

On the Demand and Effectiveness of Reward in Voluntary Contribution Mechanisms. A Note

.....
Daniela Grieco
Università Bocconi
(Italia)

1. Introduction

Peers' punishment of free riders has been shown to represent a powerful instrument to induce cooperation in public good games (Fehr & Gaechter, 2000) and common pool resources (Ostrom, Walker & Gardner, 1992). Interestingly, the symmetric instrument of rewarding high contributors has exhibited significantly lower demand and efficacy in promoting cooperation. This result is robust across experimental settings (Dickinson, 2002; Andreoni, Harbaugh & Vesterlund, 2003; Walker & Halloran, 2004; Sefton, Shupp & Walker, 2006; Vyrastekova & Van Soest, 2008). Experimental subjects emerge - in general - not to elicit the option of rewarding cooperators as a device to drive peers' behaviour in favour of the group, neither altruistic behaviours are significantly stimulated by the possibility of being rewarded. As Offerman (2002) points out, "hurting hurts more than helping helps."

This note aims at discussing the reasons why individuals - in situations of anonymous interaction with strangers - do not decide to remunerate pro-social behaviours, whereas

sanctioning anti-social ones is quite commonly used and efficient in its impact on contributions levels. This result is suboptimal because rewards present both monetary and non-monetary benefits. Like sanctions, rewards are costly; however, they represent a transfer from the rewarder to the rewardee, whereas sanctions are costly for both the punisher and the punishee. Furthermore, and crucially, rewards do not induce in the recipient negative feelings like the anger and the desire of revenge that might undermine the cohesion of a group in case of punishment. An institution where rewards are fruitfully used as instruments to promote cooperation could represent a form of "virtuous" institution that rules out any constrictions and threats and where interactions occur in a more helpful environment. The absence of perverse effects and its nature of transfer are the reasons why an efficient use of rewards could represent a significant improvement in the devices put at work to prevent the decay in contributions. We believe that a better comprehension of this phenomenon could help in detecting the conditions where rewards could play a significant role.

In what follows, we address the issue of understanding the reasons behind the low demand and the ineffectiveness of reward resulting in economic experiments. The economic perspective is enlarged with insights from cognitive and social psychology. In particular, we analyse how rewards are perceived from a cognitive point of view and how the possibility of being rewarded does affect the motivations driving altruistic behaviours. Two distinct explanations for low demand

and ineffectiveness are discussed. On the one hand, rewarding is not elicited by subjects because it does not enable to yell out the negative emotions that instead can be expressed by assigning a sanction. On the other hand, rewards are ineffective in promoting cooperation because they are received by subjects holding already an intrinsic motivation to contribute: the extrinsic motivation built up by a reward or a sanction could be differently effective if an intrinsic motivation is already at work.

2. Economic Experiments On Reward

Experimental studies show the emergence of cooperation under conditions that strongly favour egoism. In real context of repeated, *non-anonymous* interactions, unlike any other species, humans cooperate with non-kin in large groups: this behaviour is puzzling from an evolutionary perspective because cooperators face individual costs to confer benefits on unrelated group members. In laboratory experiments, people interacting *anonymously* elicit positive contributions levels that decline as decisions are repeated: when given the opportunity to prevent the decay in contributions by costly punishing non-cooperators, subjects *do punish* even in one-shot interactions. Ethnographic data suggest that such altruistic punishment helps to sustain cooperation in human societies (Boyd, Gintis, Bowles & Richerson, 2003).

