and short? *Br. J. Polit. Sci.* 3:383-407
- Platt, G. 1973. Social traps. *Am. Psychol.* 28:641-51
- Pruitt, D. G., Kimmel, M. J. 1977. Twenty years of experimental gaming: critique, synthesis, and suggestions for the future. *Ann. Rev. Psychol.* 28:363-92
- Rapoport, Amnon. 1967. Optimal policies for the prisoner's dilemma. *Psychol. Rev.* 74:136-48
- Rapoport, Anatol, Chammah, A., Dwyer, J., Gyr, J. 1962. Three-person non-zero-sum nonnegotiable games. *Behav. Sci.* 7:30-58
- Robertson, D. 1974. Well, does Leviathan...? *Br. J. Polit. Sci.* 4:245-50
- Rubinstein, F. D., Watzke, G., Doctor, R. H., Dana, J. 1975. The effect of two incentive schemes upon the conservation of shared resource by five-person groups. *Organ. Behav. Hum. Perform.* 13:330-38
- Schelling, T. C. 1973. Hockey helmets, concealed weapons, and daylight saving: a study of binary choices with externalities. *J. Conflict Resolut.* 17:381-428
- Smith, A. 1776, 1976. *The Wealth of Nations.* Chicago: Univ. Chicago Press
- Trivers, R. L. 1971. The evolution of reciprocal altruism. *Q. Rev. Biol.* 46:35-57
- Tyszka, T., Grzelak, J. L. 1976. Criteria of choice in non-constant zero-sum games. *J. Conflict Resolut.* 20:357-76
- Wade, N. 1974a. Sahelian drought: No victory for Western aid. *Science* 185: 234-37
- Wade, N. 1974b. Green revolution (1): a just technology, often unjust in use. *Science* 186:1094-96