Gemeinschaftsgüter: Recht, Politik und Ökonomie

Preprints
aus der Max-Planck-Projektgruppe
Recht der Gemeinschaftsgüter
Bonn
2002/11

A Game Theoretic Taxonomy of Social Dilemmas

von
Martin Beckenkamp

A game theoretic taxonomy of social dilemmas[1]

Abstract

Both social psychology and experimental economics empirically investigate social dilemmas. However, these two disciplines sometimes use different notions for very similar scenarios. While it is irrelevant for economists whether an experimental public-good game is conceptualised as a take-some or give-some game - i.e. whether something is conceptualised as produced or extracted - it is not irrelevant for some psychologists: They grasp public-good games as "give-some" games. And, whereas most economists define social dilemmas in reference to a taxonomy of goods, some psychologists think that dominant strategies are a necessary attribute. This paper presents a taxonomy that relies on a formal game-theoretic analysis of social dilemmas, which integrates and clarifies both approaches. Because this taxonomy focuses on the underlying incentive structure, it facilitates the evaluation of experimental results from both social psychology and experimental economics.

Motivation or "Why introduce another distinction"?

Diverse views of social dilemmas have been suggested in recent years. Unfortunately, some of these definitions are incompatible with each other. But some consensus nonetheless exists in the different fields of research nonetheless. One reason for this might lie in the fact that there are now conceptualisations of social dilemmas that integrate numerous different views under the umbrella of one very general perspective. One such general definition is offered by Liebrand, van Lange and Messick (1996, p. 546): "Social dilemmas are characterised as situations in which private interests are at odds with the collective interests." These authors point out that the existence of a dominant strategy is not an essential feature of social dilemmas. Their definition thus stands in contrast to many popular definitions presented in the literature, such as those found in Dawes (1991) or Yamagishi (1995).[2] While making dominant strategies integral to the definition of social dilemmas is widespread, it nevertheless leads to a definition which is too narrow to grasp all interesting forms of social dilemmas. Not even Hardin's famous illustration of the tragedy of the commons implies a dominant strategy, i.e. a strategy "which yields the person the best payoff in *all* circumstances" (Liebrand, van Lange and Messick 1996, p. 547).

1 I'm very grateful to the VW-foundation who financed our research project "Protection Policies Using Incentives". The considerations offered in this paper have been developed in the context of this project. I am also very grateful to Axel Ostmann, who was a tremendous support for this paper and in our common work. This paper contains excerpts from my „Habilitationsschrift" (Beckenkamp 2001) at the University of Saarland, Germany.

2 Dominant strategies yield the person the best payoff in all circumstances, i.e. the best response is always the same option in all circumstances. The example below shows that this requirement is too strict to apply to all social dilemmas.

Important concepts in defining social dilemmas

In order to introduce some important concepts and to clarify why dominant strategies are not essential in defining social dilemmas, I shall here refer to a constructed example concerning a fishery. This example presents a social dilemma which is rich enough to introduce and to distinguish between relevant game-theoretic concepts that characterise a social dilemma. In this constellation let us assume that the payoff (that is, the yield minus costs) for all the participants taken together simply depends on the number of cutters on the sea. The underlying conflict structure between the interests of the individual and the interests of the group is more complicated than the structure of prisoners' dilemmas. The advantage of using this example instead of the prisoner's dilemma is that it entails a little more realistic view of certain environmental problems, specifically the way in which the payoff of the group depends on the total contribution. For the sake of clarity, this illustration ignores iterated games, time dependencies, and the dependence of prices on supply and demand.

Let's assume the following scenario: In a fishing village each fishermen has the opportunity to fish with up to four cutters. The following figure shows the resulting payoffs in dependence on the total number of cutters used.

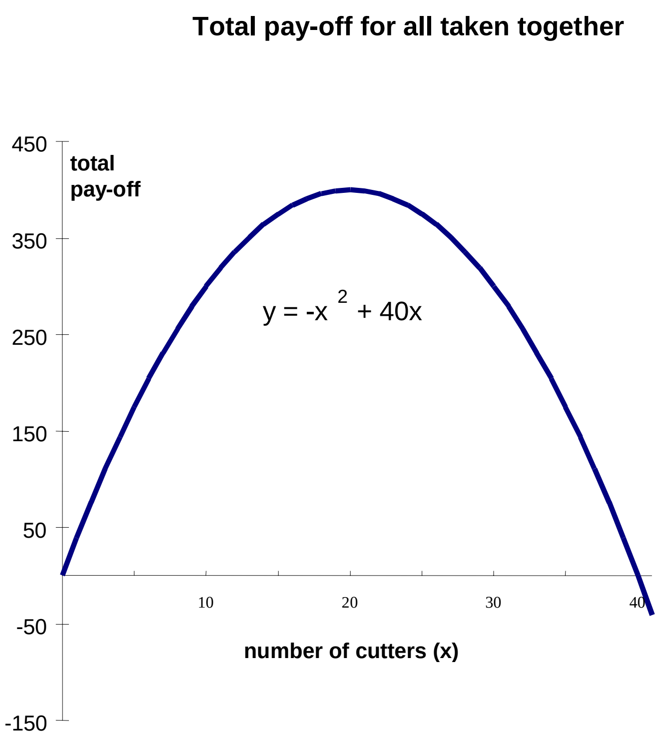

Figure 1: An example for a common payoff function (i.e. total payoff) that yields a social dilemma without the existence of a dominant strategy.

The maximum total or common payoff[3] measured in an arbitrary currency is reached with twenty cutters. 400 units of the currency is the resulting payoff in this example. These 400 units are divided among the fishermen, depending on the number of cutters they use in relation to the total number of cutters used. For example, if a fisherman uses four cutters and the total number of cutters used is twenty, his payoff is 4/20 * 400 units = 80 units. The following table gives the yields from the fishermen's point of view. Additionally, in the rightmost column the total payoff for all fishermen can be found again (corresponding to the parabola in figure 1). The grey cells in the table present the optimal decision of the fishermen, given the total number of cutters, as presented in the leftmost column.

Table 1: An example: an individual payoff table of a social dilemma without dominant strategies (compare also figure 1)

Total sum of cutters used	Number of cutters used from a particular fisherman and their respective payoffs					
	0	1	2	3	4	Total Payoff
.....						
19	0,0	21,0	42,0	63,0	84,0	399
20	0,0	20,0	**40,0**	60,0	80,0	400
21	[illegible]	[illegible]	38,0	57,0	76,0	399
22	[illegible]	[illegible]	36,0	54,0	72,0	396
....						
29	0,0	11,0	22,0	33,0	44,0	319
30	0,0	10,0	20,0	30,0	40,0	300
31	0,0	9,0	18,0	27,0	36,0	279
32	0,0	8,0	16,0	24,0	32,0	256
....						
36	0,0	4,0	8,0	12,0	16,0	144
37	0,0	3,0	6,0	9,0	12,0	111
38	0,0	2,0	4,0	6,0	8,0	76
39	0,0	1,0	2,0	3,0	4,0	39
40	0,0	0,0	0,0	0,0	0,0	0

Symmetric *Welfare-Optimum* with (2,2,2,2,2,2,2,2,2,2)

Asymmetric *Equilibrium* (stable responses) e.g. with (4,4,4,4,4,4,4,4,3,2) (sum 37)

3 Depending on the context, in the following, I shall sometimes refer to the total payoff and sometimes to the common payoff function, the former providing a number and the latter providing the whole payoff function relevant for the group.

