

Ono Academic College

The Business School Working Paper series

Working Paper No. 2008-26

Is punishment of Non-cooperators essential? The role of Altruistic Punishment

Miriam Dishon-Berkovits¹ and Richard Berkovits²

¹ *Ono Academic College, Kiryat Ono, Israel*

² *Department of Physics, Bar-Ilan University, Ramat-Gan 52900, Israel*

Abstract

In human society unrelated individuals cooperate with each other; for example by donating old clothes to charity, by conserving water or by not littering a public park. Recent studies reveal that *Altruistic Punishment* is a key component in maintaining cooperation. Altruistic Punishment is the human behavior of punishment of noncooperators even at a personal cost to the punisher. *Altruistic Punishment* is seen by some researchers (e.g., Barclay, 2006; Fehr & Gächter, 2002; Gintis, Bowles, Boyd & Fehr, 2003; Guzmán, Rodríguez-Sickert & Rowthorn, 2007; O’Gorman, Wilson & Miller, 2004) as an evolutionary puzzle because it is implausible for the punisher to expect that his or her costs will be repaid.

In the current study we investigate why it is at all necessary to punish noncooperators, and why admonishment is not sufficient to deter noncooperators and to promote cooperation among strangers.

STUDY 1

To this aim we created a mathematical model of a large population with the size of n individuals ($n = 10,000$). There are two behavioral types: Activists and Noncooperators. The rules of the model are as follows: m packets of tokens ($m = 1,000,000$) are randomly distributed among the n individuals. In 80% of the cases each packet contains only one token which the receiver can keep. In 20% of the cases the packet contains 5 tokens. The receiver is permitted to keep only one token. Activists distribute the remaining 4 tokens among strangers (randomly). Noncooperators discard the 4 tokens (i.e., waste a common good that could be shared by society).

Next, consider what happens when another individual observes the behavior of the Noncooperator. A fellow Noncooperator would do nothing, and the 4 tokens would be removed from the game (hence, a resource of potential common use is wasted). An Activist would pay one token from his personal assets (of tokens) in order to (randomly) distribute the 4 tokens among the population (Activists and Noncooperators alike). We name this behavior *Altruistic Admonishment* because the Noncooperators suffer no punishment but only receive rebuke (note, they can still keep their token).

RESULTS

Consider Panel A of Figure 1. Results indicate that as the percentage of Activists in society increases, the mean number of tokens of each behavioral group (Activists and Noncooperators) increases as well. Yet, on average Noncooperators receive a larger number of tokens than Activists do. At the same time above a critical percentage of Activists in society, the Activists group also benefits: each individual Activist receives more tokens than he or she would have received without *Altruistic Admonishment*.

DISCUSSION

Three conclusions are warranted: First, Noncooperators gain the most benefits: they are not being punished for wasting potential common resources, as well as benefiting from the Activists' behavior. Hence, Noncooperators in this model are in essence free riders. Second, a certain percentage of Activists within a population is beneficial for society. Third, above

a critical percentage of Activists in society, although each Activist incurs a personal loss in a single interaction of punishing a Noncooperator, the Activists group as a whole *incurs benefits in the long run*. (Note that in the model Activists are acting in an altruistic fashion because choosing the second alternative - i.e. being a Noncooperator - would only bring benefits to oneself).

STUDY 2

In Study 2 we asked what would happen if Noncooperators were to be punished (and not only rebuked). We created a mathematical model identical to the first, with these changes: When an Activist observes a Noncooperator discard 4 tokens, the Activists pays 1 token to punish the Noncooperator by taking away his packet of 5 tokens (4 of which are then redistributed among the rest of the population).

RESULTS AND DISCUSSION

Consider Panel B of Figure 1. When the percentage of Activists in a society is small results are similar to those of study 1. When the percentage of Activists in a society reaches a critical threshold the trend changes: Activists earn more tokens than do Noncooperators (which are being punished). This is why punishment is vital to society, and without it cooperation can not exist.

STUDY 3

In Study 3 we asked what would happen if Activists were to be rewarded (and Noncooperators only rebuked). We created a third mathematical model identical to the first, with the following changes: When an Activist observes a Noncooperator discard 4 tokens, the Activists pays 1 token to reward the Activists ingroup - by taking away the 4 tokens from the Noncooperator and distributing them only among the Activists group. (Note: Noncooperators are not punished).

RESULTS AND DISCUSSION

Consider Panel C of Figure 1. Results are similar to study 2: When the percentage of Activists in society is small it is not beneficial to the individual to be an Activist, because the probability of paying a token to punish a Noncooperator is larger than the probability that someone else would punish a Noncooperator and you would gain (one of the 4 redistributed tokens). Yet, once there is a critical mass of Activists things change: Noncooperators are punished not directly, but indirectly, by the fact that their group is not included in the redistribution of the goods.

