

Running head: TOWARDS ANOTHER PARADOX OF SANCTIONING SYSTEMS

Towards Another Paradox of Sanctioning Systems in Social Dilemmas:

A Self-Determination Theory Approach

Hein Huyghe

Willy Lens

University of Leuven

Acknowledgement: I would like to thank Edward Deci and Maarten Vansteenkiste for their comments on the study.



### Abstract

Previous social dilemma research has shown that sanctioning defection may enhance cooperation. However, recent studies have shed some light on detrimental effects of sanctioning defection as well. Taking a self-determination theory perspective, this study points at another paradox of sanctioning systems: while sanctioning defection may push people to cooperate through external regulation, it may undermine internalisation of cooperative behavior and intrinsic motivation to cooperate by thwarting people's psychological need for autonomy. A Web experiment was conducted and showed that having been confronted with a sanction on defection makes people experience less choice about what to do in a social dilemma which is on its turn associated with less cooperation.

## Towards Another Paradox of Sanctioning Systems in Social Dilemmas:

## A Self-Determination Theory Approach

Personal interests often conflict with group interests. Take for example the issue of smog. In order to mitigate this kind of air pollution, Flemish policy makers often decide to temporarily lower the speed limit on highways. However, drivers might be tempted to exceed the lowered speed limit for their own benefit (e.g. saving time driving) through which a conflict between personal interests (gaining time) and group interests (smog reduction) arises. Situations like these are often referred to as *social dilemmas*. Social dilemmas are defined by two properties (Dawes, 1980<sup>1</sup>): (1) acting on personal interests (exceeding the lowered speed limit) yields higher outcomes for individuals (gain of time) than acting on group interests (complying with the temporal measure). However, (2) each individual attains a higher outcome (smog reduction) when everybody furthers group interests rather than personal interests (the goal of smog reduction won't be achieved when all violate the temporal measure). With *acting in favor of group interests* usually being referred to as *cooperation* and *acting in favor of personal interests* as *defection*.

The central question in social dilemma research is how people can be encouraged to cooperate in this kind of situations. Generally, two approaches can be distinguished: the motivational approach and the structural approach (Chen, 1999). The motivational approach focuses on changing people's perception of the situation, and therefore their motives to cooperate. The structural approach focuses on the impact of altering characteristics of the social dilemma itself. (For an overview of research classified according to this distinction, see e.g. Chen, 1999; Kollock, 1998; Mulder, 2004.)

Sanctioning defection is considered to be an example of the structural approach: it changes the social dilemma itself by altering its payoff structure. Experimental research on sanctioning systems in social dilemmas has been carried out in various ways (Mulder, 2004).

Some researchers fined the least cooperative individual(s) (McCusker & Carnevale, 1995; Mulder, van Dijk, De Cremer, & Wilke, 2006a; Van Vugt & De Cremer, 1999; Yamagishi, 1992) or provided participants with the opportunity to punish a defecting group member by subtracting a certain amount of his or her gains (Caldwell, 1976; Fehr & Gächter, 2002, Rapoport & Au, 2001). Other researchers designed the sanctioning system in such a way that the outcomes of defection were unconditionally reduced (Eek, Loukopoulos, Fujii, & Gärling, 2002; Gneezy & Rustichini, 2000; Mulder, van Dijk, De Cremer, & Wilke, 2006b; Mulder, Van Dijk, Wilke, & De Cremer, 2005; Tenbrunsel & Messick, 1999; Wit & Wilke, 1990). Results suggest that there are indeed circumstances under which the presence of a sanction on defection, if large enough, raises the level of cooperation (Caldwell, 1976; Eek, et al., 2002; Fehr & Gächter, 2002; McCusker & Carnevale, 1995; Mulder, et al., 2006a, 2006b; Rapoport & Au, 2001; Tenbrunsel & Messick, 1999; Van Vugt & De Cremer, 1999; Wit & Wilke, 1990; Yamagishi, 1992). The most straightforward explanation of the cooperation-increasing effect of a sanctioning system is that sanctioning defection makes the outcomes of defection less beneficial to an individual, which in turn makes cooperation a relatively more attractive option which encourages people to cooperate. Fining those who exceed the speed limit on highways for example, makes violation of the speed limit less lucrative and, therefore, drivers will be more inclined to comply with the speed limit.