Let us assume there are ten fishermen in the village. In such a case it would be optimal for the village (respectively the group or the "national economy") if a total of 20 cutters were be used. The *welfare-optimum*[4] is achieved when a total of 20 cutters is used. The resulting total payoff would then be 400 units. There are many possible distributions for reaching this optimal outcome; for one, it could be achieved if each fishermen used two cutters. This would also correspond to an equity norm principle, because the total catch would be distributed in equal shares to each fisherman. In this case the yield of each fisherman would be 40 units.

But one of the ten fisherman could have the following consideration: "Actually I should send one more cutter out on the sea. Instead of 20 cutters, the total number would then be 21. But then I would get 57 units instead of 40. And if I sent all four of my cutters out, there would be a total of 22 cutters and my payoff would be 72 units." It can easily be comprehended that this consideration is not devious at all if one puts oneself in the fisherman's position, who has to support his family and himself from his revenues. But the fatal fact is of course, that each of the ten fishermen in our example could search for such a "*best response*" . Further, the decision is rational in that it maximizes one's own payoff. In table 1 the best responses are tagged by grey shades: These should not be confused with the cell tagged by bold font and grey shades, which shows the welfare-optimum. These latter cell does not represent the best response.

If each of the fishermen acts on the basis of considerations of his own best response, the best responses may result in an *equilibrium*. In our example there are three different equilibria, two with a total sum of 37 cutters and one with a total sum of 36 cutters. In the first case there is equilibrium when eight fishermen use 4 cutters, one fisherman uses 3 cutters and one fisherman uses 2 cutters, or when seven fishermen use 4 cutters and three fishermen use 3 cutters. In the second case there is an equilibrium when six fishermen use 4 cutters and four fishermen use 3 cutters.[5] In an equilibrium there is no incentive for any of the eight fishermen to make another choice, that is, to increase or to reduce the number of cutters he uses, because using 4, 3 and 2 cutters is the best response when there are a total of 37 cutters, and the use of 4 and 3 cutters is the best response when there are a total of 36. The lack of incentives for fishermen to reduce the number of cutters they use may bring about harmful environmental consequences. In other words, the consequence of such an equilibrium is that, as long as the others abide by their choices, a new choice by a single fisherman would result in lower payoffs for that fisherman, independent of the direction of change (i.e. increase or decrease of cutters used).

This example also illustrates the severity of the conflict between mixed motives in a social dilemma. If the fishermen co-operated and came to an agreement to settle on an optimal total catch, in which a total of 20 cutters were used, each fisherman would gain 40 units — so long as

4 This is the economic term for pareto-optimal solutions. In the following, other notions from game-theory and economics will be introduced. These notions will be tagged by italics.

5 In this example the decisions of the fisherman correspond to integer values (e.g. a decision to send 1.58 cutters cannot be made). This is why there is more than one equilibrium. The consequence is that this example contains aspects of two different games: a social-dilemma game and a coordination game like "The Battle of Sexes". In this case it may be difficult to hit the equilibrium accurately. Nevertheless it can be expected that the total crop empirically found is near the total crops given by the equilibria.

the agreement consisted in an equal division between each fisherman, with each using 2 cutters. In an equilibrium, in which there are a total of 36 cutters, the gain would lie between 12 and 16 units for each fisherman. But, as already mentioned above, when there is an optimised total catch each fishermen is tempted to defect and to use 4 instead of 2 cutters, thereby earning 72 instead of 40 units. In the case under discussion the other fishermen (using 2 cutters) would earn 36 units. Economists speak about pecuniary *externalities* in such cases: these are costs caused by one individual, which have to be borne by the others. In this case, each of the other fishermen, would have to bear the costs of 4 units. The total costs would be 9*4 = 36 units. The gain by the defecting fisherman (32 units) is less than the total costs caused by him. This is why economists think that social dilemmas are a very delicate problems, and refer to them as instances of a "*market failure*". Thus the virtues of the market's "invisible hand", which were postulated by Adam Smith (1776, new edition 1988), is at least challenged by common goods and public goods.

One function of law is to define norms and rights such that defection becomes irrational. One problem with regulating fishing is that property rights are missing, and it seems impossible to introduce property rights. (This is at least true of offshore fish.) However, in many cases norms and rights can stabilise cooperative solutions. If this happens, it can even be quite rational to tolerate some defection at the margin, as long as there is reason to believe that others will not defect. This is especially clear if the possible consequences of defection and the breaching the norm seem so unpleasant that they override the advantages of a defection.

Are time lags an essential feature of social dilemmas?

Instead of focussing on conflicts between private and collective interests, in accord with the methodology of behaviorism another definition of social dilemmas has been given. A social dilemma is defined as a kind of a "social trap" or "social fence":

> The term [social trap M. B.] refers to situations in society that contain traps formally like a fish trap, where men or organizations or whole societies get themselves started in some direction or some set of relationships that later prove to be unpleasant or lethal and that they see no easy way to back out [of] or avoid.
>
> Two recent descriptions of traps of this kind have already been widely quoted and discussed. The first is Garrett Hardin's (1968) article... A converse type of situation might still be regarded as a generalized trap, but perhaps is more accurately called a countertrap. The consideration of individual advantage prevents us from doing something that might nevertheless be of great benefit to the group as a whole. It is, so to speak, a social fence rather than a social trap (Platt 1973, S. 641).

Social traps are conceptualised as reinforcement plans, immediately reinforcing long-term noxious behaviour. On the other hand, overcoming social fences requires behaviour that entails short-term aversive consequences, but that, in the long run, leads to positive consequences for oneself and the community.

Researchers are aware of social dilemmas that can be characterised as social traps. Nevertheless, good reasons can be given for why a conceptual distinction makes sense. If social dilemmas were defined as social traps, then they would only account for situations in which the positive and negative consequences of a decision are spread over different points of time. However, this is not an essential attribute of many situations typically characterized as social dilemmas. For example, there are a lot of experimental studies on prisoners' dilemmas which do not provide for such time lags.

