

Emotional appraisal of different situations in commons dilemmas depending on social value orientation

Martin Beckenkamp; Wiebke Gmbel – University of Saarland – Psychology

How to escape the tragedy of the commons

A usual common consists of a common pool resource and members interacting and managing the resource (cf. Ostrom 1990). It is well known that the resource is endangered to be overexploited if the members of the common behave self-interested or "economically rational" ("tragedy of the commons"; Hardin 1968). In order to counter this dilemma it is often suggested to establish a sanctioning regime on the incentive structure. Hardin himself postulated that "ruin is the destination toward which all men rush, each pursuing his own best interest..." (Hardin 1968, p. 1244). He suggests to prevent ruin by "mutual coercion mutually agreed upon" (Hardin 1968, p. 1247). Bell and others seized this idea in different psychological studies and demonstrated empirically the efficiency of punishment and rewards (Bell, Peterson and Hautoluoma 1989; Martichuski & Bell 1991; Harvey, Bell & Birjulin 1993). These studies were essentially orientated towards Platt's approach of "social traps" (Platt 1973), which is basically a behaviouristic approach.

Research which is not theoretically orientated towards behaviourism also investigated different forms of "coercion" which might successfully prevent the ruin of common-pool resources, like moral suasion, commands and prohibitions. Using such measures in practice requires control and sanctioning regimes. In many cases sanctioning takes place by raising fines. Such regimes have also been examined in experimental studies as in Ostrom, Gardner and Walker (1994).

By using control and sanctioning regimes in order to prevent the depletion of common pool resources the original situation is changed with respect to two important aspects: a legal and an economic one. From an economic point of view introducing a control and sanctioning regime involves a change of the incentive structure. The individual's gain by defection is reduced by the expected value of the fine, which has to be paid if control is successful. This expected value can be calculated by multiplying the amount of the fine with the probability, that control will be successful. As a consequence game-theoretic parameters of common-pool resources like equilibria or best replies are changed. In other words: sanctioning regimes change the incentive structure of the common-pool resource. Such changes have to be considered in the interpretation of experimental results like Ostrom, Gardner and Walker (1994) do by giving complete game-theoretic analyses of the structural changes. This is the economic aspect of introducing control and sanctioning regimes.

Engel (in press) argues that economic analysis of such sanctioning regimes is a useful project. But the economic model might not mislead to focus merely on negative incentives of sanctions imposed in case of detected defections. Though interesting questions could thus be answered, caution is indicated. Even companies in perfect competition would not compare the product of threatened sanctions and probability of detection with the advantage by defection. Two of many reasons for it which Engel cites are the influence of norms on

cognitive models and the discursive character of norms. Like Ostrom (1990) he also emphasises the significance of the context sensibility of norms.

From a psychological point of view both change in incentive structure and in external change of norms may lead to unexpected paradox-like reactions. If people believe themselves to have specific legitimate behavioural freedoms which are threatened or taken away, “the person experiences psychological reactance, which urges attempts to regain in threatened or lost freedom. Because the most direct form of reestablishing a freedom is to exercise it, reactance will frequently result in an intensified form of the behavior whose freedom has been threatened” (Brehm 1996; p. 466). Besides the introduction of norms, loss of freedom may also be caused by changes in incentive structures like the introduction of taxes, which prevents people from buying things at low prices.

How much of a threat each of these measures is depends on the persons’ perceptions of pressure. “Considerable research has shown that intensified behavior or increased preference for the behavior whose freedom has been threatened is directly proportional to the magnitude of the pressure to comply” (Brehm 1996; p. 466). Such reactions may also be expected when there is no insight in the necessity of such measures. Yamagishi’s (1986, 1988) results clearly show that bad experiences in the management of social dilemmas increase the acceptance to introduce countermeasures. His results suggest that without such experiences the acceptance may depend on the social value-orientations (cf. Liebrand & Dehue 1996) of the individuals. Co-operatively oriented subjects, assuming that the other participants will also behave co-operatively, have no insight in the necessity of counter-measures. Thus they are not willing to invest any amount of money in institutions like control and sanctioning regimes. Competitively oriented subjects - on the other hand – invest for such measures in social dilemmas even without former experience. One might argue that the emergence of reactance depends on the social value orientation which is projected onto the other subjects from a very egocentric position.

The efficiency of monetary sanctioning regimes – experimental results

In an experimental computer scenario we (Ostmann, Wojtyniak & Beckenkamp) simulated a self-governed common based on a scenario from Ostrom-Gardner-Walker in different series, published in Beckenkamp & Ostmann (1999) Ostmann (1998); Wojtyniak & Ostmann (1999). In this scenario each subject encounters the same payoff-function and is provided with the same possibilities of appropriation to the common. Thus in terms of game-theory we can speak of a symmetric game, and the one-period incentive structure of such a commons-dilemma can be represented by a normal form game with n players, action spaces, and a payoff function. Every player can choose a contribution (x) whose value may range from 0 to an upper limit (w), called the actor’s endowment. Like in Ostrom, Gardner & Walker (1994, p. 114ff.), the commons-dilemma is operationalised with eight appropriators ($n = 8$) and a quadratic production function. The parameters correspond to one of their experimental scenarios. Therefore let us call this scenario the “baseline-scenario”. However, the instruction or “story” for our baseline-scenario was different. For that reason a short verbal description of the scenario is given before the introduction of the formal description and the concrete values of the parameters.

The baseline-scenario

Eight subjects participated in a series of exactly 20 decision rounds. The common resource is renewed in every round. In every round, the subjects have to decide how much of their given endowment of 25 token they want to invest in a common which is defined by a quadratic production function. This endowment corresponds to what Ostrom, Gardner and Walker in their scenario (1994) call “high endowment” condition, the joint production function corresponds to their “market 2”. Thus the return from the common depends both on the amount of total investment and the own investment. Accordingly the investment of one subject yields externalities to others. However, the subject’s investment also yields costs. These costs merely depend on her or his own amount of investment. Thus these costs can be interpreted as a private good – they yield no externalities. In contrast to Ostrom, Gardner and Walker (1994) we interpret a part of the equations as costs instead of a decision to invest in a second market. Yet both scenarios are formally equivalent. By transformation of the formulas given in their description the following representation can be given.

$$(1) u(x,y) = p - cx + q(x,y)f(x+y); c > 0$$

Let x denote the amount of the own investment and y the total appropriation from all of her or his $n-1$ partners. Then $u(x,y)$ gives the payoff for a specified subject. p represents a fixed base payoff, which the subjects can save by investing nothing. The term $-cx$ represents the linear private costs of the investment. $q(x,y)$ gives the quota which determines how the joint product is to be divided. In commons dilemmas the individual’s consumption decreases others’ shares, expressed by the proportional rule:

$$q = x/(x + y)$$

The joint production $f(x+y)$ is quadratic and defined by:

$$f(x+y) = f(s) = s(a-bs).$$

By completing the equations with the following parameters $p = 125$; $c = 5$; $a = 23$; $b = 0.25$ this formal representation is isomorphic to the formal part of the experimental design with high endowment introduced in Gardner, Ostrom & Walker (1994, p. 115). This scenario served as a baseline on which a monetary sanctioning regime with different degrees of severity was introduced. We call these different experimental conditions with different monetary sanctioning regimes the **Saarbrücken scenario**. Compared to the baseline supplementary we find in this scenario an individual use limit (λ), a threshold of unbearable overuse (η) of the common where surpassing results in evoking a control, and sanctions if control was successful. Thus the incentive structure is essentially changed. Based on Hardin’s ideas of coercion and based on economic ideas of changes in incentive structures it is plausible that the following assumption is standard: By establishing an institution on the baseline-scenario prosperity of the common can be achieved or regained. In this case a suitable control and sanctioning regime with adequate inspection probabilities and sanctions has to be set up.