A bunch of studies focuses on the effectiveness of reward as an alternative or addi-

tional device to promote cooperation and prevent the decay in contributions. In two-stage games, a reward

is an increase in payoffs that a second-stage player imposes on a first-stage player, whereas a punishment is a decrease in payoff at the second stage. Under peculiar conditions, providing subjects with both the opportunity to reward *and* to punish results in higher efficiency, although more often this does not occur. Andreoni, Harbaugh and Vesterlund (2003) examine demands for rewards and punishments in a proposer-responder game. The proposer first makes an offer to split a fixed-sized pie; the responder is given or is *not* given a costly option of increasing or decreasing the proposer's payoff. The results show substantial demands for both punishments and rewards: however, rewards *alone* have little influence on cooperation, whereas punishments have some. When rewards and sanctions are combined, the effect on cooperation is relevant, suggesting that rewards and punishments are complements in producing cooperation. More specifically, offers exceed the Nash equilibrium level in all three treatments and, interestingly, this effect is *stronger in the rewards treatment* than in the sanctions treatment. However, offers are highest in the treatment where both instruments are available to the responder.

Sefton, Shupp and Walker (2003) contrast the institutions of reward and punishment with respect to their impact on cooperation and efficiency in the context of voluntary contributions mechanism with linear payoffs. In the reward treatment, contributions decrease over time reaching a level below the one ob-

served in the absence of opportunities to reward: the opportunity to reward results insufficient to sustain cooperation; in contrast, sanctions succeed in sustaining public goods provision at a level above the one observed in the absence of sanctioning opportunities, and so sanctioning appears to be a more effective mechanism for sustaining contributions. However, opportunities to sanction initially result in a loss of efficiency, as the direct costs associated with sanctioning prevail over the effect of increased contributions. The authors also observe differences in the dynamic patterns of rewarding and punishing behaviour: whereas subjects initially use rewards more frequently than sanctions, over time the use of rewards declines at a faster rate than the use of sanctions, so that in later rounds rewards are used less frequently than sanctions. These results capture the complexities involved in using a reward system for sustaining cooperation: “one might argue that a successful reward system requires continued use of rewards, and those rewards must be in the form of transfers from those allocating less to the public good to those allocating more. In our experiments, however, it was those subjects who allocated relatively more to the group account who tended to give more rewards” (*ibidem*). The significant decay in rewards across decision rounds suggests that *groups may have difficulty in maintaining a rewards system* and that subjects seemed to lack a clear focal point or consensus in regard to where rewards should be targeted. Moreover, rewards seem to work only with a critical mass of altruistic individuals that contribute in the first stage and then reward who contributed in

the second stage: however, if this is true, there would be no need of building an extrinsic motivation by means of rewarding or punishing because it would be effective for people who already behaved altruistic. In a similar flavour, Dickinson (2001) studies teams of four members in a laboratory environment where giving effort towards a “team goal” is captured by eliciting voluntary contributions towards the provision of a public good. The efficiency-improving properties of four distinct environments is tested: monetary prizes given to high contributors versus monetary fines assessed to low contributors, where being high or low contributor is defined first in terms of absolute contributions and then in terms of contributions relative to abilities (called “handicapping”). The use of carrot (positive) and stick (negative) incentives are methods of increasing effort among members of work teams. The results show that both carrot and stick can significantly increase efficiency (i.e., contributions): more specifically, handicapped incentives promise the highest efficiency levels and, when handicapping is not used, *penalties are more effective than prizes*.

Walker and Halloran (2004)’s study involves a one-shot, voluntary contribution mechanism and finds that, whereas some subjects are willing to reward and sanction others at a private cost, *the opportunity to reward or sanction is ineffective in facilitating cooperation* relative to previous experiments in which repeated game environment is employed. Andreoni, Harbaugh and Vesterlund (2003)’s and Sefton, Shupp and Walker (2003)’s experiments present anonymous, strangers treatments: however, a reputation effect about

peers' attitude towards cooperation, or at least a mechanism of learning about the rules of the game, seems to emerge and strengthen the role of punishment and, in a lower extent, of reward. This effect disappears in a one-shot game like Walker and Halloran (2004)'s.