Besides this temporal factor, many other factors have been suggested as distinguishing different social dilemmas: for example, the distinction between take-some and give-some games. Although changes in surface features[6] may lead to structural changes – as happens, for example, when the size of a group changes – such categorizations take the surface features into consideration rather than the characteristics of the underlying formal structure. This is particularly clear in numerous categorizations of social dilemmas suggested in the field of social psychology. Economic categorizations differ from these because they are orientated towards differences in the incentive structure instead. The conceptual distinction between public goods and commons is relevant in this context.

A useful way of defining social dilemmas

In various publications by Liebrand et al., a definition of social dilemmas is offered in which neither the existence of time lags nor the existence of dominant strategies is an essential feature. Instead they point to a different common feature in the underlying incentive structure of social dilemmas:

> …[S]ocial dilemmas can be defined as situations in which each decision-maker is best off acting in his own self-interest, regardless of what the other persons do. Each self-interested decision, however, creates a negative outcome or cost for the other people who are involved. When a large number of people make the self-interested choice, the costs or negative outcomes accumulate, creating a situation in which everybody would have done better had they decided not to act in their own private interest (van Lange, Liebrand, Messick & Wilke 1992, p. 4).

Whereas in this definition one might get the impression that time lags also characterize social dilemmas, the following definition is non-ambiguous with respect to this point:

> Social dilemmas are characterized as situations in which private interests are at odds with the collective interests (Liebrand, van Lange & Messick 1996; p. 546).

Based on this definition, which focuses on the incentive structure, a taxonomy of different social dilemmas will be presented in the following section. This taxonomy, based on a fundamental idea of Axel Ostmann (cf. Beckenkamp & Ostmann 1996), has been advanced in close co-

6 The term "surface features" is used in the sense of Gentner (1983). The idea is that the same abstract formal structure may be embedded in different semantics, so that surfaces similarities can be distinguished from structural similarities.

operation with him. From our point of view, in the context of the empirical research on social dilemmas, this categorization can be much more useful than distinctions that only consider surface features. Therewith the psychological virtue of surface features is not denied at all.The point is that in the analysis and presentation of the results, it is imperative that there be clarity about whether a particular case of experimental manipulation exclusively concerns surface features or whether it also concerns changes in the incentive structure.

In other words, the incentive structure defines potential conflicts in (monetary) interests, which may not become manifest. Thus an interesting research question ensues: namely, when do incentives matter? In our example in the village under discussion it may be a general habit for each fisherman to use two cutters. Presuming this contexts, let’s say that one day a fisherman covertly introduces a third cutter. Given the established routine, it cannot be expected that the other fishermen will also defect so long as the reduction in the catch of the others is not perceived as a surprise, but rather as a random variation. Thus the formal structure of interests may give the necessary condition that a manifest conflict may occur, although it is not a sufficient one.

A formal taxonomy system of social dilemmas

A general formal characterization of social dilemmas

The investigation of experimental social dilemmas is usually in a setting where subjects can earn different amounts of money (“payoffs”), the amounts depending on their decision and the decision of the others in the same group. In game theory, such payoffs to subjects can be formally represented by a normal form game with n players. Such normal-form games of social dilemmas can be created in by first specifying a general payoff function. The payoff of a subject is given by:

equation 1: $u(x, y) = b - c * x + q * f(x + y)$; with $c > 0$ and $x \in [0 \ldots 1]$.

x: the investment of the subject (“focus subject”) whose payoff is calculated by this formula. In simple cases, like the classical prisoners’ dilemma, this investment is either 0 (defect) or 1 (co-operate). In more complicated cases, where any investment between 0 and a maximal endowment **e** is allowed, x can be re-scaled in the interval [0..1] by interpreting x as the relative amount of the endowment .

y: the sum of the investments of all the other n-1 subjects of the group (that is the total sum of investments minus the investment of the focus subject).

u(x,y): the payoff that the focus subject gets, dependent on his investment and the investment of the (n-1) other subjects.

f(x + y): the total payoff of all subjects taken together. In most social dilemmas that have been experimentally investigated, this function is either a linear or quadratic polynomial. Social dilemmas with a threshold are also often investigated, mostly in “public-good” experiments.

c * x: the resulting costs of an investment *x* for the focus subject.

q: the proportioning rule (“quota”) specifying how the common payoff f(x + y) is divided up amongst the subjects. The concrete specification of the rule may depend on x, y and n.

b: A “baseline” payoff that the subjects receive independently of their investment *x* and the investment *y* of the others. This baseline can also be negative, leading to costs when the whole group defects.

Using this function, different social dilemmas can now be specified. In specifying them it is possible to categorize most of the social dilemmas experimentally investigated by merely introducing possible differences with respect both to the quota and the payoff function. Social dilemmas with fixed quotas can be distinguished from social dilemmas with proportional quotas, and social dilemmas with linear payoff functions can be distinguished from social dilemmas with quadratic payoff functions. Introducing a third type of payoff function with thresholds, also makes it possible to also integrate public goods with thresholds into this scheme. (In a certain sense, public goods with thresholds are not social dilemmas, as will be shown in detail below. This is because there is an efficient and an inefficient equilibrium in such cases.)

Is the distinction between “take-some” and “give-some” relevant from a formal point of view?

In social psychological research on social dilemmas, the relevance of take-some versus give-some scenarios has been intensively discussed. Whereas many results indicate that this distinction may be empirically relevant, in our taxonomy it is completely irrelevant for the definition of different types of social dilemmas. The distinction between the cost of an investment *x* and the shares from investment *x* makes it possible to interpret any experimental game with both costs and shares as a take-some or a give-some game as well. By focussing on the investment in the common pool, which yields to benefits from the common payoff function, a scenario could be introduced in an experimental game so that it could be labelled as a “give-some” game in social psychological research. In this case, the instruction given to the subjects could be formulated as follows: “You have to decide how many of your endowment you want to invest in the common-pool resource.” On the other hand, by focussing on the costs reducing the baseline-payoff the same scenario can be described as a “take-some” game. Such an instruction might be as follows: “There are **e** [in a specific instruction, **e** will be replaced by a number] units in the common-pool resource. You have to decide how many units you want to take from the resource.” It is also possible to describe such scenarios as investment games in two different markets, as Ostrom, Gardner and Walker (1994) do. Thus, while from an economic point of view it is irrelevant for the definition of public goods or common goods whether they involve extraction or production , in psychology public goods are often understood in the sense of “give-some” scenarios.

The fact that this taxonomy ignores differences between “take-some” or “give-some” scenarios because of their formal identity does not necessitate disapproval of empirical research investigating the effects of embedding one and the same formal game in two different scenarios. On the contrary, even if the formal structure is one and the same, useful hints about psychological effects can be found, when differences in cooperativeness can be detected between settings, where the task is to give from an endowment as opposed to settings, where the task is to take from a resource. The “endowment effect” (for a description compare Thaler 1994) can be understood in this sense. This effect has not yet been discussed or investigated within the context of social dilemma research.

