

The Pleasure of Being Nasty

by KLAUS ABBINK* and ABDOLKARIM SADRIEH**

* *University of Amsterdam*

** *University of Magdeburg*

April 2008

Abstract

Many laboratory experiments show that people are often altruistic or care for fairness. We present data that reveal a darker side of human nature. We introduce the joy-of-destruction game. Two players each receive an endowment and simultaneously decide on how much of the other player's endowment to destroy. Subjects play this game repeatedly. In one treatment, subjects can hide their destruction behind random destruction. In this treatment, money is destroyed in almost 40% of all decisions. We attribute this behavior to a visceral pleasure of being nasty. Under full information destruction is also observed, but rare. In this treatment acts of destruction are followed by immediate retaliation.

Keywords

Spite, nastiness, money-burning, anti-social behavior

JEL Codes

C72, C90, D82, F51

Acknowledgements

We thank Jeroen Udo for help in conduction the experiments, and Anick Bosmans and Rik Pieters for providing us with advertisement evaluation questionnaires. We are grateful to Benedikt Herrmann, Reinhard Selten and seminar participants in Maria Laach, Nottingham and Valencia for helpful comments.

Contacting the Authors

Klaus Abbink
CREED, Faculty of Economics and Business
University of Amsterdam
Roetersstraat 11
1018 WB Amsterdam
The Netherlands
+31 20 525-4229 (fax -5283)
www.uva.nl/creed
k.abbink@uva.nl

Abdolkarim Sadrieh
Faculty of Economics and Management
University of Magdeburg
Postbox 4120
39016 Magdeburg
Germany
+49 391 67-18585 (fax: -12120)
www.wv.uni-magdeburg/e-business
sadrieh@ovgu.de

1. Introduction

The defining figure of neoclassical economic theory, the *homo oeconomicus*, has long been challenged by both experimental economics and social psychology. People are, as we now know from a plethora of experimental studies, not own-payoff maximizing egoists, but compassionate social beings who care about fairness and the well-being of others. They reject unfair offers in ultimatum games (Güth et al. (1982)), share their endowment in dictator games (Forsythe et al. (1994), Hoffman et al. (1994)), reciprocate in gift exchanges (Fehr et al. (1993), Berg et al. (1995)), or punish free-riders in public goods games (Fehr and Gächter (2000)), to name just a few instances of fairness-driven, pro-social behavior.¹ Human beings, so does the experimental evidence overwhelmingly suggest, are of far better character than neoclassical economists have given them credit for.

So far, so nice. But there is a danger of overstating the kindness of human nature. Behavioral economists study such deviations from the orthodoxy because they are interesting and, thus, focus on environments in which they hope to find altruistic behavior. This may create a selection bias, leading to a neglect of environments in which human behavior can indeed be selfish, greedy, uncooperative, or even plain nasty. There can be no doubt that nastiness exists. People suffer random violence from complete strangers. Property is arbitrarily vandalized. Malicious computer viruses are circulated solely to do harm. One can argue that such anti-social behavior is an anomaly, committed by individuals from the fringes of society. But quite possibly a pleasure of being nasty is present in the mind of everybody. This question is the focus of this study.

We introduce the *joy-of-destruction game*. The basic setup of the game is simple. There are two players. Each player first earns an endowment. The expected size of the endowment is equal for both players. Both players can mutually and simultaneously destroy each other's endowments. Destruction is costless and entails no material benefit for the destroying party. Since no pecuniary, fairness, or reciprocity motives are present, destruction is most likely due to pure spite and nastiness. Note that other conventional motives to choose destruction are also removed from the game. We avoid the experimenter demand effect and the boredom effect by embedding the destruction choices into a much more cumbersome and time-consuming task, with which the subjects earn their endowments.