However, inspired by the motivational approach,<sup>2</sup> research has explored more indirect ways through which sanctioning defection increases cooperation. More specifically, it has been argued that sanctioning defection increases people's expectations about other people's cooperativeness which in turn makes people behave more cooperatively. Indeed, on the one hand, Yamagishi (1986; 1992) showed that people tend to favour the introduction of a sanctioning system when there is little trust that other people will cooperate, suggesting that knowing that others will be sanctioned if they defect, results in more confidence that they will

cooperate. On the other hand, expectations about other people's cooperativeness have been shown to relate positively to cooperation (Bruins, Liebrand, & Wilke, 1989; Coombs, 1973; De Cremer, Dewitte, & Snyder, 2001; De Cremer & Van Dijk, 2002; Kerr, 1983; Parks, Henager, & Scamahorn, 1996; Pruitt & Kimmel, 1977; Rapoport & Eshed-Levy, 1989; Robbins, 1995; Schnake, 1991; Yamagishi, 1986). But, there is more to say. Mulder and her colleagues (2006a) showed in a series of studies that sanctioning defection may be a rather paradoxical means to promote cooperation. They refer to Deci (1971); Deci, Benware, & Landy (1974); Deci, Koestner, & Ryan (1999); and Ryan & Deci (2000) and their distinction between intrinsic and extrinsic motivation to distinguish between two bases of trust in others' cooperativeness: (1) trust in others being externally motivated to cooperate and (2) trust in others being internally motivated to cooperate. They agree that a sanctioning system may indeed increase people's expectations about others' cooperativeness, but they argue that the presence of a sanctioning system only increases trust in others being externally motivated to cooperate (i.e., trust in others being motivated by external incentives; note that Mulder et al. seem to take external regulation for extrinsic motivation). Trust in others being internally motivated to cooperate (i.e., trust in others that they will voluntarily choose to cooperate; note that Mulder et al. seem to take autonomous motivation for intrinsic motivation) decreases when there is a sanction on defection. This undermining effect of a sanctioning system on trust in others being internally motivated to cooperate does not impair peoples cooperativeness as long as the sanctioning system is present, since people will expect others to cooperate because of the sanctioning system. However, after removal of the sanctioning system, levels of cooperation of those whose trust was initially high, drop below those of people who have been confronted with the same social dilemma but without having experienced a sanction on defection. Knowing this, one may recommend not to remove a sanctioning system, were it not that sanctioning systems may actually fail to increase

cooperation. Experimental research on social dilemmas has mostly been restricted to variants with only two alternatives: a cooperative and a defective one. Real-life social dilemmas on the contrary do not necessarily need to be dichotomous which may leave sanctioning (only) one defective alternative ineffective to increase cooperation. In other words, sanctioning (only) one out of several alternative ways to defect may not have a cooperation-increasing effect. Mulder, et al. (2006b) presented participants a social trilemma in which they could choose (1) to cooperate, (2) to defect, or (3) to defect in a way that benefited the individual more than cooperation but less than the other defective alternative. Sanctioning those who chose the second option, failed to increase cooperation and collective outcomes because some people turned to the third option which favored the group less than both other alternatives.

As may have become clear by now, installing a sanctioning system does not necessarily promote cooperation and may even have detrimental effects. The goal of this study is to further investigate the impact and consequences of sanctioning defection. Several perspectives are possible. Behavioral effects of sanctions can be approached from a learning psychology point of view (e.g., operant conditioning) or can be explained via their motivational effects. We opted for the second approach. In what follows, self-determination theory (SDT) will be presented because it is a broad motivational theory that discusses the impact of sanctioning on intrinsic and extrinsic motivation.

SDT (Deci & Ryan, 1985, 2000; Ryan & Deci, 2000) is a macro-theory concerning human motivation and personality. People are assumed to be active, growth-oriented organisms rather than reactive mechanisms and are believed to have a natural propensity toward the development of an organized coherence not only among the elements of their psychological makeup but also between themselves and their social world. People are also believed to be naturally inclined to explore, to be curious, to engage in challenging activities, to extend themselves, and to master new skills. However, these growth-oriented tendencies

require the continuous satisfaction of the psychological needs for autonomy, competence, and relatedness. Within SDT, autonomy concerns experiencing a sense of volition and self-endorsement with respect to one's actions (deCharms, 1968; Deci, 1975) and is not to be mistaken for independence, detachment, or individualism. Competence refers to feeling effective in one's actions (White, 1959) and relatedness implies caring for and feeling cared by important others (Baumeister & Leary, 1995; Ryan, 1995). Satisfaction of these psychological needs can be fostered or hindered by one's social context. The interaction between the active organism and its social environment forms a basis for SDT's predictions about behavior, experience, and development.