The limits of tolerance

So in our scenario the institution was specified in the following way. The individual appropriation limits are (λ) defined such that the sum of the individual appropriations is as near as possible to the “welfare optimum”, which is the optimal level of investment (the welfare-optimum is given with a total appropriation of $s = 36$. $n\lambda = 40$ so that utilising the limits would lead to slight inefficiencies in the government of the experimental commons). Monitoring is costly. This is why control is carried out only if co-operation is judged as “endangered”. If total appropriation s surpasses the threshold $\eta = 46$, the control takes place imposing fixed costs $k_a = 40$. There are several reasons why a value of 46 was chosen for the threshold in our scenario. The joint-production-function has its maximum at a total appropriation of $s = 46$ (nevertheless, because of the opportunity-costs of the investment, the optimal usage is given with $s = 36$). Thus each token which surpasses this threshold harms the production. The threshold of control is not defined at the total usage which would give the welfare-optimum (that is at $s = 36$) because monitoring is costly. High monitoring costs may motivate to tolerate a certain degree of overuse. Therefore in our scenario the threshold for total usage is higher than the numbers of participants multiplied with their individual appropriation limits (λ), that is $n\lambda < \eta$ or $40 < 46$ respectively.

The sanctioning-rules

Total appropriation is allowed up to a limit of $\eta=46$. If this threshold is surpassed, control is evoked. The monitoring agency imposes fixed costs of 40 units ($\kappa= 40$). The monitoring agency randomly chooses members of the common. The number of subjects randomly drawn varied in our different experimental series. In the first series the monitoring agency randomly drew 2 out of the eight members of the common, in the later series only one member was randomly chosen. This corresponds to control probabilities of 1/4 and 1/8 respectively. In the following only experimental results of the later series will be reported, thus only series where only one random-control took place if the threshold has been surpassed.

From the members randomly drawn the monitoring agency controls the compliance with the individual limits. Two cases are possible: (1) no defection has been detected, that is each controlled member respected the individual limit of maximal 5 token. (2) At least one of the controlled members committed an offence against the appropriation rules by investing more than 5 units in the commons.

1. If no defection is detected the monitoring costs are distributed equally among the commoners. That is in this case each member loses 5 ($40/8$) units from the payoff which he would have received if no control would have taken place.
2. If some defection is detected, defectors have to bear the monitoring costs and additionally to pay a fine. This fine is a compensating fee to everyone of their non-detected partners. The amount of the compensation fee depends on the individual's overuse by taking the difference between the individual's appropriation and the individual's limit λ (which was set to 5; see above). This difference $D = x - \lambda$ was multiplied with a sanctioning factor σ . This sanctioning factor was an independent variable which was varied in different experimental conditions of the series reported

here. We investigated experimental conditions with the following sanctioning factors: 0, 0.5, 1, 1.5, 2, 2.5. Please note that in the experimental condition a fine has to be paid if a defection has been detected. But the fine is merely the recompense of the costs for the control. That is with $\sigma = 0$ the amount of the fine is independent of the actual overdrawing of the respective subject who has merely to pay 40 units.

Thus the game has been changed formally. A transfer component $t(x, K(s, \omega))$ has been added to the payoffs. The transfers are dependent on the respective subject's appropriation x and dependent on random control events $K(s, \omega)$. These random control events hinge on whether the total appropriation s surpasses the threshold and hinge on which kind of random-mechanism is defined. In our case the random variable is equally distributed with a probability of $1/8$ for each subject to be detected.

Now the utility function can be formulated in the following way:

$$u(x, y, \omega) = u(x) + t(x, K(s, \omega)).$$

The control regime does not change the optimal level of investment. Like in the baseline-scenario the optimal level of investment (the welfare optimum) is given with a total usage s of 36 token. Thus the welfare optimum can theoretically be achieved by an investment of 4,5 units from each subject (in the actual experiments subjects had to choose integer values, so they could either invest 4 or 5 token in order to achieve the welfare optimum).

Theoretic and empirical efficiency of monetary sanctions

Indeed, the theoretic efficiency of such changes can be shown by game-theoretic analysis of the structural changes. In our discussion we will focus on two payoffs:

(1) the overall maximum yield which can be achieved if each subject (symmetrically) appropriates in such a way as to enable them to achieve the welfare optimum. From this payoff the base-payoff p is subtracted, because p is constant and independent of which amount of appropriation is chosen by a subject. Thus the corresponding payoff is calculated by the following formula:

$$u_{\max}(x, y) = -5x + 1/8 f(8x)$$

(2) the mean payoff which results if each subject chooses his best reply function such that an equilibrium of the individual choices can be found. Like before in point (1) the base-payoff p is subtracted and ignored in further analysis. In contrast to the baseline-scenario the resulting Nash-equilibria may become asymmetric and multiple equilibria can be found, depending on how many subjects defect. Therefore in the following intervals of the equilibria under the different sanctioning-conditions will be given with the lowest and highest possible values respectively.

In a next step the mean payoff in equilibrium is divided into the payoff in welfare-optimum. The resulting fraction can be interpreted as the "theoretic efficiency" of the equilibrium-solution compared to the welfare optimum. That is a fraction of 0.25 means that in the corresponding equilibrium only $1/4$ of the payoffs can be attained compared to the payoffs in

the welfare optimum. The term “efficiency” corresponds to what Gardner, Ostrom and Walker (1994) name “rent”. In this context we prefer the word “efficiency” making clear that a comparison of different monetary sanctioning regimes is intended by using these values. Besides the theoretic efficiencies we will also present efficiencies empirically found. In this case the mean payoff found in different experimental conditions is divided into the payoff in welfare optimum. By comparing theoretic and empirical efficiencies conclusions with respect to the intended effects of different sanctioning regimes can be drawn. Theoretically both high control probability and high sanctioning factors lower the incentives to deviate from the limiting standard. In terms of efficiency an inspection cost of 40 means an efficiency loss of 0.123 if the agency is called for inspection.