We are also interested in the dynamics of play in this environment, which is why pairs of subjects play the game repeatedly. We wish to test two competing hypotheses. The opportunity to retaliate could trigger the escalation into an ongoing pointless vendetta. It

¹ The space in this paper, indeed any paper, would not suffice to list all the relevant literature in this field. So we only mention the pioneering studies and send our apologies to all other contributors.

could, however, also have a deterrent effect, such that subjects refrain from destroying money for fear of retaliation. The results show that the latter is the case.

We use two variants of the game. In the complete information setting (*open treatment*), the destruction is ex-post perfectly observable. Hence, after destruction there is common knowledge about the act and the extent of destruction. In the incomplete information setting (*hidden treatment*), the destruction action is veiled by an additional random destruction. The targeted person can only observe the total damage, but typically cannot identify which part was due to the action of the destroyer and which part due to nature. Thus, it is possible to hurt someone else without being recognized.

The results show very little destruction in the open treatment. Destruction rates start already low and then fall to zero. However, the comparison with the hidden treatment shows that this is due to fear of retaliation rather than genuine kindness. When people can hide their nasty actions behind the veil of random destruction, chosen destruction rates shoot up. Two out of every five decisions (39.4%) destroyed money. This shows that many individuals enjoy a pleasure of being nasty, even in our subject pool of generally well-mannered students.

To our knowledge, this is the first experimental study that examines the pleasure of being nasty in its pure form, but it does not completely stand in isolation. Despotic behavior has been observed in public good games with punishment. Though the norm is for cooperators to punish free-riders, there is also a good deal of perverse punishment, i.e. selfish individuals punishing contributors (Anderson and Putterman (2005)). Gächter et al. (2004) find that this kind of punishment is particularly pronounced among subjects who have been socialized in the rough economic climate of the post-Soviet world. In the complex environment of these games it is of course difficult to isolate single motives. Like in our study, Zizzo and Oswald (2001) set up a game in which participants can reduce others' income without gain. In their experiment subjects indeed burn substantial money, but they do so to equalize payoffs. So the seemingly spiteful act stems from a "good" motive, namely aversion to unfairness. In our game we have removed inequality aversion as a possible explanation for nasty behavior.

2. The experimental design

The experiment was conducted in two sessions with 40 subjects at the CentERlab of Tilburg University. It was run as pen-and-paper experiment. Each subject was allowed to participate in one session only and no subject had participated in experiments similar to the present one. The subjects were undergraduate students, who had been recruited through the laboratory's email recruiting system. Subjects interacted anonymously and were paid confidentially.

To avoid boredom and, more importantly, experimenter demand effects the joy-of-destruction game was nested into a much more time-consuming evaluation task, which subjects performed individually. The task was to view and evaluate full-page ads published in popular Dutch magazines. The questionnaire used in the evaluation task was developed by the

marketing department as part of a PhD thesis project. Subjects received a fixed payment of 80 euro cents for each completed evaluation questionnaire in the open treatment, €1.20 in the hidden treatment (the higher payment accounted for the random destruction). Note that because subjects had to actually work (fill-out lengthy questionnaires) for their endowments, we also avoid the house-money effect. The number of evaluation questionnaires that were given to a subject varied randomly between 1 and 3. The procedure was as follows. Each subject drew a card out of a deck of three cards. The card drawn indicated which number of ads and evaluation questionnaires the subject received.² After having completed all evaluation tasks of a round, subjects were asked to indicate what part of the income of the other subject they would like to destroy. Destruction choices were restricted to multiples of 10 cents between 0 and 80.

We aimed at playing 10 rounds of evaluation and destruction tasks. In the hidden treatment time constraints did not allow us to play out the full 10 rounds, such that the experiment had to be shortened to eight rounds. The matching of subjects for the destruction tasks was anonymous and random, but remained fixed in all rounds. Before the first and the sixth round, the subjects completed mood questionnaires for the marketing study.

The random destruction in the hidden treatment was implemented using cards drawn from a deck of 17 differently colored cards, where each color was associated with a destruction level between zero and €1.60, in steps of €0.10. Subjects did not know which card carried which value until the very end, such that the random draw could be credibly implemented without revealing the outcome during the game. Note that the total destruction was simply the sum of the destructions inflicted by Nature and the partner. Thus, hiding behind Nature was only partially possible. Very high total destruction would reveal to a subject that some of the destruction must have come from the partner.