Vyrastekova and Van Soest (2007)'s work studies the effectiveness of costly rewards in mitigating excess extraction in a standard Common Pool Resource game experiment. Two treatments are implemented: in the first, rewards are a pure transfer from one player to the other ("transfer reward"); in the second, the benefits of receiving a reward are higher than the cost of providing it ("net positive reward")¹. In this perspective, the authors suggest an explanation for Andreoni, Harbaugh and Vesterlund (2003)'s results on reward effectiveness: the difference with the ineffectiveness emphasized by other experiments' results might lie in the *parameterization*. In fact, Andreoni, Harbaugh and Vesterlund (2003)'s study uses either 1:5 sanctions, or 1:5 rewards, or both. Vyrastekova and Van Soest find that, when the benefits of receiving a reward exceed the cost of providing it, the use of extraction tokens in the CPR game is significantly closer to the social optimum than in case of transfer rewards. This difference in effectiveness arises due to the difference in profitability of receiving rewards (and consequently in the opportunity costs of not receiving them), and due to differences in who rewards whom. When there are positive net gains from exchanging reward tokens, estab-

lishing bilateral exchange by means of reward is in the interest of both the players. As subjects who cooperated have no incentives to send reward tokens to free-riders, bilateral exchange is established among subjects using the same number of extraction tokens. This represents for free-riders an incentive to decrease the number of extraction tokens used to attract rewards. Such considerations are absent when there are zero net gains from exchanging rewards, as in the case of the transfer treatment. This kind of analysis is important because in many situations, a rewardee's evaluation of a reward is unlikely to be equal to the costs incurred by the rewarder. In case of rewards taking the form of monetary transactions, this can be due to subjects' differences in income or wealth; in case of rewards taking the form of the exchange of goods or services, marginal utilities may differ considerably between individuals.

Vyrastekova and Van Soest (2007)'s contribution is very relevant in the topic and sheds light on the reasons why sanctions are generally more effective than rewards. However, it lacks in investigating explicitly the difference between sanctions and rewards. Although the role of the parameterization and the prize-to-fee ratio are significant, a change in the impact ratio has been found to substantially affect the effectiveness of sanctions (Carpenter, 2006; Egas & Riedl, 2005) as well. As Casari (2007) shows, sanctions are sensitive to parameterization. Rewards behave coherently in the same directions: to prove lower effectiveness of the rewards, what would matter is a relatively higher sensitivity

to parameterization that has not been documented yet.

3. The Psychological Literature On Positive Reinforcement

Psychological studies strongly differ from economics in predicting the role of rewards versus sanctions. The psychological literature conceives rewards (sanctions) as positive (negative) reinforcements of actions. A huge amount of studies within the operant tradition (since the seminal work by Skinner, 1953 on) has documented that *extrinsic rewards do shape behaviour significantly*.

Psychological studies on punishment and reward are concerned with adaptive learning and the contrast between intrinsic and extrinsic motivation. Altruistic and pro-social behaviour are forms of conduct that derive from intrinsic motivations but that can be stimulated by means of extrinsic devices like reward or punishment. Four types of outcomes can occur in operant research (Burgess and Bushell, 1969): (1) positive reinforcement, i.e. the addition of an appetitive stimulus to a situation; (2) negative reinforcement, i.e. the removal of an aversive stimulus from a situation; (3) positive punishment, i.e. the addition of an aversive stimulus to a situation; and (4) negative punishment, i.e. the removal of an appetitive stimulus from a situation. The two types of reinforcement can be grouped under the general label of “reward,” while the two types of punishment can be grouped under the general label of “cost” (Gray & Tallman, 1987). By 1971, hundreds of studies within the oper-

ant tradition had established that extrinsic rewards can control behaviour. In fact, when administered closely subsequent to a behaviour, rewards were reliably found to increase the likelihood that the behaviour would be emitted again, an effect that persisted as long as the reward contingency was operative. When rewards were terminated, the likelihood that the behaviour would be emitted eventually returned to the pre-reward baseline. This general finding led to the widespread advocacy of rewards as a motivational strategy, and behaviour-change programs based heavily on the use of rewards were introduced into a variety of applied settings (Deci, Koestner & Ryan, 1999).