Measuring the performance of the sanctioning system in a first step involves a description of the structural changes in the incentive structure by the different sanctioning regimes respectively. As already mentioned above the changes lead to multiple asymmetric equilibria. Increasing both the inspection probability and the sanctioning-factor (σ) leads to improved theoretic efficiency. By varying sigma it is possible to compare sanctioning regimes that differ in the strength of sanctioning. In Wojtyniak and Ostmann (1999) experimental results from studies with fixed inspection probabilities of $p=0.125$ and different sanctioning factors can be found. The sanctioning factors in the different experimental conditions gradually reached from weak monetary sanctions ($\sigma = 0$) to very severe sanctions ($\sigma = 2.5$) with steps of 0.5. Please remember that $\sigma = 0$ represents a regime with fixed sanctions where a detected defector merely has to pay the inspection costs independent of his overdrawing.

Table 1 shows in the second row the intervals of the equilibria in the different experimental conditions. In the fourth row the corresponding efficiencies of these equilibria are given. In a first step the observed mean total usages can be compared with the theoretically expected total usages (row 3 vs. row 2). The mean total usage is above the interval of equilibria in high-sanctioning-factor conditions. However, merely looking at the mean total usage during the course of the experiment is not a sensible measure of effective management of the common. A fictive example may illustrate this point. Imagine one group which in half of the rounds has total usages of 0 and in the other half of the rounds total usages of 72. The mean total usage is 36. A total usage of 36 corresponds to the welfare-optimum with a payoff of 324 for each subject (if the fixed-base payoff p of 125 units is ignored, cf. equation 1). The group however obtains 0 units both in every round and all over the rounds. Thus the efficiency of the group is 0 although the mean total usage corresponds to a total usage of 1. For that reason, from the efficiencies in every round we have calculated the mean efficiency (which does not correspond to the efficiency of the mean total usage!). These efficiencies can be compared with the theoretic efficiencies which can be found in the equilibria. The distance from the empirical efficiencies to the interval – also calculated in each round – gives a measure we call the “co-operation-shift” (4th row). The main result of these series is that observed means of the co-operation shift are negative for all sanctioning factors with the only exception of a moderate value.

Table 1. Efficiency targets and observed co-operation shifts

sigma	0	0.5	1	1.5	2	2.5
equilibrium range	60.25	58.94	55.67	55.00	53.33	52.17
	62.50	61.04	58.80	57.40	55.00	53.69
observed mean (s.d.)	60.79 (12.6)	59.42 (12.9)	53.49 (9.1)	61.67 (11.7)	57.96 (27.3)	55.57 (9.9)
efficiencies of equilibrium range	0,423	0,470	0,578	0,598	0,645	0,675
	0,335	0,393	0,475	0,523	0,598	0,635
co-operation shift	-0,06	-0,13	0,09	-0,31	-0,13	-0,13
# trials	6	5	6	5	5	7

These results might suggest that moderate sanctioning might have the same (or even better) effects than high sanctioning. We followed this idea in the first part of the experimental series presented here, when we had only results from the sanctioning conditions $\sigma = 0$, $\sigma = 1$, and $\sigma = 2$. Based on these data in Beckenkamp & Ostmann (1999) we presented a “U-shaped” effect. We discussed the possible influence of cognitive and socio-emotional processes on observed behaviour. We supposed that the loss of freedom experienced in the severe sanctioning-condition might evoke reactance phenomena. The results presented before have cast doubt on these interpretations. The prominent role of the experimental condition might solely be based on a statistical random error which occurs in less than 5% of the cases with the significance level chosen in our analyses. In other words: perhaps the results of the experimental condition were based on a seldom outlier effect. The raising doubts with respect to our first interpretations about the cognitive and socio-emotional processes stimulated a new experiment with a series of trials which should serve with answers to the following questions: Is there – from a psychological point of view – optimal sanctioning factor which is severe enough to be accepted and moderate enough to obviate counteracting? In other words: Can the prominent role of the experimental condition $\sigma = 1$ be replicated? How do subjects react emotionally during the course of the experiment? Are there dependencies between the social value-orientation of the subjects (competitors, co-operators and individualists) and their emotional appraisal of the experimental situation? We searched for clarification of these questions in a new experiment.

How social value orientation influences emotional appraisal and appropriation behaviour in monetary sanctioning regimes

Psychological research in context of social dilemmas grasped one idea of Kelley and Thibaut’s (1978) interdependence theory. A given outcome matrix – which can be interpreted for our purposes as the payoff matrix the subjects encounter in our experiments – can be transformed by the subjects into an effective matrix. The transformation process depends on preferences of the respective subjects. Thus one and the same outcome matrix may lead to very different effective matrices. The effective matrix may represent preferences, motivations

and values of the subjects. Particularly the social value orientation has been intensely studied in social dilemma research.

The influence of social values in social dilemmas

“Social dilemmas are characterised as situations in which private interests are at odds with the collective interests” (Liebrand, van Lange & Messick 1996, p.546). This definition includes the definition of commons-dilemmas, where the pursuit of self interest produces external costs which are higher than the individual gain. For this reason it is useful to consider results from social-dilemma research in the context of this study.

Deutsch (1958) reasoned that “individuals systematically differ in the kind of decision they make in settings of outcome interdependency. Deutsch reasoned that decisions are guided by interpersonal orientations, or certain rules people use to evaluate own and others’ outcomes, and distinguishes three different interpersonal orientations: Co-operation refers to valuing joint outcomes for both self and other, competition to valuing relatively higher outcomes for self than for the other, and individualism to valuing own outcomes independent of the outcomes of the other” (Liebrand & Dehue 1996, p 609).

By taking into account two different weighting factors which transform the original payoffs in individual utilities these different types of decision-making can be represented formally by a simple linear equation:

$$u_i = w_1(p_i) + w_2(p_j)$$

The utility for a subject i is given by adding the weighted payoff for the subject and the weighted payoff for the other subjects. According to Liebrand and Dehue (1996, p. 610) subjects with a strict co-operative orientation can be represented by setting w_1 and $w_2 = 1$; subjects with an individualistic orientation can be represented by setting $w_1 = 1$ and $w_2 = 0$ and competitive subjects by setting $w_1 = 1$ and $w_2 = -1$. Many other social value orientations and transformation scales have been suggested. Nevertheless, the predominant important social values which are also discussed mostly are co-operation, competition and individualism.