Social dilemmas with fixed quotas

A proportioning rule with fixed quotas is given if the proportion get from the common payoff $f(x + y)$ is clear *a priori*. In many experiments on social dilemmas the corresponding situations are symmetrical, with each subject consequently getting $1/n$ of the common payoff $f(x + y)$. Yet it is also possible to define asymmetrical situations with different fixed quotas, mapping the subjects $1 \ldots n$ on quotas $q_1 \ldots q_n$. What alone is decisive is that these quotas are defined *exogenously*, that is, independently from the investment x of the respective subject. In other words, what is important is that these quotas are fixed. The quotas must add up to 1. This means that the common payoff is completely divided up among the subjects.

N-Person Prisoners‘ Dilemma

Specifying equation 1 by setting both fixed quotas (with $q = 1/n$ in the simplest case) and a linear payoff-function creates experimental settings like prisoners dilemmas or public-good games (without thresholds). In this specification, the payoff function is

$$f(x + y) = a(x + y).$$

In order to construct a dilemma between self-interest and group-interest, the following conditions in equation 1 must hold:

$$c > 0,\ a < n*c \text{ and } a > c \text{ e.g. } 0 < c < a < n * c.$$

In order to induce a dilemma situation, the value of *c* must be greater than 0. This is because *c* presents the costs of co-operative investments and is multiplied by –1 in the general formula $u(x,y)$. The condition $a < n * c$ ensures that the resulting payoff from an individual's co-operation is less than if the person retains the co-operative investment (ceteris paribus). The condition $a > c$ ensures that complete co-operation renders a better total payoff than complete defection, which is an essential feature of the dilemma.

In the simplest case each subject can only decide whether to co-operate or not. This is represented by a dichotomous variable x:

$$x \in \{0,1\}.$$

The resulting situation is widely known as a N-person prisoners' dilemma. It can be graphically represented as follows:

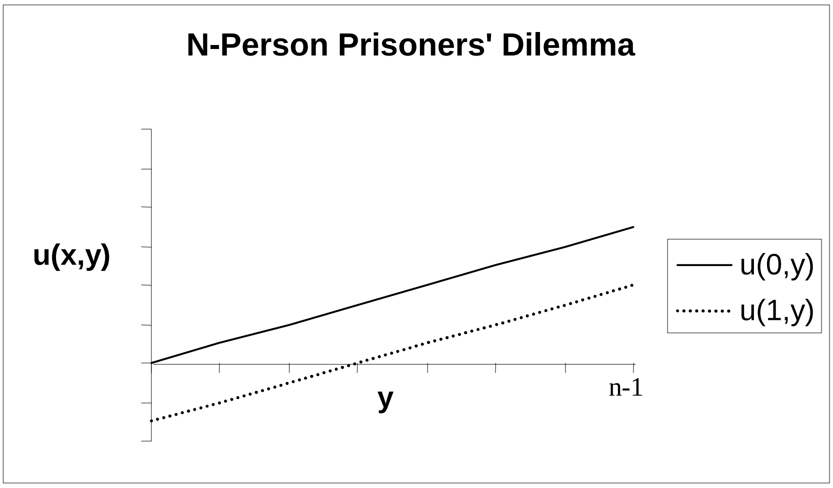

Figure 2: A graphical representation of the n-persons prisoners' dilemma in a generalized form

The graphic shows clearly that defection is in fact a dominant strategy in this social dilemma: the defection line is always above the co-operation line. If the parameters of the formula are interpreted , then *a* gives the slope of both the lower co-operation line and the upper defection line. Because the common payoff function is linear, both lines are parallel to each other. The parameter *c* gives the distance between both lines. But be careful, this is the distance in the mathematical sense, i.e. the shortest link between the two lines, the perpendicular distance.[7] Therefore *c* is not equal to the temptation to defect. The temptation is given by taking the (longer) distance between the two lines on the ordinate. Consequently, the temptation to defect depends both on *a* and on *c* (the formulae will follow below).

Instead of dichotomous decisions, it is also possible to admit any relative amount of co-operation. What holds is $x \in [0, 1]$, and not $x \in \{0, 1\}$, as in the dichotomous case. Thus, for example, an x-value of 0.4 could be interpreted as a co-operative contribution of 40%. If we are dealing with a public good, this means that a subject contributes 40% of his endowment to the public good.

A short example will illustrate how a N-person prisoners' dilemma can be constructed using the formal representation given above. For example, setting n = 4, a = 8, b = 2 and c = 3 results in the following payoff function:

$$u(x, y) = 2 - 3 * x + 1/4 * 8 * (x + y); \quad \text{with } c = 3 \text{ and } x \in \{0,1\}.$$

7 c would give the vertical distance between the two lines (instead of the perpendicular distance), if the abscissa represented x+y instead of y. In this case, the lines would not be strictly parallel. The line representing defection would begin at an abscissa-value of 0 and end at n-1, the line representing co-operation would begin at an abscissa value of 1 and end at n.

The corresponding payoff table is the following:

Table 2: Payoff matrix with n = 4, a = 8, b = 2, c = 3 and a proportioning rule with fixed quotas.

Subject's decision \ **Total number** of altruistic choices (x + y)	0	1	2	3	4
Altruistic, co-operative	-	1	3	5	7
Egoistic, defective	2	4	6	8	-

Some formulas make it possible to calculate interesting characteristics of this game. The temptation to defect t depends both on a and c:

$$t = u(0, y) - u(1, y) = c - a/n \, .$$

In our example the temptation to defect is expressed in one unit, $t = 1$. Thus in this game defection earns a fixed marginal income of one unit; that is, independent of how many other participants cooperate, one's private income will always be increased by one unit if the respective subject decides not to cooperate but to defect instead.

The advantage of co-operation for the whole group g is calculated by:

$$g = n * u(1, n-1) - n * u(0,0) = f(n) - f(0) - n * c = n * (a - c).$$

In our example $g = 20$ units. If all the participants defect, the total payoff is 8 units; if all the participants co-operate, the total-payoff is 28 units. This parameter is also psychologically relevant, as many experiments have shown. In many experiments on social dilemmas it has been demonstrated that chances of co-operation co-variate with g: In other words cooperation decreases if g is small and increases if g is high (cf. Dawes 1980).

As already mentioned above, asymmetries can be introduced in the dilemma. The subject's share of the earnings in the common payoff f(x+y) would be dependent on his fixed quota, whereby the sum of the quotas must add up to 1:

$\sum q_i = 1$.