A wide range of studies has explored the effects of various types of rewards like tangible rewards such as money (Deci, 1971; Deci, 1972a, Deci, 1972b), positive feedback or verbal rewards. Additional studies explored the effect of external effects such as classroom climates (Deci, Schwartz, Sheinman & Ryan, 1981) and the combined effect of reward and interpersonal contexts (Ryan, Mims & Koestner, 1983). Interpersonal contexts represent the social ambience of settings like homes, classrooms, work groups and play a role in the extent to which people within the context feel pressured to behave in a particular way. When studied in laboratory experiments, the interpersonal climate is usually instantiated as the interpersonal style used by the experimenter to administer rewards (Ryan, 1982). An important work by Miller, Butler and McMartin (1969) on social power emphasizes the role of reward and punishment as a form of social control whose strength and effect de-

rive from the ability to deliver reward and punishment. The authors remark an asymmetry between reward and punishment based on two arguments. First, considerations on notions like exchange (e.g. Blau, 1964), reciprocity (Gouldner, 1960) and equity (Adams, 1964) suggest that reward and punishment implemented by high level of power subjects should imply distinct (instead of similar) outcomes due to the fact that high power subjects that use reward should reach their goals because their use of rewards is perceived as altruistic and generous. On the contrary, high power subjects that exercise their power by means of sanctions could have difficulties in getting their ends and, if they do, “in their case their dessert are aversive stimuli” (Miller, Butler & McMartin, 1969). Furthermore, reward and punishment differ in their effect on the learning and maintenance of human behaviour (Estes, 1944; Skinner, 1953): mild punishment seems to be less permanent in its effects, is more likely to produce unwanted emotional behaviour than does reward and the behaviour that it elicits may be hard to specify. However, when subject face stochastic punishment or reward, subjective inferences on the perceived probability of being punished or rewarded play a role. The huge amount of work generated from Tversky and Kahneman and their colleagues (e.g., Kahneman et al., 1982), suggests that a given probability of punishment may be a greater deterrent against crime than an equal probability of rewards for choosing a non-criminal activity. This follows since the probability of negative punishment for a non-criminal activity would be greater, given that success at non-criminal

activity was a rare event; in such situations the stick may be more powerful than the carrot.

4. Discussion

The insights deriving from the experimental studies and the psychological investigations reported above shed light on the discrepancy in demand and effectiveness between rewards and sanctions. In this section, the possible explanations of this phenomenon are presented and discussed.

An explanation for low demand: Motivational factors. Psychological studies focus on the fact that an extrinsic device such as reward might undermine intrinsic motivation to cooperate. As emphasized by Charness (1999), recent models of non-pecuniary motives to pro-social behaviour can be classified as either altruism-based, equity-based, or reciprocity-based. One explanation is altruism, where people care not only about their own material well-being but also about the material well-being of others: an agent is assumed to assign a constant weight to the welfare of every other individual. Another approach expands the altruism principle by incorporating distributive concerns: these models, known as equity or fairness models, one’s regard for another person’s monetary well-being depends on the other person’s monetary payoff relative to one’s own⁴. Finally, reciprocity theories expand the principle of altruism in a different direction, asserting that attention for someone else’s payoff depends on how much altruist the other is perceived to be: the prin-

ciple of reciprocal altruism emphasizes the relationship between one's action and one's beliefs about the intentions of the other agents⁵.

4 Current models in this category are Bolton and Ockenfels (1997) and Fehr and Schmidt (1997). 5 The leading model in this category is Rabin (1993), while Levine (1996) offers an alternative formulation.