Decomposed games were developed by Messick & McClintock (1968) in order to give a valid assessment of the subject’s social value. In these games the consequences of a choice for oneself and for the other person are made explicit. By giving a sophisticated series of decomposed games subjects’ social value orientation can be classified. Liebrand and Dehue (1996, p. 611) give one example:

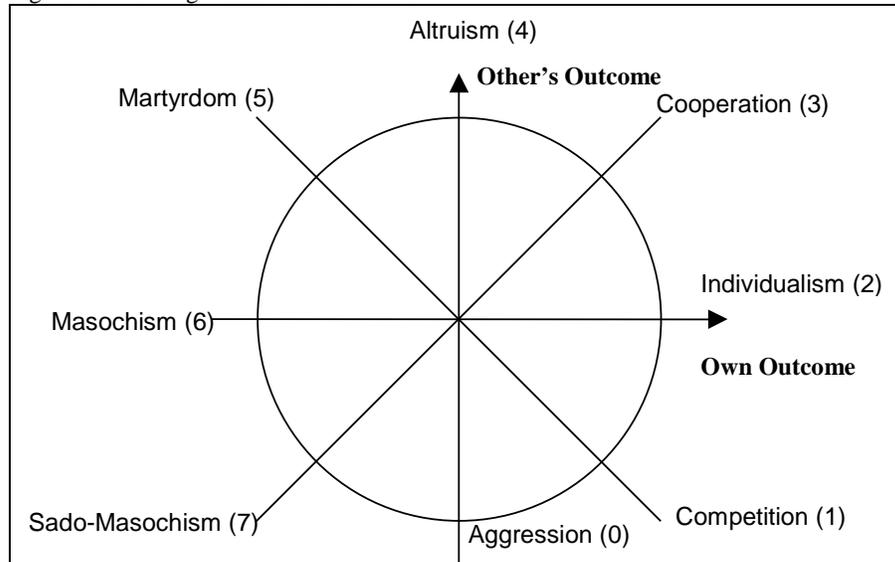
Table 2: An example of a decomposed game (from Liebrand & Dehue 1996)

	A	B	C
Self	490	560	500
Other	90	300	500

This game could keep apart co-operators from competitors from individualists in the sense of the definition given before choosing C, A and B respectively. Based on this idea of decomposed games Liebrand & McClintock (1988) developed the Ring Measure of Social

Values which allows to rank measures of social values within a circle with its centre in the point of origin of a co-ordinate system. The abscissa represents the weight one gives the own payoff, the ordinate gives the weight one imputes the other. The closeness to the circle and accordingly the distance to the centre gives the reliability of the measure or the consistency of the subject's answers. the values of the subjects are given as polar coordinates. The first value gives the type with 1 for competition, 2 for individualism, 3 for co-operation, and so on. Thus the first value represents the orientation in the circle. The second value which may lay between 0 and 100 gives the length of the vector and represents the coherence of the subjects' answers. Thus this value represents the reliability of the measure.

Figure 1: The Ring Measure Value Scale



In social-dilemma research many studies demonstrated the influence of social value orientations on choice behaviour and expectations about others' choices, the temporal stability of social value orientations and its external validity beyond laboratory experiments (cf. Liebrand & Dehue 1996).

The influence of different social value orientations on the appropriation behaviour in our scenario: Theoretical expectations

Liebrand (1984) examined which presumptions subjects have with respect to other subjects if no further information about each other exists. Based on the egocentric model and the results of Kuhlman & Wimberley (1976) he assumed that subjects suppose that others have the same social value orientation as themselves. In an experiment with group sizes of six to seven subjects Liebrand (1984) could confirm his hypothesis. However in another experiment with 19 subjects he could not. Possibly with growing group size co-operative subjects may increase their doubts on the unconditional co-operation of the others. This hypothesis may be supported by findings that co-operators react on non-co-operation (cf. Kelley & Stahelski 1970; Kuhlman & Marshello 1975). The same studies also suggest that the behaviour of pro-selfish subjects (that is individualist and competitors) is primarily non-co-operative "and only weakly influenced by the other's choice behaviour" (van Lange 1991, p. 29). When others don't verify the assumptions in iterated games, then co-operative subjects co-operate with co-

operative others but don't co-operate with non-co-operative others. Individualists and competitors however never begin to co-operate during iterated games, independent of the social value orientation of the other participants. We expected similar results in our experiment. Based on the results of our former experimental series and the analysis of appropriation behaviours and verbal remarks from subjects we can add some assumptions. In contrast to pro-selfish subjects (that is individualist and competitors) we suppose that in progression of the experiment with resulting experiences of others' defections co-operators try to find opportunities to give revenge. In other words we suppose that co-operators don't try to choose best replies (like individualists) or best relative replies (like competitors) but some modifications of tit-for-tat in their strategic behaviour after having gained experience of defections in the course of the experiment.

Let us summarise these assumptions in some specific hypotheses:

1. In the first round we expect both that co-operative subjects expect that others' will also co-operate and pro-selfish subjects expect competition. The appropriation of co-operators should be maximal one eighth of the estimated total appropriation (that is if a subject estimates that the total usage is 40 she or he will appropriate with maximal 5 token). In other words we expect that co-operators don't lay claim to better quotas than the other participants. Pro-selfish subjects however should not be guided by such considerations. They should try to find best (or relatively best) replies to the estimated total usage of the others. Therefore – and due to the properties of the experimental commons - their appropriation behaviour should not correlate linearly with their estimations. A positive correlation between own appropriations and estimated appropriations of the others should be found for co-operative oriented subjects.
2. Individualists should use strategies which correspond to best replies. The strategies they choose should depend on their estimation of the appropriation of the others and lie near best replies (the exact calculation of best replies may be very difficult, especially because the calculation of expected values is difficult in case of over-cropping).
3. Competitors try to maximise their relative advantage to others. That's why we expect that competitors will not try to choose best replies but will lay claim on even higher appropriations, leading to better relative advantages for them. This will also be true during the whole course of the experiment (but interpretation may become complicated if a defection has been successful. At this point we formulate hypotheses ignoring the respective experimental sanctioning condition).
4. Co-operators should increase their appropriations in the course of the experiment (this is only true if defections of other subjects occur and total usage often is relatively high. But this restriction is met in almost every experiment both in our series as well as in Ostrom, Gardner & Walker (1994)).
5. We expect that co-operatives' anger increases in the course of the experiment because the behaviour of the others disappoints their expectations.

Method

Subjects, Design

Subjects consisted of 72 students of psychology, 27 male and 45 female. Most of them participated in an introductory psychology class. Subjects were run in groups of eight. Some days before the groups were composed the subjects filled out the ring-measure value-scale. With the questionnaire subjects gave preferences for different dates which were given for the accomplishment. The subjects' social value orientations were evaluated. Three blocks of three groups were constituted. Within each block the social-value orientations were paralleled as good as possible between the three groups (subjects serving as substitutes had to stand in for subjects who dropped out at the date of the experiment. In this case paralleling failed). Within a block the allocation of the sanctioning-factor (0, 1, 2) was randomised. There were three groups per sanctioning-factor with 8 subjects respectively playing 20 rounds.

Procedure

The experiment took place in a computerised setting. The screens of the subjects were protected from sight of other subjects. After each round the subjects received a feedback of their appropriation, the total usage, whether a control has been performed and whether the control has been successful. If control was successful the uncontrolled subjects received the number of the subject who has been detected and the amount of transfer-payoff they receive. The detected subject received feedback of the amount of the fine. The subjects could not associate the numbers to a specific person. After the feedback as well as before the experiment subjects filled out a questionnaire for the assessment of emotional states (EMO 16; Schmitz-Alzert & Hüppe 1996). Before every round subjects gave estimation of the total usage in the next round.