The reason for engagement in a certain kind of behavior has been shown to affect persistence and behavioral quality. SDT differentiates among five regulatory styles which all reflect qualitatively different reasons to act. A first distinction can be made between intrinsic and extrinsic regulation of behavior. The term extrinsic motivation refers to the performance of an activity in order to attain some separable outcome and, thus, contrasts with intrinsic motivation, which refers to doing an activity for the inherent satisfaction of the activity itself (Ryan & Deci, 2000). Intrinsically motivated behavior is volitional or autonomously regulated by definition. Engagement happens spontaneously and is not triggered by an external force pulling one's strings. However, extrinsically motivated behavior is not necessarily characterized by the lack of personal endorsement. SDT proposes that it can vary greatly in how much it is externally controlled or autonomously regulated (e.g. Ryan & Connell, 1989; Vallerand, 1997) and accordingly subdivides extrinsic motivation into four qualitatively different regulatory styles. External regulation is the first of the four. People feel controlled by external contingencies when their behavior is externally regulated. Engagement is merely a means to attain something they want or to avoid something unpleasant. It's in fact the only type of regulation recognized in operant theory (e.g. Skinner, 1953). Introjected regulation is

a second type of extrinsic motivation. Both external and introjected regulation are characterized by an external perceived locus of causality (E-PLOC; deCharms, 1968). However, with introjected regulation, contingencies have been somewhat internalized. External contingencies have been replaced by internally administered contingent self-worth. Threats of guilt and shame are in charge now. So, while your neighbor might turn in bottles only because he gets a deposit when he does (i.e. behaving through external regulation; external control), you may recycle because you would feel guilty if you don't (i.e. behaving through introjected regulation; internal control). Identified and integrated regulation are the two remaining types of extrinsic motivation. Both are characterized by an internal perceived locus of causality (I-PLOC; deCharms, 1968). People recognize and accept the underlying value of a behavior as their own through the process of identification. However, internalization is not complete until integration has occurred. Identifications do not necessarily concord with the other aspects of one's self and that's where integration comes into play and completes internalization (Pelletier, Tuson, & Haddad, 1997; Ryan, 1995). When the regulation is identified or integrated, the motivation is still extrinsic but autonomous. I may for example volitionally separate glass from my other garbage, not because it is fun doing so (i.e. behaving through intrinsic regulation), but because I personally value taking care of the environment. In doing so, I behave through identified regulation when I for instance drive a big polluting car as well, since taking care of the environment is not consistent with driving a polluting car. However, I behave through integrated regulation when my behaviors are assimilated into my self and brought into alignment with all my values and goals.

The regulatory style closest to the idea of sanctioning defection as a means to get people to cooperate, is external regulation. Pushing people to cooperate by sanctioning them when they don't, may indeed make them cooperate. We therefore hypothesize (1) that people

will cooperate more in a social dilemma when there is a sanction on defection than when there isn't. However, unlike operant theory, SDT stresses that there is a reverse side of the medal (see Figure 1). One may advocate that cooperation through external regulation engendered by a sanctioning system is better than no cooperation at all, but SDT warns us for a paradoxical effect of sanctioning systems.<sup>3</sup> Controlling circumstances, like sanctioning a certain kind of behavior, have been shown to thwart people's need for autonomy and therefore forestall internalization and undermine intrinsic motivation. We therefore hypothesize (2) that people will be less cooperatively when they are presented a second social dilemma without a sanction on defection after having been presented a social dilemma with a sanction on defection, than people who had not been confronted with a sanction on defection before. We furthermore hypothesize (3) that this effect will be mediated by perceived choice during the first social dilemma (see Figure 2) since lack of autonomy induced by a sanctioning system is held responsible for undermining initial autonomous motivation to cooperate.

## Method

### *Participants and Design*

The Web experiment was conducted on a Web site<sup>4</sup> which was linked to the Web Experiment List,<sup>5</sup> Social Psychology Network,<sup>6</sup> and Psychological Research on the Net.<sup>7</sup> 60 (22 male and 38 female) undergraduates participated to fulfill a (partial) course requirement or to get extra course credit. The age of the participants ranged from 18 to 30 years old ( $M = 21.4$ ,  $SD = 3.3$ ). Each participant was randomly assigned to the sanction or no-sanction condition in a split-plot design with *sanction* as a between-subjects factor (sanction, no sanction) and *phase* as a within-subjects factor (first social dilemma, second social dilemma). Participation lasted about 10 minutes.