These issues of altruism and selfishness drive some crucial questions on our evolutionary origins, our social relations, and the organization of society. Experimental evidence indicates that, although agents are self-interested and maximize their earnings, human altruism is a powerful force. However, there is heterogeneity across individuals: economic agents are not generally selfish but might belong to further types, i.e. be altruists or reciprocators. In experiments, different player "types" seem to emerge: (1) some subjects are standard selfish maximizers, confirming that the *homo oeconomicus* exists; (2) other subjects act altruistically, as if they were driven by other-regarding preferences; (3) a third group of players are willing to non-strategically punish and/or reward others, i.e. are driven by some form of "strong reciprocity", based on evaluation of the opponent's intentions (Camerer & Fehr, 2006). It is this interaction between altruist and selfish individuals that drives human cooperation (Fehr & Fischbacher, 2003). The fact of belonging to a specific type is related to the presence or absence of intrinsic motivations (of being altruistic): only altruists have an intrinsic motivation that moves contribution to a public good. Following the psychological debate on the role of positive, extrinsic reinforcement on

intrinsic motivations, we argue that a possible explanation of the lower effectiveness of reward relies on the fact that rewarding an altruistic behavior undermines the presence of an intrinsic motivation to altruism. If people are altruist and believe that their pro-social behavior is interiorly grounded and this should be true also for peers, for instance due to the importance of a common project, the presence of a device that should stimulate pro-social behavior could eliminate this feeling towards the project. If a reward is directed to pay off altruists, they would probably argue that there should be some reason not to be altruistic. Any spontaneous effort is ruled out. On the contrary, sanctions may build an (extrinsic) motivation of behaving altruistically for individuals who have no intrinsic motivation and, in absence of sanctions, would free-ride. In this perspective, it would be interesting to explore more specifically whether the ineffectiveness of reward (contrasted to sanction) is due to the fact that it undermines the intrinsic motivation of a purely altruistic choice. Hence, a comparison between altruists and selfish/reciprocators in their reactions of altruists to reciprocators to rewards should be carried out: only altruists' behavior should be affected by such a reduction in intrinsic motivation.

An explanation for ineffectiveness: The role of emotions. A relevant source of asymmetry between reward and punishment lies in the type and strength of emotions that originate them and that they provoke in targets. Fehr and Rockenbach (2003) point out that in human societies, social order and cooperation rely on

both the use of rewards and sanctions, “which ensures the compliance of self-

interested actors, and on the presence of people willing to perform altruistic acts”. Their experiment present a setup based on dyadic interaction in trust game. In each round of the experiment two mutually anonymous subjects (an investor and a trustee) are involved. First, the investor has the chance of choosing a costly trusting action; then the trustee is informed about the investor’s action and can honour the investor’s trust by taking a costly cooperative action. If the investor chooses a trusting action and the trustee responds cooperatively, both players increase their monetary payoff, but the trustee has the option of not honouring the investor’s trust, saving the costs of cooperation but causing the investor to be worse off. Consequently, if the investor expects that the trustee will not honour his trust, he will not choose a costly trusting action at the beginning of the game. Therefore, the subjects face a dilemma because trust and cooperation are advantageous for both subjects but the trustee faces the temptation not to cooperate and, therefore, the investor is tempted not to trust. The main results in Fehr and Rockenbach’s experiments is that sanctions’ effectiveness relies on the motivation that has previously driven the sanction: whereas altruistically motivated sanctions implemented for the benefit of the group enhance cooperative behaviour, sanctions that are imposed to enforce an unfair distribution of resources have the opposite effect. Previous research on public goods experiments has shown that altruistic punishment is a highly effective means of enforcing

cooperation. Sanctions in the public good context enhance cooperation only when the punishment of free-riders is an altruistic act that is considered as morally legitimate, whereas rewards should not present the same need of being legitimated.