After raising the social-value orientation it was aspired to compose groups of eight subjects with as much as possible equal proportions of social value orientations. Subjects with reliability-measures lower 50 (10 cases out of 72) were not considered in the analyses of social-value orientations. The following distribution of social value orientations resulted from this procedure:

Table 3: Reliable Social Value Orientation*Sanctioning-Factor: Contingency - Table

		Sanctioning-Factor			Sum
		sigma = 0	sigma = 1	sigma = 2	
Social Value Orientation	1,13			1	1
	1,23		1		1
	1,38	1	1		2
	1,63		1		1
	1,88			1	1
	2,13	2	1	2	5
	2,23		1		1
	2,38	3	3	4	10
	2,50	2	1	2	5
	2,63	5	3	2	10
	2,88	5	2	1	8
	3,13	4	1	8	13
	3,23		1		1
	3,38		2	1	3
Sum of reliable measures		22	18	22	62
Total Group Size		24	24	24	72

In context of social dilemma research social-value orientations are often dichotomised in pro-selfish and pro-social (cf. van Lange 1996). One advantage from dichotomising is that the power of statistical tests can be increased leading to a better proof of statistical effects. However one problem of such a procedure is to find the adequate borderline. Just in our data with the frequencies presented above the problem is severe, because many cases ($n = 5$) lay exactly on the borderline separating the space of individualists and co-operators (value = 2.5). We decided to categorise these subjects as “pro-selfish”, because a social value of 2.5 means that those subjects focus with a weight of $\frac{3}{4}$ on their own payoff and a weight of $\frac{1}{4}$ to the payoff of the others. So selfishness outweighs co-operation in this case. So far the theoretic arguments – unfortunately we will see that empirical results are reluctant to our considerations.

Results

Hypothesis 1: co-operators' accordance of estimation and claim in the first round

We suspect that co-operators' appropriations are lower in the first round than the appropriations of individualists and competitive subjects. We also suspect that co-operators' estimations of the appropriations of the others are lower in the first round than the estimations of individualists and competitors. The mean appropriations and the mean estimations of the usage of the others (that is estimated total usage minus own appropriation) in the first round is given in the following table:

Table 4: Mean realised and estimated others' appropriations in the first round depending on social value orientation

Mean own realised appropriations in the first round depending on social value orientation					Mean estimated sum of others' appropriations in the first round depending on social value orientation				
Social Value Orientation	Mean	Median	N	standard deviation	Social Value Orientation	Mean	Median	N	standard deviation
1,13	9,00	9,00	1	,	1,13	31,00	31,00	1	,
1,23	5,00	5,00	1	,	1,23	40,00	40,00	1	,
1,38	7,50	7,50	2	,71	1,38	38,50	38,50	2	2,12
1,63	5,00	5,00	1	,	1,63	50,00	50,00	1	,
1,88	6,00	6,00	1	,	1,88	38,00	38,00	1	,
2,13	6,20	5,00	5	2,17	2,13	36,20	35,00	5	2,39
2,23	6,00	6,00	1	,	2,23	30,00	30,00	1	,
2,38	5,80	5,00	10	1,99	2,38	36,60	35,00	10	2,41
2,50	9,60	6,00	5	8,65	2,50	36,60	37,00	5	2,70
2,63	6,80	6,00	10	3,74	2,63	38,40	37,50	10	5,97
2,88	7,25	5,00	8	5,28	2,88	36,00	35,00	8	4,11
3,13	5,62	5,00	13	1,19	3,13	36,38	35,00	13	2,40
3,23	7,00	7,00	1	,	3,23	35,00	35,00	1	,
3,38	5,67	6,00	3	,58	3,38	36,67	36,00	3	2,08
Altogether	6,55	5,00	62	3,57	Altogether	36,87	35,00	62	3,92

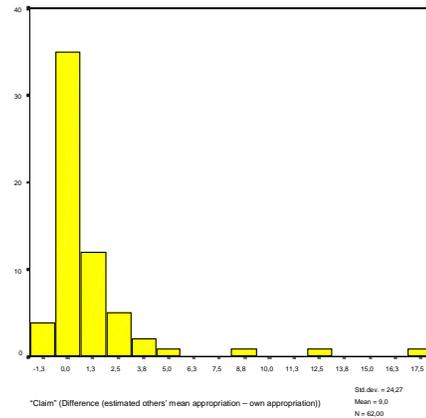
(shaded cells represent values of “pro-social” subjects)

Our expectations with respect to the comparison of estimations of pro-selfish subjects vs. other subjects cannot be confirmed. Pro-selfish subjects seem to give the same predictions as pro-social subjects. More important in our theoretical discussion was the expectation that the estimations of pro-selfish subjects should correspond to their appropriation behaviour. Indeed, all the estimations lie in an interval that corresponds to the others will appropriate with 5 or maximal 6 token. In other words, in the first round pro-selfish subjects expect no crisis in the management of the common. What is odd is that subjects with other social value orientations also don't.

We also expected that pro-social subjects behave restrained. Their own behaviour should correspond to the estimated total usage near the welfare optimum. In other words: we expect that co-operators behave in symmetrical co-operation, they behave like they expect the others to behave towards them, particularly because no bad experiences have yet been made. Calculating the subjects' mean claims by taking the difference of the mean estimated appropriation of the others (that is the estimated total usage of the others divided by 7) minus the own appropriation yields the results shown in table 5. No differences between pro-selfish and pro-social subjects can be found. Even worse, subjects at the borderline between individualism and co-operation seem to behave very greedy. The high value can be explained by the behaviour of one subject appropriating with 25 units which also explains the high standard deviation. Nevertheless, there is no significant indicator for more restrained behaviour of pro-social subjects compared to pro selfish subjects. Please note that in table 5 the median of claims over all social value orientations is 0, the mean however is 1,28. This means that 50% of the subjects demand the same appropriation or less than the others. The other 50% however who demand more are inclined to demand for higher defections worthwhile the risk of possible detection.

Table 5: Mean claims in the first round depending on social value orientation

Social Value Orientation	Mean	Median	N	standard deviation
1,13	4,5714	4,5714	1	
1,23	-,7143	-,7143	1	
1,38	2,0000	2,0000	2	,4041
1,63	-2,1429	-2,1429	1	
1,88	,5714	,5714	1	
2,13	1,0286	,0000	5	1,9017
2,23	1,7143	1,7143	1	
2,38	,5714	,0000	10	2,0636
2,50	4,3714	1,0000	5	8,5777
2,63	1,3143	,5714	10	3,3031
2,88	2,1071	,0000	8	5,0414
3,13	,4176	,0000	13	,9358
3,23	2,0000	2,0000	1	
3,38	,4286	,4286	3	,4286
Altogether	1,2811	,0000	62	3,4676



Let's summarise the results with respect to hypothesis one so far: no differences can be found with respect to the three variables estimated total usage of the others, realised appropriation and realised claims between pro-selfish subjects and pro-social subjects. Thus the results contradict the egocentric hypothesis.

However, this is not the whole story. A significant pearson correlation between the estimated total usage of the others and the own appropriation can be found for pro-social subjects ($r = 0,502$; $n = 35$), but not for pro-selfish subjects ($r = -0,01$; $n = 27$; the same is true for finer classifications of "pro-selfish subjects).