GENERAL DISCUSSION

Consider Figure 1. A striking result is that the average number of payoffs in society is the same regardless of Altruistic Punishment, Altruistic Admonishment or reward for activists. We have tested it for different values of n , m , number of tokens in the package and percentage of packages larger than one. Although the quantitative results change, the qualitative behavior does not. This creates an intriguing puzzle: It is common in many societies to punish individuals who abuse common goods. Yet, if the average gains for society stay constant regardless of punishment, what is the goal of punishment? The answer may be a moral one: to deter free riders. The results have many applied implications to situations of conserving common property resources and public goods. A theoretical implication of the study is that in certain situations *Altruistic Punishment* does not violate rational-choice theory.

REFERENCES

- Barclay, P. (2006). Reputational benefits for altruistic punishment. *Evolution and Human Behavior*, 27, 325-344.
- Fehr, E., & Gächter, S. (2002). Altruistic punishment in humans. *Nature*, 415, 137-140.
- Gintis, H., Bowles, S., Boyd, R., & Fehr, E. (2003). Explaining altruistic behavior in humans. *Evolution and Human Behavior*, 24, 153-172.
- Guzmañ, R. A., Rodriguez-Sickert, R., & Rowthorn, R. (2007). When in Rome, do as the Romans do: the coevolution of altruistic punishment, conformist learning, and cooperation. *Evolution and Human Behavior*, 28, 112-117.

O'Gorman, R., Wilson, D. S., & Miller, R. R. (2005). Altruistic punishing and helping differ in sensitivity to relatedness, friendship, and future interactions. *Evolution and Human Behavior*, 26, 375-387.

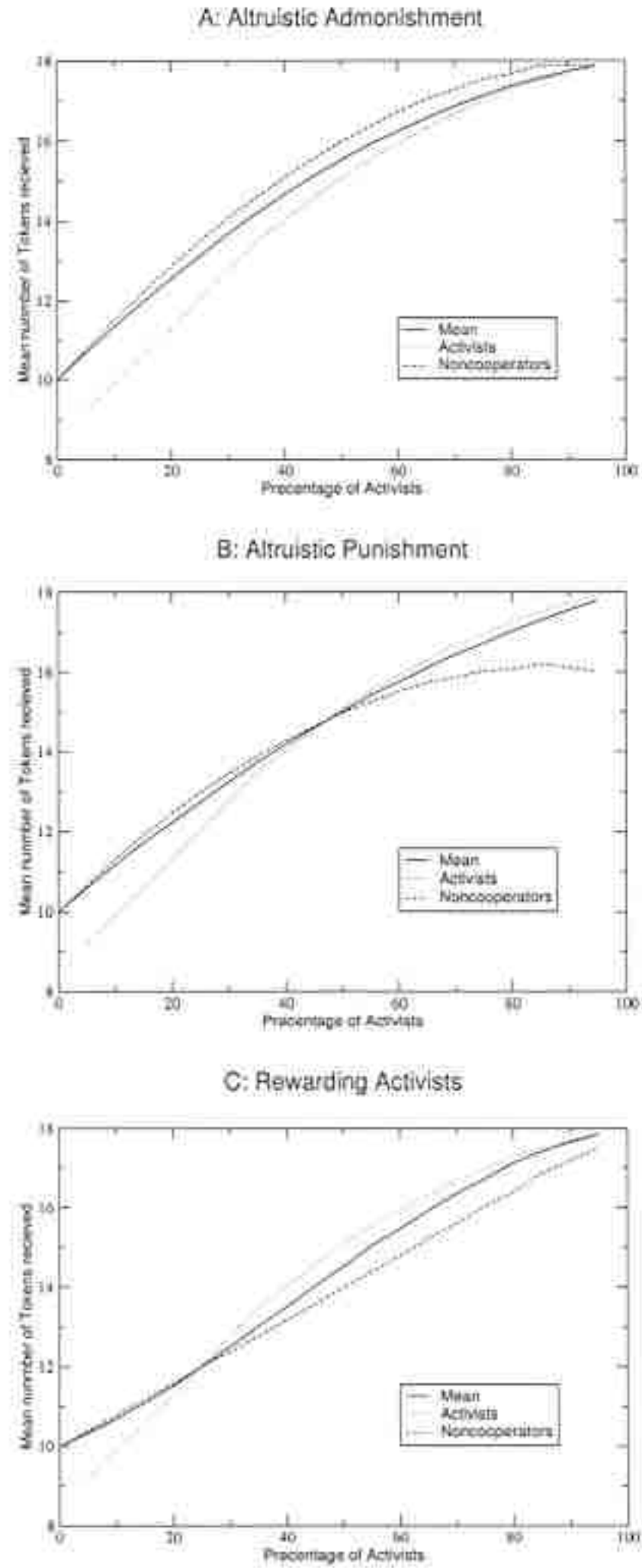


FIG. 1: Mean number of tokens earned as a function of the percentage of activist in the society