Offerman (2002)’s experiments present simple proposer-responder games that demonstrate that responders are willing to depart from own-earnings maximization by rewarding more generous proposers or sanctioning less generous proposers. The main hypothesis tested in this paper is that negative intentions provoke stronger reciprocal responses than positive intentions. The potential asymmetric effect of positive and negative intentionality is assessed in the “hot response game”. In this game the first mover makes a choice between a helpful and a hurtful choice where the helpful choice increases the payoff of the first mover by less guilders than the hurtful choice and increases the payoff of the second mover by the same number of guilders. The second mover observes the choice of the first mover, before choosing between a cool, remunerative response and a hot, reciprocal response. People easily feel insulted after a hurtful choice. In accordance with this conjecture, subjects experience stronger negative emotions when a hurtful choice is intentional than when it is unintentional. In addition, they show more restraint in responding to their emotion when the hurtful choice was unintentional. Both an increase in the negative emotion experienced and a decrease in the restraint of responding to the emotion explain the strong effect of negative intentionality on the probability of reciprocation. Negative intentionality matters

considerably more than positive intentionality.

Over the years, numerous studies of law and economics exposed the effect of legal sanctions on individual behaviour, whereas neuroeconomics studies have been contributing to indicate that the desire to punish is often strong and non-strategic, and that it is a largely automatic response triggered by emotional forces (see Singer and Fehr, 2005). Among others, the positron emission tomography has been used to examine the neural basis for altruistic punishment of defectors in an economic exchange in an experiment where subjects could punish defection either symbolically or effectively. Symbolic punishment did not reduce the defector's economic payoff, whereas effective punishment did reduce the payoff. Effective punishment, as compared with symbolic punishment, activated the dorsal striatum, which has been implicated in the processing of rewards that accrue as a result of goal-directed actions. Moreover, subjects with stronger activations in the dorsal striatum were willing to incur greater costs in order to punish. These findings support the hypothesis that people derive satisfaction from punishing norm violations and that the activation in the dorsal striatum reflects the anticipated satisfaction from punishing defectors. This result is confirmed by Casari and Luini (2008)'s finding that punishment is expressive, and individuals do not consider sanctions inflicted by peers as a substitute of their own sanction, although sanctioning is costly.

They seem to accrue an intrinsic pleasure to punish by themselves, and the possibility to

set out negative emotions by punishing deviators should play a role.

In this perspective, it would be interesting to explore whether the effect of the negative emotions in strengthening the efficacy of sanctions compared to rewards could be removed by implementing virtuous institutions where punishment cannot be used to implement anti-social behaviours like targeted or blind revenge that grounds on negative emotions. This idea is confirmed by Haigner, Kocher and Sutter (2006), who analyze an experimental public goods game in which group members can endogenously determine whether they want to supplement a standard voluntary contribution mechanism with the possibility of rewarding or punishing other group members. They find a large and positive effect of endogenous institutional choice on the level of cooperation in comparison to exogenously implemented institutions: democratic participation rights enhance cooperation in groups. Moreover, with endogenous choice, groups typically vote for the reward option.

References

- Adams, J.S. (1965). Inequities in social changes, *Advances in Experimental Social Psychology*.
- Antoci, A. and Zarri, L. (2008). On Punishing Non-Punishers. When (Not So) Nice Guys Deserve the Stick, not the Carrot, Science24.com
- Andreoni, J., Harbaugh, W. and Vesterlund, L. (2003), The Carrot or the Stick: Re-