Thus – contradictory to all expectations – no differences with respect to the realised appropriation behaviour could be demonstrated between pro-social subjects and subjects with other social value orientations. However – according to our hypothesis - their behaviour is guided in a symmetrical way by their expectations of what others will do. A plausible interpretation is that co-operators have more fear to be exploited than the subjects with other social value orientations, yet the results of an analysis of the emotional questionnaire with respect to fear/anxiety before the first round in the experiment don't indicate any difference.

The lack of evidence might be explained by the rather unspecific item in the questionnaire which is not especially constructed for the context of social-dilemma research. According to Yamagishi (1986) the "fear" and the "greed" in social dilemma games are rather "motivational bases" than emotional states. Another explanation might lay in the fact that subjects have enough emotional distance to the situation by perceiving it as a game (which is a fundamental problem of experimental games). This explanation then again raises the question why different correlation coefficients depending on social value orientations can be found.

Hypotheses 2 and 3: Do individualists choose best replies and are competitors the greediest subjects?

Transferring the idea which lays behind the definition of social value orientations to our scenario leads to the following assumptions. For any given payoff matrix individualists try to maximise their payoff. In other words, individualists look for the best reply.

The consequence from this consideration is that a group of individualists playing against each other should lay near the theoretic equilibria. On the other hand getting a precise overview of best replies from their payoff matrices is difficult for the subjects because of the control-scenario which requires calculating with expected values. Competitors should even appropriate beyond the best replies, because the maximal relative advantage of the won payoff compared to the mean payoff of the others doesn't lay in the calculated equilibria but above. From the point of view of the social value orientation co-operators should appropriate less than the others because co-operators like to maximise the total payoff in the symmetric scenario given here. In other words: co-operators' preference lies in the welfare optimum. However we also expect that co-operators defend themselves against the defections by giving up co-operation. The specific strategic ideas which co-operators try to develop in order to achieve this goal are difficult to predict, because co-operators are confronted with the problem that their counter-aggression is directed towards a whole group and not against specific subjects. Thus counter-aggression might lead to unintended injuries. Verbal reports from the former experimental series and our interpretations given there suggest that co-operative subjects try to defend themselves by seldom but severe defections, as soon as they see a good chance to do so.

A different use of strategies should express itself in different distribution patterns in contingency tables of own appropriations and estimated usages of the others. A glance at such contingency tables with estimated total usages of the others and the own appropriations in dependency of social types doesn't show noticeable differences for the different social value orientations.

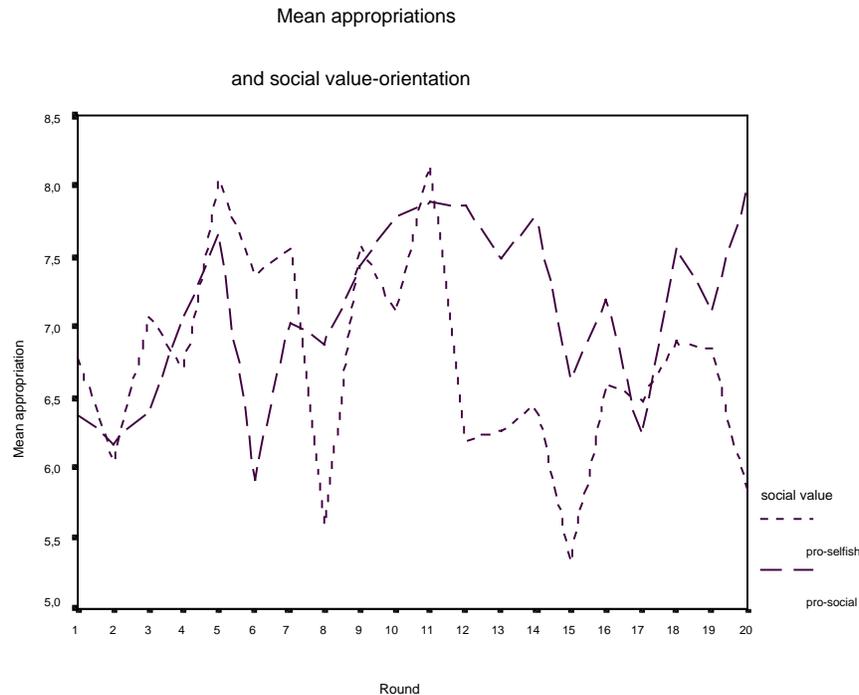
From the considerations given before we would also expect that the appropriations of competitors are higher than the appropriations of individualists. No predictions can be made with respect to the magnitude of the appropriations of co-operators, but we expect some differences in their distribution pattern.

A data analysis shows that the mean appropriation of competitors is 7,28 token compared to 6,53 token of individualists. A non-parametric Mann-Whitney test is significant, a t-test nearly misses significance ($t = 1,815$; $df = 126,71$; $p = 0,072$). The inhomogeneity of the variances might explain this difference between the two tests. In this context it is interesting to note that the standard-deviation of appropriations correlates with the social value orientations with the lowest standard-deviations for competitors (3,23, $n = 803$), then individualists (3,84; $n = 460$) and co-operative subjects (3,96, $n = 700$). This might indicate that co-operative subjects try to defend themselves by paying back. This interpretation can be sustained by a further analysis by a look at the appropriation patterns over 20 rounds in dependency of the social value orientation, which leads us directly to hypothesis 4.

Hypothesis 4: co-operators raise appropriations in the course of the experiment

Pro-selfish subjects increase their mean appropriations in the course of the experiment. The next figure makes clear how drastically appropriations change:

Figure 2: Mean appropriations in the course of the experiments



Besides the fact that the mean appropriations in the first rounds are significantly lower than in the rest of the experiment (eg. for the first two rounds vs. the rest $t = 2.447$; $df = 100.74$; $p = 0,016$) the change in the correlation between estimated total usage and own appropriation is remarkable: as we have seen in context of hypothesis 1 the appropriations of pro-socials correlate with their estimations. This is still true for the second round ($r = 0.35$; $p = 0.04$), but already in the third round the correlation breaks down ($r = -0.08$, $p = 0.645$). Taking the second part of the experiment (rounds 11 to 20) we get a significant negative correlation between the estimated usage of the others and the own usage ($r = -0.241$, $p = 0.00$). The corresponding correlation of pro-selfish subjects which was near zero in the first round ($r = -0.01$, $p = 0.97$) also becomes negative ($r = -0.267$; $p = 0.00$).

Thus we can conclude: even in the first rounds there is only a loose connection between social value orientations with appropriation behaviour and expectations which dissolve quickly by the experiences of defection. Subjects adapt their strategic behaviour by acting more and more anti-cyclical. This strategic behaviour pattern develops independent of the social value orientations of the subjects.