### *Procedure*

On entering the Web site of the Web experiment, each participant was met by a welcome page. This page introduced the study, announced that it had been approved by the University of Rochester Research Subjects Review Board, and emphasized that responses would be kept confidential. Participants were told that the study was conducted in order to explore the dynamics of decision making. Furthermore, they were informed that participation was voluntary, proceeded anonymously, and that they could discontinue at any time.

After having given consent to participate, random assignment to the sanction or no-sanction condition took place. Participants were told that they would be presented two situations (A and B), which were in fact the first and second social dilemma. They learned that they formed a group together with three other participants and were asked to imagine that each member of their group owned in both situations a certain amount of money which was divided into 10 units. One unit was told to be worth USD 1.20 in situation A and USD 1.00 in situation B for participants in the sanction condition. Participants in the no-sanction condition were told that each unit would be worth USD 1.00 in both situations. Participants were informed that they would have to decide individually what they would do with their units. They could (1) keep their units, (2) keep part of them and invest the other part for the good of their group, or (3) invest all their units for the good of their group. They were told that each unit invested by any group member would rise in value up to USD 2.00 and would be divided evenly among all members of their group. So if a group member chose to invest a unit, all group members, including the one who invested the unit, would get USD 0.50 (i.e. USD 2.00 divided by 4) in return. Keeping a unit instead of investing it would not imply an alteration of its value and those who chose to keep, would simply keep. However, participants in the sanction condition were told that choosing to keep units rather than to invest them would be punished. For each unit that they kept for themselves, the experimenter would take a USD 0.20 fine. In the no-sanction condition, there was no mention of a sanction on keeping units.

Now, participants were asked to think about situation A and to choose the number of units they wanted to invest in the group (ranging from *0 units* to *10 units*). Next, participants were presented 7 additional statements concerning their reason for making the choice they just made (e.g., “I made the decision I made because I wanted to.”) and were asked to indicate how true they were to them on a 7-point answering scale ranging from 1 (*not at all true*) to 7 (*very true*). These items were taken from the Intrinsic Motivation Inventory (IMI; McAuley, Duncan, & Tammen, 1989; Ryan, 1982; Ryan, Mims, & Koestner, 1983) and formed a reliable scale ( $\alpha = .84$ ) which we will refer to as *perceived choice*.

Then, the second social dilemma was presented. As has been told before, this second social dilemma was exactly the same as the first one for participants in the no-sanction condition. Participants in the sanction condition got a social dilemma which differed in two ways from the first one they got: (1) a unit was told to be worth USD 1.00 instead of USD 1.20 and (2) keeping a unit would not be sanctioned anymore. Note that the pay-off structure was held constant in both situations across the sanction and no-sanction condition (see Table 1) (cf. Wit & Wilke, 1990). Again, participants were asked to think about the situation and to choose the number of units they wanted to invest in the group (ranging from *0 units* to *10 units*).

Next, participants were asked to fill out some demographics and after that, the experiment was finished. Participants were debriefed, all details about the experimental set-up were explained, and they were thanked for their participation.

## Results

### *Preliminary analyses*

ANOVA's indicated that ethnicity did not predict perceived choice ( $F(5, 50) = 1.95, p = .10$ ), cooperation in the first phase ( $F(5, 50) = 1.41, p = .24$ ), or cooperation in the second phase ( $F(5, 50) = .71, p = .62$ ). Gender did not predict perceived choice ( $F(1, 56) = .02, p =$

.90), nor cooperation in the first phase ( $F(1, 56) = 2.28, p = .14$ ), but females cooperated less in the second phase ( $M = 5.6$ ) than males ( $M = 7.4$ ) ( $F(1, 56) = 6.43, p = .01$ ).

#### *Primary analyses*

The level of cooperation did not significantly differ between the sanction and no-sanction condition in the first phase ( $F(1, 58) = .70, p = .41$ ) nor in the second ( $F(1, 58) = .22, p = .64$ ). However, perceived choice was lower in the sanction ( $M = 5.3$ ) than in the no-sanction condition ( $M = 6.1$ ) ( $F(1, 58) = 8.23, p = .006$ ). See Table 2. Perceived choice correlates positively with cooperation in the first phase ( $r = .27, p = .034$ ) and second phase ( $r = .29, p = .027$ ).