In order to gain an absolutely clear insight in the taxonomy presented here, it is important to see that the quotas are fixed or *exogenous*; that is, they are completely independent of the subjects' investments, although there are different quotas for different subjects in this case. Besides reducing the slope of the two lines and the distance between the two lines, this kind of asymmetry may also soften the severity of a social dilemma. So, for example, if one subject is assigned to a quota of 0.9, the incentive to co-operate is very high for that subject, because he will receive a big portion of the profit of the common payoff function, whereas for the other subjects, with correspondingly reduced quotas, defection remains a dominant strategy. In other words: asymmetry may overcome a prisoners' dilemma, if cooperation is a dominant strategy for at least one powerful participant. He is expected to decide independently of what the others. If

$a > n_i c$, with $n_i = 1/q_{i,}$,

then co-operation – instead of defection – becomes a dominant strategy for a subject *i* in the asymmetric case.

Indexing *n* in this formula supports the suggestion that the introduction of different quotas can be interpreted, for example, as a variation in the power of votes. In this case the following condition needs to be met: $\sum n_i = N$, whereby N does not represent the number of subjects but the number of votes.

Public-good games with a threshold

From the point of view of the underlying formal structure, simple public-good games without a threshold are equal to N-person prisoners' dilemmas. Thaler (1994) also takes this in his consideration:

> …A typical public goods experiment uses the following procedures. A group of subjects (often, college students) is brought to the laboratory. Groups vary in size, but experiments usually have between four and ten subjects. Each subject is given a sum of money, for example, $5. The money can either be kept and taken home, or some or all of the money can be invested in a public good, often called *group exchange*. Money invested in the group exchange for the n participants is multiplied by some factor k, where k is greater than one but less than n. The money invested, with its returns, is distributed equally among all group members. Thus, while the entire group's monetary resources are increased by each contribution (because k is greater than one), each individual's share of one such contribution is less than the amount she or he invests (because k is less than n). To take a concrete example, suppose k=2 and n=4. Then if everyone contributes all $5 to the public good, each ends up with $10. This is the unique Pareto efficient allocation: no other solution can make everyone better off. On the other hand, any one individual is always better off contributing nothing, because in exchange for a player's $5 contribution that player receives $2.50, while the rest of the payoff ($7.50) goes to the other players. In this game, the rational, selfish strategy is to contribute nothing and hope that the other players decide to invest their money in the group exchange … These conditions constitute what is sometimes called a 'social dilemma'
> (Thaler, 1994; S. 9f.).

This citation also shows that the taxonomy presented here includes public-good games as understood by economists, for, after all, in experimental economics Thaler is a well known author and researcher.

Thus one can directly turn towards the question of how to represent public-good games with a threshold γ, where an additional payoff is given if the threshold is surpassed. This scenario can be represented by the following formula for the common payoff $f(x + y)$:

$$f(x,y)=\begin{Bmatrix} 0 \; iff \; (x+y)<\gamma \\ U \; iff \; (x+y)\geq\gamma \end{Bmatrix}$$, whereby *U* gives the total payoff for the whole group.

As in a prisoners' dilemma, the proportioning rules in a public-good game with thresholds can be represented by fixed quotas, and they are thus independent of the respective contributions of the individual subjects. Therefore, the essential difference between the prisoners' dilemma game and the public-goods game with thresholds lies in the common payoff function f(x + y). A game-theoretic analysis shows that – similarly to prisoners' dilemmas –in such cases there is also one inefficient equilibrium namely where $x_i = 0$ applies to all subjects. On the other hand, in contrast to prisoners dilemmas and to the definition of social dilemmas, there is also a set of efficient equilibria. All of these efficient equilibria fulfil the condition $(x + y) = \gamma$. This is why public-goods games with thresholds are not social dilemmas in the strict sense, but a hybrid of a social dilemma and a co-ordination game. When no communication is allowed, it may become difficult to achieve such an efficient equilibrium; and this may suade the subjects to make non-co-operative choices, as is also the case in the example presented above with respect to arriving at the inefficient equilibrium. If communication is allowed, it seems irrational for subjects to agree to the inefficient equilibrium, because in the efficient equilibrium everybody gets the best bargain and there is no temptation to defect at all (otherwise it would not be an equilibrium). Nevertheless, psychological problems may arise with respect to the evaluation of fairness, especially if the threshold cannot be achieved by symmetric contributions (as in the volunteers game). In this case, the bargain of those who contribute to achieving the threshold will be less than the bargain of those who do not contribute, although the payoff for everybody will be higher than if the total contribution were below the threshold. Issues of fairness in the distribution of social gains are relevant in this case.

Prisoners' dilemmas and public-good games with quadratic payoff functions

By taking a quadratic instead of a linear common payoff function and maintaining the proportioning rule with fixed quotas, a simple change is introduced. This change can be understood as an extension of a prisoners' dilemma game or a public-good game; the formal equivalence of both types of game has been shown above. The particular interpretation of the formalism as a public good or as a prisoners' dilemma may depend on the semantic embedding of the underlying incentive structure or the contextual components. However, in most of the cases such scenarios with common quadratic payoff functions will be interpreted as public-goods games. One reason for that is that these games, as will be shown below, do not have a dominant strategy. Nevertheless they have an inefficient and a unique equilibrium (cf. Keser 1996); they can thus be defined as a social dilemma.

In this case, the common payoff function is:

$$f(x + y) = (x + y)(a-d(x+y)).$$

Another parameter *d* must be introduced: In order to specify the parbola, three different parameters are necessary. Besides the additional constant, one parameter determines the linear part and the other the quadratic part of the polynom. In this case, because of the absence of a dominant strategy, calculating the equlibria requires that the best responses be considered. Let us construct another example by choosing the following values for the parameters of the general formula for social dilemmas. Setting a = 64, b = 0, d = 8, the size of the group n = 8 and the private linear costs c = 2, leads to the following subject's payoff function:

$$u(x, y) = -2x + 1/n*(x + y)(64 - 8*(x + y).$$

Taking the formula for the common payoff function, the welfare optimum can be calculated by renaming the total appropriation (x + y) with s, and calculating the first derivative in order to find the optimal total appropriation s, yielding $s^* = (a - c)/2d$. Generally, the symmetric welfare optimum is given by a contribution x^* with $x^* = (a - c)/(2nd)$ and a total contribution $(x^* + y^*) = (a - c)/(2d)$. In our example this is 0.484375 and 3.875 respectively.

Multiplying formula for the common payoff function with the quota 1/n gives the individual payoff-function. The calculation of the first derivative with respect to x gives the best-reply function x' = b(y) = ((a - c)/2d) - y. Equating this term with y/(n - 1) yields the symmetric Nash-equilibrium $y_{eq} = (n - 1)((a/n) - c)/(2d)$ and thus $x_{eq} = y_{eq}/(n - 1) = ((a/n) - c)/(2d)$.