Hypothesis 5: co-operators' anger increases in the course of the experiment

This raises the question why the reliability of social value orientations usually is rather high. Is there really no difference in the evaluation of the situation between pro-selfish and pro-social subjects after the experiences of some rounds in the experiment? Or are there possibly different evaluations of the situation which lead to similar behaviour patterns. We suggested in hypothesis 5 that co-operators feel more anger during the experiment by the experience of defections. Comparing the means or medians yields no significant results. However an interesting pattern could be found by a look at the respective contingency table:

Table 6: Anger in round 1 vs. round 11-20 in dependence of social value orientations

	Anger in the first round		Sum	Anger in the rounds 11-20		Sum
	pro-selfish	pro-social		pro-selfish	pro-social	
not present (0)	21	24	45	150	186	336
very weak (1)	4	6	10	46	59	105
rather weak (2)	1	3	4	31	46	77
medium (3)		1	1	26	29	55
rather strong (4)	1	1	2	14	21	35
strong (5)				3	8	11
Sum	27	35	62	270	349	619

It can be seen that high anger values are seldom, but that in the second part outbursts have a much higher probability than in the first round of the experiment. But this is true both for pro-socials and pro-selfish subjects. Differences between these social value orientations are not significant ($\chi^2=2,18$; $df = 5$; $p = 0,82$).

And what about the U-shaped effect?

Besides the hypotheses given in the chapter concerning theoretical expectations another goal was announced in the abstract. One purpose of this experiment was the replication of a “U-shaped effect” which was found in the first part of our former experimental series with the sanctioning-conditions $\sigma = 0, 1, \text{ and } 2$. Moderate sanctioning (with $\sigma = 1$) gave the best results with respect to an efficient management of the commons (cf. table 1).

In the experiment presented here the mean total appropriations in the different sigma conditions are 61.82, 54.62 and 52.08 for $\sigma = 0, 1, \text{ and } 2$ respectively. The corresponding medians are 61.50, 53.00 and 51.50. The difference of the subjects' appropriations under the three conditions is significant ($F=12,37$; $df = 2, 1437$). By comparing these values with the values given in table 1 it can be seen that the values are similar with the exception of the experimental condition $\sigma = 2$. In the actual experiment subjects reached the target; they were even better than the theoretic equilibria would predict. Comparing this result with the U-shaped effect found before this result suggests that probably it was not the sanctioning factor 1 which was an outlier. Rather it seems that in our first series the appropriations observed in the sanctioning condition $\sigma = 2$ were unusual high. Thus the sanctioning factor remains a prominent point. Increasing the sanctioning-factor $\sigma = 1$ leads to the most significant decrease in appropriation. The mean appropriation is only negligible decreased by continuing

severity of monetary sanctioning regimes. But that is only half the truth, as we will see in the analysis of the efficiencies.

The sanctioning-factor $\sigma = 1$ is also interesting with respect to estimated total usages. In the theoretic assumptions we suggested that the estimations of the subjects are based on egocentric generalisations. In this sanctioning condition we get this result over all 20 rounds (cf. figure 3), which is astonishing especially compared to the other two sanctioning conditions (cf. figure 4).

Figure 3: Mean estimated others' appropriations and social value orientation in $\sigma = 1$.

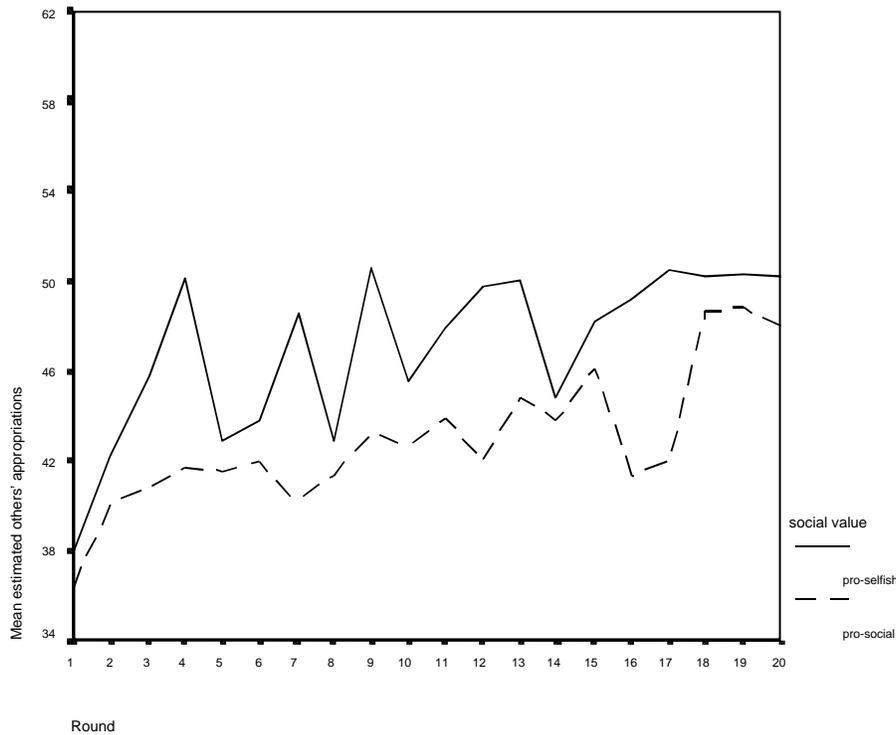
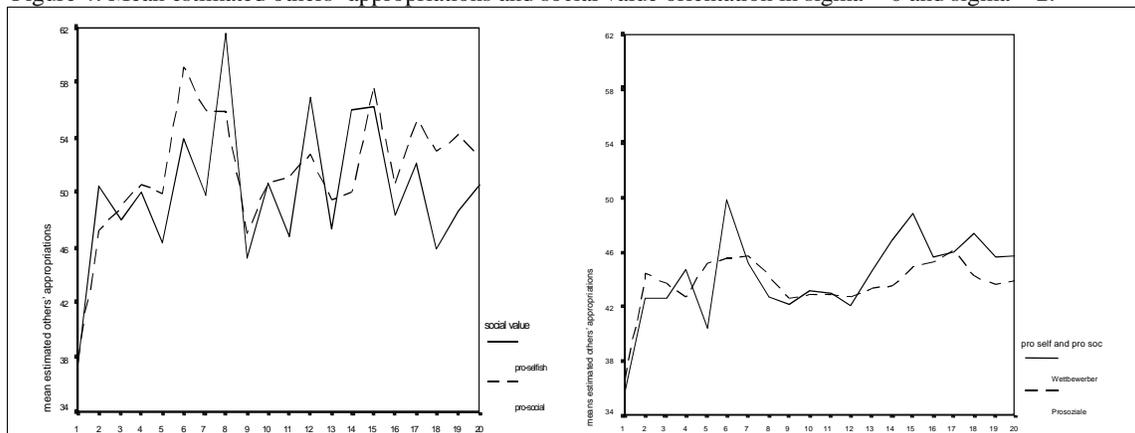


Figure 4: Mean estimated others' appropriations and social value orientation in $\sigma = 0$ and $\sigma = 2$.