A repeated measures ANOVA did not indicate a main effect of condition ( $F(1, 58) = .54, p = .47$ ), or phase ( $F(1, 58) = .58, p = .45$ ), nor an interaction between condition and phase ( $F(1, 58) = .21, p = .65$ ). Furthermore, while linear regression analyses showed that condition predicted perceived choice ( $B = -.79, p = .006$ ) and perceived choice predicted cooperation in the second phase ( $B = .69, p = .026$ ), mediational analyses to examine whether perceived choice mediates the effect of condition on cooperation in the second phase did not need to be continued, since condition did not affect cooperation in the second phase ( $B = .33, p = .64$ ). See Figure 3. However, there is nonetheless evidence for an indirect effect of perceived choice on the relation between condition and cooperation in the second phase (*Sobel Test Statistic* = 1.78,  $p = .075$ , two-tailed,  $p = .037$ , one-tailed).

#### Discussion

Results of our study do not show any support for our first hypothesis. The sanctioning system failed to increase cooperation. One can only speculate about why the sanction on defection did not lead to an increase of cooperation, since the kind of data that has been collected is limited and does not allow us to answer that question. Possibly, the sanction on

defection might have been not large enough to have an impact on behavior or the description of the social dilemma and the sanctioning system may have been not appealing enough.

Our second hypothesis got no support either, through which our third hypothesis even did not need to be tested. Participants did not behave less cooperatively when they were presented a second social dilemma without a sanction on defection after having been presented a social dilemma with a sanction on defection, than participants who had not been confronted with a sanction on defection before. The reasoning behind these hypotheses was that the sanctioning system would undermine initial autonomous motivation to cooperate. However, it might have been the case that participants actually had no autonomous motivation to cooperate to start with. Another possibility is that there actually needs to be taken a moderator into account in order to reveal an undermining effect of sanctioning defection on cooperation. Fabes, Fultz, Eisenberg, May-Plumlee, and Christopher (1989) for instance showed that rewarding prosocial behavior may undermine subsequent prosocial motivation in situations where rewards are no longer forthcoming. Rewards undermined children's helping in a free-choice period, but only for children whose mothers felt positive about using rewards.

Based on our results, one may conclude that no additional paradox of sanctioning defection has been revealed. However, while there is no evidence for a direct effect of a sanctioning system on cooperation, there is some for an indirect effect. Sanctioning defection has been shown to make people experience less choice about what to do in a social dilemma which is on its turn associated with less cooperation. Further research is needed, but this may already count as a new glimpse on the dark side of sanctioning systems.

## References

- Baumeister, R., & Leary, M. R. (1995). The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin*, *117*, 497-529.
- Bruins, J. J., Liebrand, W. B., & Wilke, H. A. (1989). About the salience of fear and greed in social dilemmas. *European Journal of Social Psychology*, *19*, 155-161.
- Caldwell, M. D. (1976). Communication and sex effects in a five-person prisoner's dilemma game. *Journal of Personality and Social Psychology*, *33*, 273-280.
- Chen, X. P. (1999). Work team cooperation: The effects of structural and motivational changes. In M. Foddy, M. Smithson, S. Schneider, & M. Hoggs (Eds.), *Resolving social dilemmas: dynamic, structural, and intergroup aspects* (pp. 181-192). Philadelphia: Psychology Press.
- Coombs, C. H. (1973). A reparameterization of the prisoner's dilemma game. *Behavioral Science*, *18*, 424-428.
- Dawes, R. M. (1980). Social dilemmas. *Annual Review of Psychology*, *31*, 169-193.
- deCharms, R. (1968). *Personal Causation*. New York: Academic Press.
- Deci, E. L. (1971). Effects of externally mediated rewards on intrinsic motivation. *Journal of Personality and Social Psychology*, *18*, 105-115.
- Deci, E. L. (1975). *Intrinsic Motivation*. New York: Plenum.
- Deci, E. L., Benware, C., & Landy, D. (1974). The attribution of motivation as a function of output and rewards. *Journal of Personality*, *42*, 652-667.
- Deci, E. L., Koestner, R., & Ryan, R. M. (1999). A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin*, *125*, 627-668.