In our example, x_{eq} is 0.375, the total contribution in the equilibrium $(x_{eq} + y_{eq})$ is 3. It can be seen that this once again constitutes a social dilemma. In the equilibrium, the total resulting contribution is 3 units. In the welfare-optimum, however, the total contribution is 3.875 units. Generally, the best response is

(a - cn)/2d - y, if the calculated value lies in the interval [0…1]. Otherwise it is 0 or 1 respectively.

Figure 3 plots the example. It includes the best-response function b(y) and three different parabolas. In the quadratic case each parabola (instead of the line in the linear case) represents the subject's payoff-function, given his/her contribution rate. In this plot, three parabolas can be found for x = 0, $x = x_{eq} = 3/8$ and x = 1. So each parabola represents a different share invested by the subject, with 0%, 37.5% and 100%. The first parabola (x = 0) gives the subject's payoff (which can be found on the left ordinate) when the subject doesn't contribute to the community at all, whereas the parabola x = 1 gives the payoff when the subject invests his whole endowment for the community. In the discussion of linear payoff functions, the latter was called the "co-operative" contribution. It can be seen that the intention of a certain proportion of the contribution is not so clear for a quadratic public good: the contribution of x = 1 units can be both co-operative and defective, depending on the investment of the others. As mentioned above, the subject's payoff is given on the left ordinate, the plot of the best reply refers to the right ordinate, which assigns values between 0 and 1. The figure shows that the best response when there is a value of y = 2.625 is x = 0.375, so that in such cases x equals y/(n-1). Thus the equilibrium can be found here. The subject gets a payoff $u(x_{eq},y_{eq})$ of about 13.36; in the welfare-optimum each subject gets a payoff $u(x^*,y^*)$ of about 15 units (this cannot be seen in the figure, because the parabola for $x^* = 0.48$ is not plotted in).

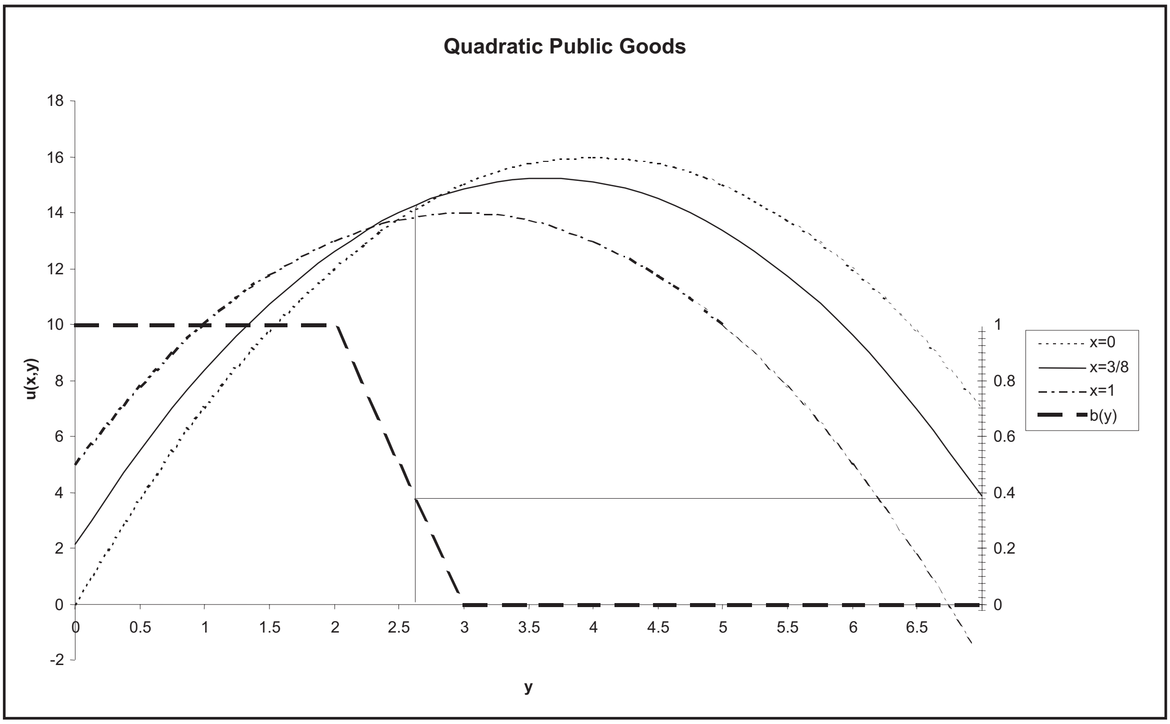

Figure 3: A public-good game with a quadratic payoff function.

As in the linear case, the situation can be complicated by introducing asymmetric quotas. And as in the linear case, this can soften the severity of the dilemma situation. Our research group (Ostmann, Wojtyniak and Beckenkamp) performed experiments with such asymmetric quotas and quadratic production functions. This scenario was used to model the over-fertilization with nitrate, which leads to harmful consequences for the groundwater. The asymmetry was used to represent the different sizes of the farmers' pastures The payoff function renders a good fit to the payoff, the payoff depending on the nitrate concentration in the soil.

Social dilemmas with proportional quotas

This section considers social dilemmas in which the share of the common payoff function is not given by fixed quotas, but by distributions, depending on the relative amount of an individual's own contribution. Then, the quota can be calculated by the following formula:

$$q_i = \frac{x}{(x+y)}.$$

As in all the cases discussed before, the quotas must add up to 1:

$\sum q_i = 1.$

The consequence of this proportioning rule is that the subject's payoff from *a given* common production depends on both his own contribution and on the contribution of the others. This means that, in contrast to a prisoners' dilemma, besides the common payoff from the common production, the amount of the individuals payoff is also dependent on the intensity of his own contribution. This seems like a first step towards privatisation and thus a first step towards a solution to the dilemma situation. As will be shown below, in effect this is true for linear common payoff functions, but in quadratic production functions the dilemma situations may be aggravavated. The resulting dilemma situation will be termed a "commons" dilemma. As will be shown, this formal categorization grasps the difference between common goods and public goods in the economic sense, as is, for example, presented in Ostrom, Gardner and Walker (1994, p.7) or in Blankart (1991, p. 56).

Commons Dilemmas

Thus far the presentation can be understood as a suggestion that avoids defining the difference between public goods and commons in reference to the existence of a dominant strategy. From the point of view of the categorization presented here, there are public-good dilemmas that lack a dominant strategy. The basic aim is, with the use of the quotas, to formally grasp the relevant distinction between public goods and commons (with respect to rivalry and excludability) and, by employing the form of the common payoff function, to introduce a special sub-case. To do that a distinction between first-grade and second-grade polynomials has been introduced, that is, a distinction between linear and quadratic functions. The use of polynomials as a model for production functions is also very common in economic theory. A consideration of third-grade polynomials, which are also very common in economics, has been omitted. There are two reasons for that: first, the cubic part is irrelevant for the generation of pecuniary externalities (and thus for the dilemma), and second, to my knowledge, there are no experiments with cubic payoff functions in the dilemma-context.