3. Results

Figure 1 shows the frequency of destruction over the course of the 8 or 10 rounds of the experiment. The figure shows that overall frequencies in the open treatment are low (on average in 8.5% of all decisions money was destroyed). Some destruction can be observed in the early rounds, but destruction rates quickly fade away.

In the hidden treatment, however, destruction frequencies are surprisingly high. On average, 39.4% of all decisions involve the destruction of at least some of the partner's endowment. The figure shows no sign of falling over time, but destruction remains constantly close to the

² It can be argued that the random endowment creates some inequality and thus subjects may destroy money to reduce this inequality. However, subjects were told only their own endowment, not the one of their partners. Further, we do not see different destruction rates depending on the own endowment, so we can rule out inequality reduction as a motive for destruction.

40% mark. The difference between the two treatments is significant both for all rounds and for the first round separately (Fisher’s two-sample randomization test, $\alpha = 0.05$ one-sided).

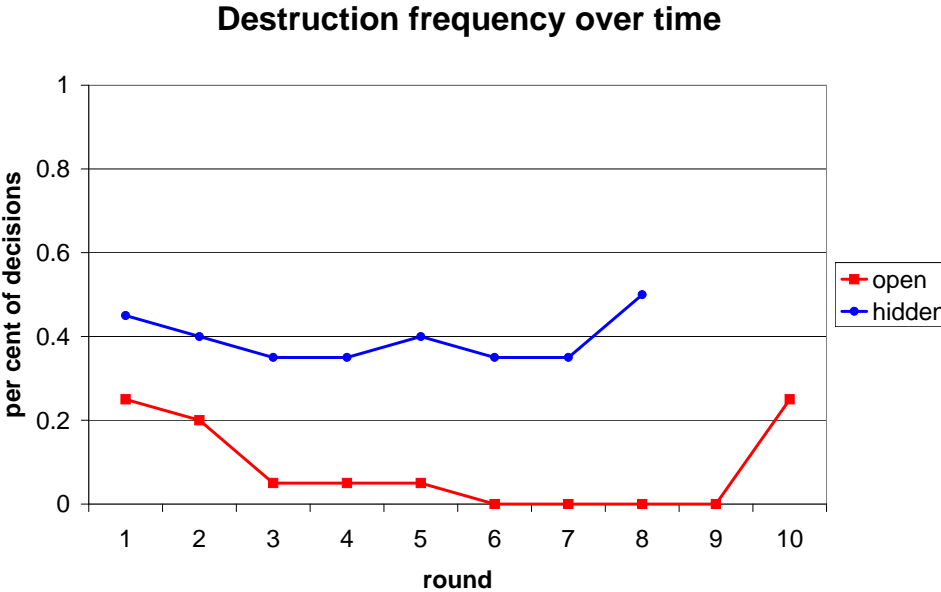


Figure 1

Both treatments show a neat end-game effect, as destruction frequencies shoot up in the final round. This is reminiscent of the breakdown of cooperation in the final round of social dilemma games. This effect is most often interpreted as a selfish act to maximize own payoff. In our game there is of course nothing to gain from burning the other player’s money.

Figure 2 shows the evolution of destruction in money units (euros). The general tendencies are the same. In the open treatment, destruction is almost negligible, while in the hidden treatment substantial amounts are being burned (in total 20.4% of the maximum allowed).

In each round, each individual’s endowment was determined by the number of ads the subject had to evaluate. The question arises whether subjects condition their destructive acts on the endowment they receive. Figure 3 shows that this is not the case. No systematic pattern can be detected. A subject receiving a low endowment can expect to be poorer than the partner (with a two-third probability), so the subject may wish to destroy to establish (expected) equity (like in Zizzo and Oswald (2001)). The lack of any such effect refutes this explanation. In the hidden treatment, destruction is even highest when the own endowment is the same as the opponent’s expected endowment, though no difference is significant. This result thus refutes the possible explanation of an aversion to expected inequality.