Perhaps this is the reason why $\sigma = 1$ yields the best decrease in total usages, when we gradually rise the sanctioning factor. The social value orientations remain stable in this condition so that pro-socials can soften escalation. On the other hand, the sanctioning condition $\sigma = 2$ leads to less excited estimations. Although the decrease in mean total usages from the sanctioning factor $\sigma = 1$ to the sanctioning factor $\sigma = 2$ is not very noteworthy, the decrease in efficiency in this experimental series is! (remember that efficiencies measure both mean total usage and variance of usages, cf. table 1). The efficiencies are 0.19, 0.58 and 0.67. The observation of this result strengthens the assumption, that in this experiment the results from the sanctioning condition $\sigma = 2$ were extreme, so that an isolated interpretation seems rash. Therefore we intend to make an integrated data-analysis from all experimental data taken together and stop the interpretation here.

Discussion

In this experiment we investigated the efficiency of monetary sanctioning and the influence of different social value orientations. The results suggest that social value orientations become less and less important with increasing experiences of defection. Subjects learn to behave anti-cyclical. The management of a commons reprobates to a stock market game. Uncertainty on what happens in the next round leads to jumpiness both in expectations as well as in appropriation behaviour. Introducing monetary controls which consider the amount of overdraft (as it is done in the sanctioning conditions $\sigma = 1$ and 2) moderate this jumpiness. From former experimental series taken together with the results presented here it seems that especially the introduction of a moderate sanctioning regime leads to most noticeable positive consequences. Nevertheless, especially in the experiment presented here, continuing increase in sanctioning leads to wider decrease of defections, but the effect becomes less distinct.

Defections cause anger – even in the experimental setting. Neither the median nor the mean change with statistical significance, but nevertheless the progression in the scenario seldom leads to strong anger. Further analyses will be made to explore reasons for such anger. It is not yet clear whether co-operators are angry because of the defections, whether subjects' anger is high when their defection has been detected and fines have to be paid or whether anger is high when co-operators try to take pay back and are detected just in this moment.

The efficiency of moderate and higher sanctions may have two reasons: (1) the sanctions are tolerated by the co-operative subjects in consideration of the defections they encounter. This interpretation is obvious if you also take in account Yamagishi's research (1988) on sanctioning systems. Pro-socials feel relief of the strain to take revenge. Thus oscillations are smothered. (2) These oscillations are also damped by making the range of best replies smaller for individualists as well as for competitors. At this point further analyses will be made.

References

Beckenkamp, M. & Ostmann, A. (1999). Missing the target? Sanctioning as an Ambiguous Structural Solution. In M. Smithson (Ed.), *Resolving Social Dilemmas*. Philadelphia: Taylor & Francis.

- Bell, P. A., Peterson, T. R. & Hautoluoma, J. E. (1989). The effect of punishment probability on overconsumption and stealing in a simulated commons. *Journal of Applied Social Psychology, 19*, 1483-1495.
- Brehm, J. W. (1995). Reactance. In A. S. R. Manstead, & M. Hewstone (Eds.), *Blackwell Encyclopedia of Social Psychology* (pp. 466-468). London: Blackwell.
- Deutsch, M. (1958). Trust and Suspicion. *Journal of Conflict Resolution, 2*, 265-279.
- Engel, C. (iP). Die Grammatik des Rechts (Funktionen der rechtlichen Instrumente des Umweltschutzes im Verbund mit ökonomischen und politischen Instrumenten). In H. W. Rengeling (Hrsg.), *Instrumente des Umweltschutzes im Wirkungsverbund* (7. Osnabrücker Umweltgespräche). o.O..
- 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.
- Harvey, M. L., Bell, P. A. & Birjulin, A. A. (1993). Punishment and type of feedback in a simulated commons dilemma. *Psychological Reports, 73*, 447-450.
- Kelley, H. H. & Stahelski, A. J. (1970). Social interaction basis of cooperators' and competitors' beliefs about others. *Journal of Experimental Social Psychology, 16*, 66-91.
- Kelley, H. H. & Thibaut, J. W. (1978). *Interpersonal Relations: A Theory of Interdependence*. New York: Wiley.
- Kuhlman, D. M. & Marshello. (1975). Individual differences in game motivation as moderators of preprogrammed strategic effects in prisoner's dilemma. *Journal of Personality and Social Psychology, 32*, 922-931.
- Kuhlman, D. M. & Wimberley, D. C. (1976). Expectation of choice behavior held by cooperators, competitors, and individualists across four classes of experiments. *Journal of Personality and Social Psychology, 34*, 69-81.
- Liebrand, W. B. (1984). The effect of social motives, communication and group size on behaviour in a N-person multi-stage mixed motive game. *European Journal of Personality, 14*, 239- 264.
- Liebrand, W. B. G. & Dehue, F. M. J. (1996). Social Values. In A. S. R. Manstead, & M. Hewstone (Eds.), *The Blackwell Encyclopedia of Social Psychology* (1st, pp. 609-614). Oxford: Blackwell.
- 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.
- Martichuski, D. K. & Bell, P. A. (1991). Reward, punishment, privatization, and moral suasion in a commons dilemma. *Journal of Applied Social Psychology, 21*, 1356-1369.
- Messick, D., M. & McClintock, C. G. (1968). Motivational Bases of Choice in Experimental Games. In (Ed.), *Journal of Experimental Social Psychology* (Vol. 4, pp. 1-25). o.O..
- Ostmann, A. (1998). External control may destroy commons. *Rationality and Society, 10*, 135-154.
- Ostmann, A., Wojtyniak, B. & Beckenkamp, M. (1997). *Control and Sanction may destroy*

- Commons* (Experimental Results and some Microanalytic Explications). Discussion Paper Serie "Game Theory and Experimental Economics", Universität Karlsruhe/Institut für Statistik und mathematische Wirtschaftstheorie, Karlsruhe.
- Ostrom, E. (1990). *Governing the Commons* (The Evolution of Institutions for Collective Action). Cambridge: Cambridge University Press.
- Ostrom, E., Gardner, R. & Walker, J. (1994). *Rules, Games and Common Pool Resources*. Ann Arbor, Mich.: University of Michigan Press.
- Schmitz-Alzert, L. & Hüppe, M. (1996). Emotionskalen EMO 16. Ein Fragebogen zur Selbstbeschreibung des aktuellen emotionalen Gefühlszustandes. *Diagnostica*, 42, 242-267.
- Van Lange, P. A. M. (1991). *The Rationality and Morality of Cooperation* (The Influence of own Social Value Orientation and Perceptions of other's Personality Characteristics on Cooperation in a Social Dilemma) Unveröff. Diss., Rijksuniversiteit, Groningen. (Proefschrift)
- Wojtyniak, B. & Ostmann, A. (1999). Limits to accept - accepting the limits. Experimental tests of reduction policies for collective goods. *Submitted to: Environment & Behavior*.
- Yamagishi, T. (1988). Seriousness of social dilemmas and the provision of a sanctioning system. *Social Psychology Quarterly*, 51, 32-42.
- Yamagishi, T. (1986). The provision of a sanctioning system as a public good. *Social Psychology Quarterly*, 51, 110-116.
- Yamagishi, T. & Sato, K. (1986). Motivational bases of the public goods problem. *Journal of Personality and Social Psychology*, 50, 67-73.