- Deci, E. L., & Ryan, M. (1985). *Intrinsic motivation and self-determination in human behavior*. New York: Plenum.
- Deci, E. L., & Ryan, M. (2000). The “what” and the “why” of goal pursuit: Human needs and the self-determination of behavior. *Psychological Inquiry*, *11*, 227-268.
- De Cremer, D., Snyder, M., & Dewitte, S. (2001). ‘The less I trust, the less I contribute (not)?’ The effects of trust, accountability and self-monitoring in social dilemmas. *European Journal of Social Psychology*, *31*, 93-107.
- De Cremer, D., & Van Dijk, E. (2002). Reactions to group success and failure as a function of identification level: A test of the goal-transformation hypothesis in social dilemmas. *Journal of Experimental Social Psychology*, *38*, 435-442.
- Eek, D., Loukopoulos, P., Fujii, S., & Gärling, T. (2002). Spill-over effects of intermittent costs for defection in social dilemmas. *European Journal of Social Psychology*, *32*, 801-813.
- Fabes, R. A., Fultz, J., Eisenberg, N., May-Plumlee, T., & Christopher, F. S. (1989). Effects of rewards on children’s prosocial motivation: A socialization study. *Development Psychology*, *25*, 509-515.
- Fehr, E., & Gächter, S. (2002). Altruistic punishment in humans. *Nature*, *415*, 137-140.
- Gneezy, U., & Rustichini, A. (2000). A fine is a price. *Journal of Legal Studies*, *29*, 1-17.
- Kerr, N. L. (1983). Motivation losses in small groups: A social dilemma analysis. *Journal of Personality and Social Psychology*, *45*, 819-828.
- Kollock, P. (1998). Social dilemmas: The autonomy of cooperation. *Annual Review of Sociology*, *24*, 183-214.
- Komorita, S. S., & Parks, C. D. (1994). *Social dilemmas*. Madison: Brown and Benchmark.

- McAuley, E., Duncan, T., & Tammen, V. V. (1989). Psychometric properties of the intrinsic motivation inventory in a competitive sport setting: A confirmatory factor analysis. *Research Quarterly For Exercise and Sport*, *60*, 48-58.
- McCusker, C., & Carnevale, P. J. (1995). Framing in resource dilemmas: Loss aversion and the moderating effects of sanctions. *Organizational Behavior and Human Decision Processes*, *61*, 190-201.
- Mulder, L. B. (2004). *Throwing light on the dark side of sanctions. Sanctioning systems in social dilemmas re-examined*. Unpublished doctoral dissertation, University of Leiden, Leiden.
- Mulder, L. B., Van Dijk, E., De Cremer, D., & Wilke, H. A. M. (2006a). Undermining trust and cooperation: The paradox of sanctioning systems in social dilemmas. *Journal of Experimental Social Psychology*, *42*, 147-162.
- Mulder, L. B., Van Dijk, E., De Cremer, D., & Wilke, H. A. M. (2006b). When sanctions fail to increase cooperation in social dilemmas: Considering the presence of an alternative option to defect. *Personality and Social Psychology Bulletin*, *32*, 1312-1324.
- Mulder, L. B., Van Dijk, E., Wilke, H. A. M., & De Cremer, D. (2005). The effect of feedback on support for a sanctioning system in a social dilemma: The difference between installing and maintaining the sanction. *Journal of Economic Psychology*, *26*, 443-458.
- Parks, C. D., Henager, R. F., & Scamahorn, S. D. (1996). Trust and reactions to messages of intent in social dilemmas. *Journal of Conflict Resolution*, *40*, 134-151.
- Pelletier, L. G., Tuson, K. M., & Haddad, N. K. (1997). Client motivation for therapy scale: A measure of intrinsic motivation, extrinsic motivation, and amotivation for therapy. *Journal of Personality Assessment*, *68*, 414-435.

- Pruitt, D. G., & Kimmel, M. J. (1977). Twenty years of experimental gaming: Critique, synthesis, and suggestions for the future. *Annual Review of Psychology*, 28, 363-392.
- Rapoport, A., & Au, W. T. (2001). Bonus and penalty in common pool resource dilemmas under uncertainty. *Organizational Behavior and Human Decision Processes*, 85, 135-165.
- Rapoport, A., & Eshed-Levy, D. (1989). Provision of step-level public goods: Effects of greed and fear of being gyped. *Organizational Behavior and Human Decision Processes*, 44, 325-344.
- Robbins, T. L. (1995). Social loafing on cognitive tasks: An examination of the “sucker effect”. *Journal of Business and Psychology*, 9, 337-342.
- Ryan, R. M. (1982). Control and information in the intrapersonal sphere: An extension of cognitive evaluation theory. *Journal of Personality and Social Psychology*, 43, 450-461.
- Ryan, R. M. (1995). Psychological needs and the facilitation of integrative processes. *Journal of Personality*, 63, 397-427.
- Ryan, R. M., & Connell, J. P. (1989). Perceived locus of causality and internalization : Examining reasons for acting in two domains. *Journal of Personality and Social Psychology*, 57, 749-761.
- Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development and well-being. *American Psychologist*, 55, 68-78.
- Ryan, R. M., Mims, V., & Koestner, R. (1983). Relation of reward contingency and interpersonal context to intrinsic motivation: A review and test using cognitive evaluation theory. *Journal of Personality and Social Psychology*, 45, 736-750.
- Schnake, M. E. (1991). Equity in effort: The “sucker effect” in co-acting groups. *Journal of Management*, 17, 41-55.