- wards, Punishments, and Cooperation, *American Economic Review*, 93 (3), 893-902. 15
- Blau, P. M. (1964), *Exchange and Power in Social Life*. John Wiley & Sons, New York.
- Boyd, Gintis,
- Bowles and Richerson, (2003), The evolution of altruistic punishment, *PNAS*, March 18, 100 (6), 3531–3535.
- Burgess, R.L. and Bushell, D. (1969), *Behavioral Sociology: The Experimental Analysis of Social Process*. Columbia University Press.
- Camerer, C. and Fehr, E. (2006). When does “Economic man” dominate social behaviour?, *Science*, 311, 5757, 47-52.
- Casari, M. (2005). On the design of peer punishment experiments, *Experimental Economics*, 8, 107-115.
- Charness, G. and Haruvi, E. (2002). Altruism, equity, and reciprocity in a gift-change experiment: An encompassing approach, *Games and Economic Behavior*, 40(2), 203-231.
- Dawes, R.M. (1980). Social dilemmas, *Annual Review of Psychology*, 31, 69-93.
- Deci, E. L. (1971), Effects of externally mediated rewards on intrinsic motivation, *Journal of Personality and Social Psychology*, 18, 105-115.
- Deci, E. L. (1972a). Effects of contingent and non-contingent rewards and controls on intrinsic motivation, *Organizational Behavior and Human Performance*, 8, 217-229. 16
- Deci, E. L. (1972b). Intrinsic motivation, extrinsic reinforcement, and inequity, *Journal of Personality and Social Psychology*, 22, 113-120.
- Deci, E. L., Koestner, R. and Ryan, R.M. (1999), A meta-analytic review of experiments examining the effect of extrinsic rewards on extrinsic motivation, *Psychological Bulletin*, 125(6), 627- 668.
- Dickinson, D. (2000), The Carrot vs. the Stick in Work Team Motivation, Working Paper ERI 2000-06 Utah State University.
- Fehr, E. and Fischbacher, U. (2003), The nature of human altruism, *Nature*, 425, 785-91.
- Fehr, E. and S. Gächter (2000), Fairness and retaliation: The Economics of reciprocity, *Journal of Economic Perspectives*, 14, 159-81.
- Fehr, E. and Rockenbach, B. (2003), Detrimental effects of sanctions on human altruism, *Nature*, 137-140.
- Fehr, E. and Schmidt, K. (1999), A theory of fairness, competition, and cooperation, *Quarterly Journal of Economics*, 114, pp. 817–868
- Flache, A. (1996), *The Double Edge of Networks. An Analysis of the Effect of Informal Networks on Cooperation in Social Dilemmas*, Amsterdam: Thesis Publishers
- Gouldner, A.W. (1960), The norm of reciprocity: A preliminary statement, *American Sociologic Review*, 25(2), 161-178. 17
- Gray, L.N. and Tallman, I. (1987), Theories of choice: Contingent reward and punishment applications, *Social Psychology Quarterly*, 50 (1), 16-23.
- Haigner, S., Kocher, M. and Sutter, M. (2006), Choosing the Stick or the Carrot? Endogenous Institutional Choice in Social Dilemma Situations, *CESPR Discussion Paper*.

- Kahneman, D., J. Knetsch and R. Thaler (1986)., Fairness and the assumptions of economics, *Journal of Business*, 59, pp. S285–300.
- Miller, N. Butler, D.C. and McMartin, J.A. (1969) ,The ineffectiveness of punishment power in group interaction, *Sociometry*, 32 (1), 24-42.
- Molm, L.D. (1997), *Coercive Power in Social Exchange*.
- Offerman, T. (2002), Hurting Hurts more than Helping Helps, *European Economic Review*, 46(8), 1423-1437.
- Ostrom, E., J. Walker and R. Gardner (1992). „Covenants with and without a sword: Self- governance is possible’, *American Political Science Review*, 86, 404-417.
- 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. 18
- Ryan, R. M., Mims, V. and 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.
- Sefton, M., Shupp, R. and Walker, J. (2003)., The effect of rewards and sanctions in provision of public goods’, Indiana University Working Paper.
- Sell, J., Chen, Z., Hunter-Holmes, P. and Johansson, A. C. (2002)., A Cross-Cultural Comparison of Public Good and Resource Good Settings, *Social Psychology Quarterly*, 65 (3), 285-297.
- Skinner, B. F. (1953). *Science and human behavior*. New York: Macmillan.
- Son, J. and Sell, Y. (1995)., Comparing public goods with common pool resources: Three experiments, *Social Psychology Quarterly*, 60 (2), 118-137.
- Vyrastekova, J. and Van Soest, J. (2007)., On the (in)effectiveness of rewards in sustaining cooperation, *Experimental Economics*, 11, 53-65.
- Walker, J.M. and Halloran, M. (2004)., Rewards and sanctions and the provision of public goods in one-shot settings’, *Experimental Economics*, 7, 235–247.