In analogy to the games with fixed quotas, two different situations now have to be discussed that refer to the proportional quotas: linear and quadratic payoff functions. However, with respect to dilemma situations the linear case is irrelevant, because linear production functions with proportional quotas lead to the privatisation of the common production so that no pecuniary externalities remain. This is the situation for “private goods”. There is thus no social dilemma in this case. Each subject receives exactly the proportion he has produced, as can be seen in the following formal illustration.

If $f(x + y) = a(x + y)$ and the quota is $q = (x/(x + y))$, then in this case $u(x,y) = b – cx + qf(x + y)$, which can be simplified to $u(x) = b – x(a – c)$. $u(x)$, is independent of y; that is, the payoff is individually determined and independent of the other subjects. The situation can be visualized as a common pool situation where there is parcelling so that each individual has his own territory, where he can operate independently of others. This is exactly the situation for private goods (rivalry and exclusion in consumption).

Let us now consider the commons dilemma situation - that is, situations with proportional quotas and quadratic payoff functions. Whereas for linear payoff functions the situation becomes easier with proportional quotas, for quadratic payoff functions it becomes more complicated compared to the quadratic public goods with fixed quotas. As already mentioned above, the subject’s contribution now not only influences the common payoff and the costs, but also the share that the subject receives from the common payoff. The situation is a social dilemma with a unique and inefficient Nash-equilibrium. Thus this taxonomy also integrates experiments that have been presented in the literature on commons dilemmas, for example, in the literature from the Bloomington group (especially by Gardner, Ostrom and Walker), which has especially focussed on such scenarios (cf. Ostrom, Gardner und Walker 1994; Gardner, Moore und Walker 1994), and from others, whose experimental scenarios are based on their work (like Beckenkamp & Ostmann 1996; Beckenkamp & Ostmann 1998; Casari & Plott 2000, Beckenkamp 2001).

The latter also present experiments on non-renewable resources. This makes it clear that, besides the formal attributes presented here, in iterated games further formal criteria for distinguishing between experimental situations can be introduced. So it is possible to distinguish between functions in which the common payoff f_t depends on the payoff from the period before f_{t-1}. Analysing such experimental data is very complicated, and in order to make an adequate evaluation of the subject’s behaviour and thus to ensure that the differences between the features of the scenario and the features of the subjects are unambiguous, clarity with respect to the best responses and equilibria in each period is necessary. If such analysis is not carried out, then it is impossible to evaluate whether it is still possible for a subject to maintain a flow of extraction from the resource or whether the payoffs will inevitably break down, even if the subject makes the best possible decisions in order to maximise the common profit or his individual profit.

Although such time-dependencies are not considered here, since this taxonomy facilitates the determination of corresponding parameters, it may nevertheless be a first reasonable step towards evaluating sub-games.

For the sake of completeness, let us also give a formal characterization of this case. As in the case of public goods with quadratic payoff-functions, the common payoff function is:

$f(x + y) = (x + y)(a - d(x + y)) – cx.$

Therefore, the welfare optimum x^* is the same in both cases, as given by the formula:

$x^* = (a - c)/(2nd)$ with total contribution $(x^* + y^*) = (a - c)/(2d)$.

The best response is given by the formula:

$b(y) = (a-dy-c)/(2d)$.

The resulting symmetric equilibrium is:

$y_{eq} = (1/d)(a-c)(n-1)/(n+1)$ with $x_{eq} = y_{eq}/(n-1) = (1/d)(a-c)/(n+1)$.

With this formula the parameters of experimental scenarios as presented in Ostrom, Gardner and Walker (1994) can be easily computed. However in doing so it is necessary to "translate" their payoff functions into the general form that has been proposed here. For example, in one of their scenarios the payoff-function is introduced as the investment in two markets. This "translation" can be fruitful, because it would make it much easier to compare the results of the different research groups. The game-theoretic parameters could be calculated by simply applying the formula, without any deep knowledge of the underlying mathematic background. To resume, a graphical overview of the categorization system is given below.