- Skinner, B. F. (1953). *Science and human behavior*. New York: Macmillan.
- Tenbrunsel, A. E., & Messick, D. M. (1999). Sanctioning systems, decision frames, and cooperation. *Administrative Science Quarterly*, *44*, 684-707.
- Vallerand, R. J. (1997). Toward a hierarchical model of intrinsic and extrinsic motivation. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 29, pp. 271-360). San Diego: Academic.
- Van Vugt, M., & De Cremer, D. (1999). Leadership in social dilemmas: The effects of group identification on collective actions to provide public goods. *Journal of Personality and Social Psychology*, *76*, 587-599.
- White, R. W. (1959). Motivation reconsidered: The concept of competence. *Psychological Review*, *66*, 297-333.
- Wit, A., & Wilke, H. A. (1990). The presentation of rewards and punishments in a simulated social dilemma. *Social Behaviour*, *5*, 231-245.
- Yamagishi, T. (1986). The provision of a sanctioning system as a public good. *Journal of Personality and Social Psychology*, *51*, 110-116.
- Yamagishi, T. (1992). Group size and the provision of a sanctioning system in a social dilemma. In W. B. G. Liebrand, D. M. Messick, & H. Wilke (Eds.), *Social dilemmas: Theoretical issues and research findings. International series in experimental social psychology* (pp. 267-287). Oxford: Pergamon.

## Footnotes

<sup>1</sup>Remarks on Dawes' definition of a social dilemma can be found in e.g. Komorita & Parks (1994).

<sup>2</sup>Indeed, the motivational and structural approach can't be absolutely separated because structural solutions may work through their effects on motivation (Chen, 1999).

<sup>3</sup>Other paradoxical effects of sanctioning systems have been shown by Mulder, et al. (2006a) and Tenbrunsel & Messick (1999).

<sup>4</sup><http://www.courses.rochester.edu/surveys/huyghe> (no longer active)

<sup>5</sup><http://genpsylab-wexlist.unizh.ch/>

<sup>6</sup><http://www.socialpsychology.org/expts.htm>

<sup>7</sup><http://psych.hanover.edu/research/exponnet.html>

Table 1

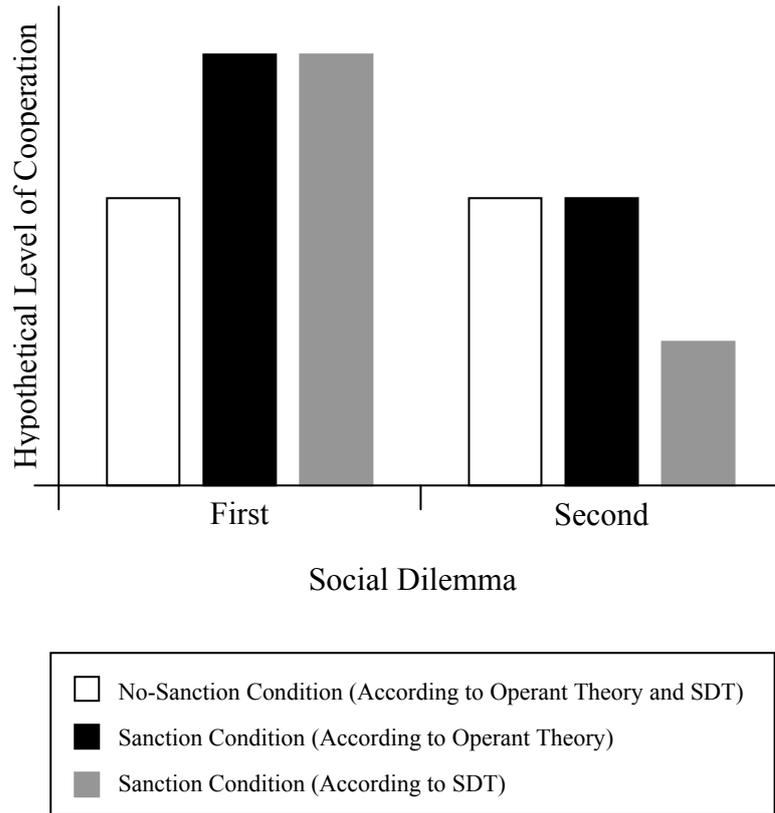
*Value of a Unit (in USD) as a Function of its Allocation by a Group Member*

Phase	Unit allocated to:	In condition:	Value to self:	Value to each of the three other group members	Collective outcomes of unit (value to self + [3 x value for other])
1 <sup>st</sup>	Group	all	0.50	0.50	2.00
	Self	no-sanction	1.00	0.00	1.00
		sanction	1.20 - 0.20	0.00	1.00
2 <sup>nd</sup>	Group	all	0.50	0.50	2.00
	Self	all	1.00	0.00	1.00

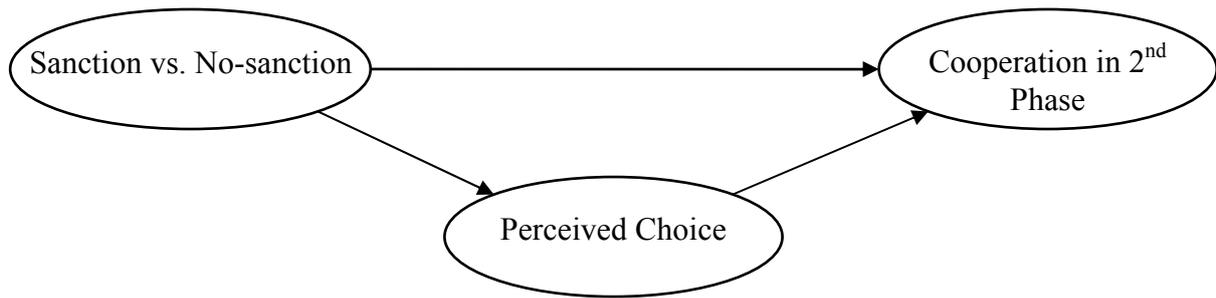
Table 2

*Means and Standard Deviations for Cooperation and Perceived Choice*

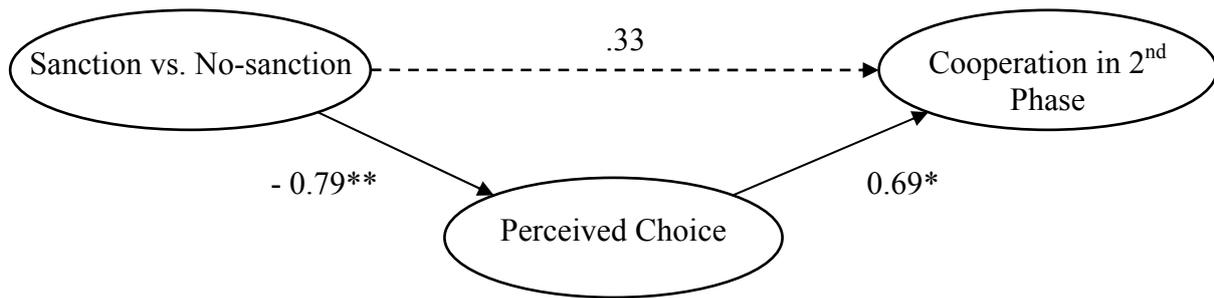
Variable	Phase	Sanction	No-sanction	$F(1, 58)$	$p$
		condition	condition		
		$M (SD)$	$M (SD)$		
Cooperation	1 <sup>st</sup>	6.3 (3.1)	5.7 (2.8)	.70	.41
	2 <sup>nd</sup>	6.4 (3.0)	6.1 (2.5)	.22	.64
Perceived choice	-	5.3 (1.1)	6.1 (1.0)	8.23	.006



*Figure 1.* Hypothetical levels of cooperation based on predictions of operant theory on the one hand and SDT on the other hand. Both sanction and no-sanction condition include two social dilemmas, with a sanction on defection in the first social dilemma of the sanction condition.



*Figure 2.* Hypothetical mediating role of perceived choice in explaining the relation between having experienced a sanction on defection and cooperation in the second phase.



*Figure 3.* Unstandardized regression coefficients for the mediation model of perceived choice in explaining the relation between having experienced a sanction on defection and cooperation in the second phase.

\* $p < .05$ . \*\* $p < .01$ .