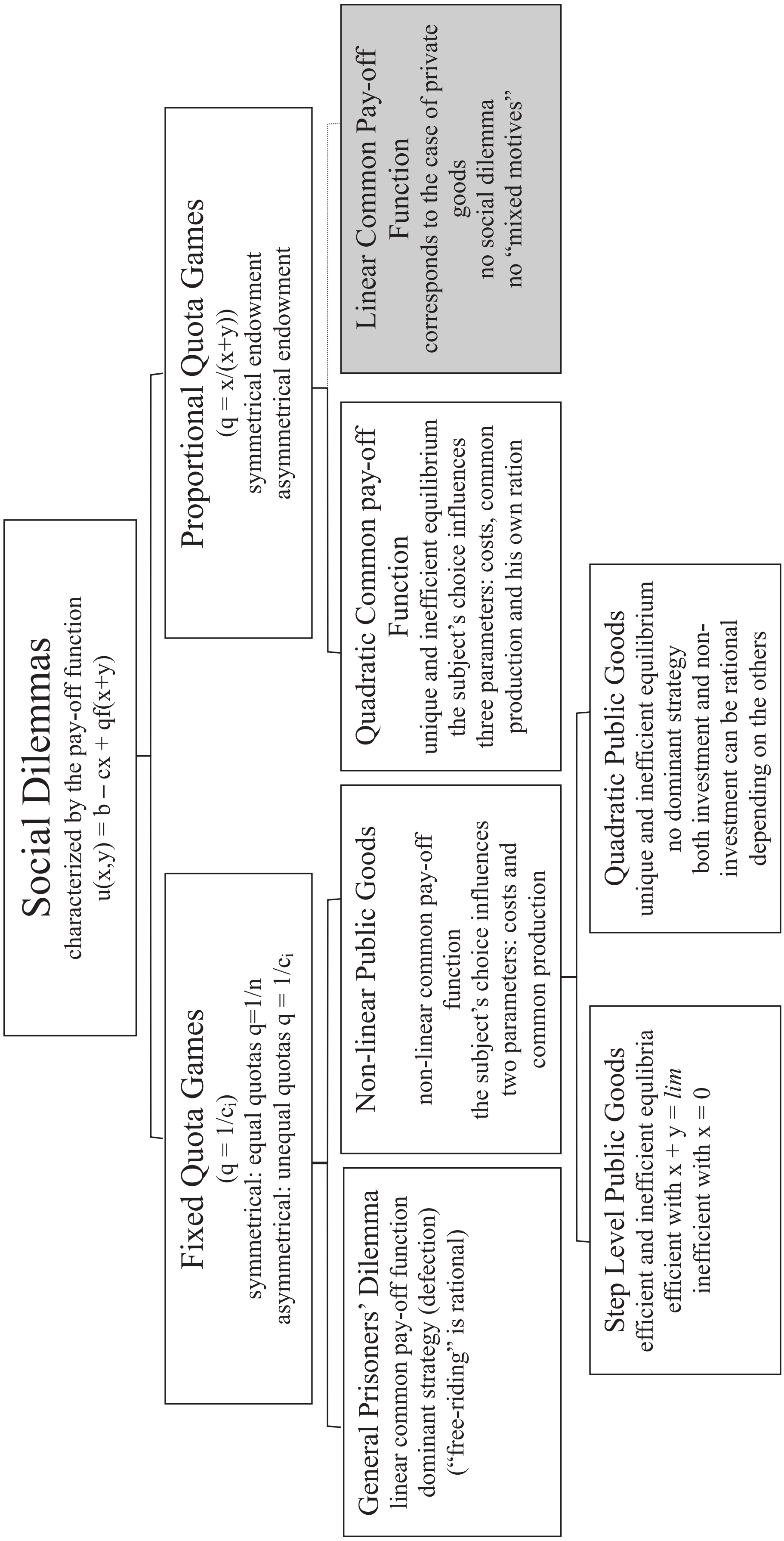

Figure 4: A graphical overview of the categorization of social dilemmas

Summary

In summary, the taxonomy presented here makes it possible to easily integrate the results from different experimental scenarios for social dilemmas, specifically, both from experimental economics and from social psychology. Although it is based on a game-theoretic conception, with its two (nonformal) criteria — i.e. excludability and rivalry in consumption — it takes into account the traditional classification of goods. By using game-theoretic criteria to distinguish between the experimental scenarios, it fosters the attempt to integrate results from both social psychology and experimental economics.

Literature

Axelrod, R. & Hamilton, W. D. (1981). The Evolution of Cooperation. *Science*, *212*, 1390-1396.

Beckenkamp, M. (2001). *Soziales Dilemma Umwelt - Spieltheoretische und psychologische Aspekte von Umweltproblemen.* . Habilitationsschrift. Philosophische Fakultät der Universität des Saarlandes, Saarbrücken.

Beckenkamp, M. & Ostmann, A. (1996). A member's view of a common: Cognitive and socio-cognitive factors. In Christine Roland-Lévy (Ed.), *Social and Economic Representations* (pp. 1203- 1220). Paris: IAREP 1996; Université René Descartes.

Beckenkamp, M. & Ostmann, A. (1999). Missing the target? Sanctioning as an Ambigous Structural Solution. In Michael Smithson (Ed.), *Resolving Social Dilemmas* (pp. 165- 180). Philadelphia: Taylor and Francis.

Blankart, C. B. (1991). *Öffentliche Finanzen in der Demokratie* (1991. Aufl.). o.O.: Vahlen.

Casari, M. & Plott, C. (2000). *Keeping an eye on your neighors* (Agents monitoring and sanctioning one another in a common-pool resource environment). http://129.79.82.27/IASCP00/Papers/casarim040800.pdf. (Paper for The 8th Biennial Conference of the International Association for the Study of Common Property (IASCP))

Dawes, R. (1980). Social Dilemmas. *Annual Review of Psychology, 31*, 169-193.

Dawes, R. M., McTavish, J. & Shaklee, H. (1977). Behavior, communication, and assumptions about other people's behavior in a commons dilemma situation. *Journal of Personality and Social Psychology*, *35*, 1-11.

Dawes, R. M. (1991). Social dilemmas, economic self-interest, and evolutionary theory. In D. R. Brown, & K. J. E. Smith (Eds.), *Frontiers in Mathematical Psychology* (Essays in Honor of Clyde Coombs, Vol. 35, pp. 53-79). New York: Springer.

Dörner, D., Kreutzig, H., Reither, F. & Stäudel, T. (1983). *Lohhausen: Vom Umgang mit Unbestimmtheit und Komplexität* (3. Aufl.). Stuttgart: Kohlhammer. [Lohhausen. On dealing with uncertainty and complexity]

Eisentraut, R. (2001). *Machtgefälle, Informationssuche und soziales Wissen in einem zweifachen sozialen Dilemma: Problemlösen unter der Bedingung struktureller Asymmetrie*. Aachen: Shaker.

Ernst, A. M. (1999). *Information in de Klemme: Informationsdilemmata und die Nutzung natürlicher Ressourcen*. Habilitationsschrift. Psychologisches Institut der Albert-Ludwigs-Universität Freiburg.

Funke, J. (1987). Problemloesen in komplexen computersimulierten Realitaetsbereichen. *Sprache und Kognition, 4*, 113-129.

Gardner, R.; Moore, M. R.; Walker, J. M. (1994*). Racing for the Water: Laboratory Evidence on Subgame Perfection*. Research Report. Workshop in Political Theory and Policy Analysis. Bloomington, Indiana University.

Gentner, D. (1983). Structure-Mapping: A Theoretical Framework for Analogy. *Cognitive Science,* 7, 155-170.

Hardin, G. (1968). The Tragedy of the Commons [The Population Problem has no Technical Solution; It Requires a Fundamental Extension in Morality]. *Science*, *162*, 1243-1248.

Keser, C. (1996). Voluntary Contributions to a Public Good When Partial Contribution is a Dominant Strategy. *Economic Letters, 50,* 359-366.

Liebrand, W., Messick, D. & Wilke, H. (Eds.). (1992). *Social Dilemmas* (Theoretical Issues and Research Findings). Oxford: Pergamon Press.

Liebrand, W. B. G., van Lange, P. & Messick, D. M. (1996). Social Dilemmas. In A. S. R. Manstead, & M. Hewstone (Eds.), *The Blackwell Encyclopedia of Social Psychology* (1st, pp. 546-551). Oxford: Blackwell.

Liebrand, W. B. & McClintock, C. G. (1988). The Ring Measure of Social Values: A Computerized Procedure for Assessing Individual Differences in Information Processing and Social Value Orientation. *European Journal of Personality, 2,* 217-230.

Ostrom, E., Gardner, R. & Walker, J. (1994). *Rules, Games and Common Pool Resources*. Ann Arbor, Mich.: University of Michigan Press.

Platt, J. (1973). Social Traps. *American Psychologist*, *28*, 641- 651.

Smith, A. (1988). *Der Wohlstand der Nationen*. München: dtv. (im Original erschienen 1776 - Aus dem Englischen von Horst Claus Recktenwald)

Thaler, R. H. (1994). *The Winner's Curse* (Paradoxes and Anomalies of Economic Life). Princeton, New Jersey: Princeton University Press.

Yamagishi, T. (1995). Social Dilemmas. In K. S. Cook, G. A. Fine & J. S. House (Eds.), *Sociological Perspectives on Social Psychology* (pp. 311-335). Boston: Allyn and Bason.