Destruction over time

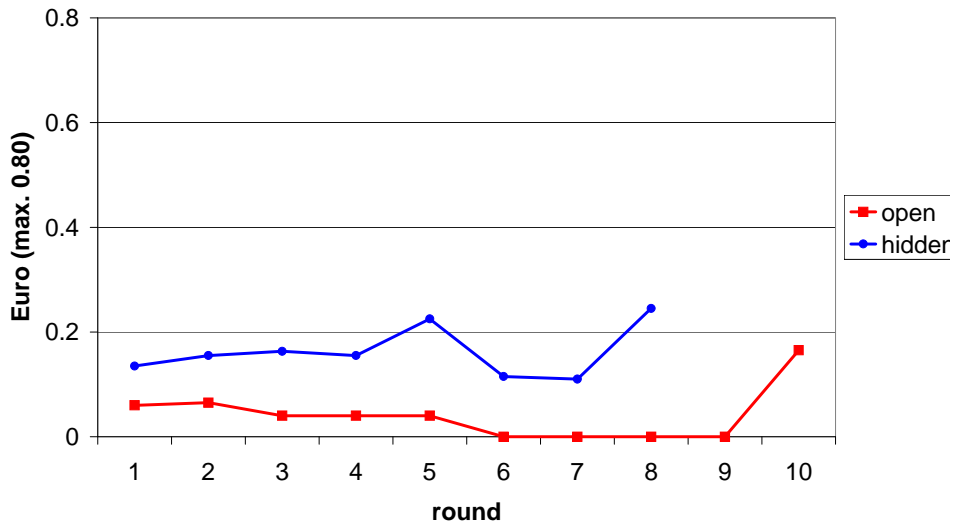


Figure 2

Destruction as a function of endowment

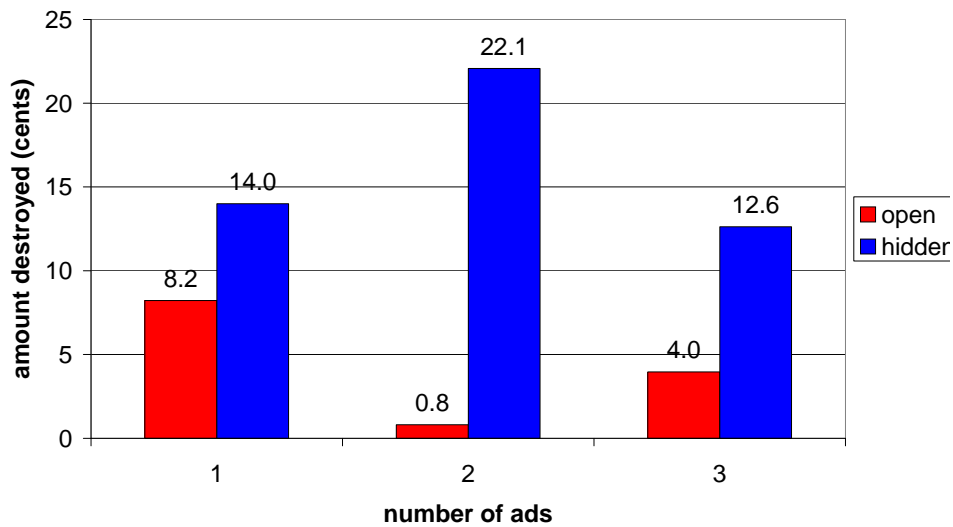


Figure 3

Finally, we check whether subjects in the hidden treatment conditioned their destruction choices on the total destruction they observed in the previous round. Recall that high levels of total destruction make some human component more likely, very high levels even reveal that some part must be of human origin. However, figure 4 shows that subjects do not noticeably respond to previous round feedback. There is no correlation between the observed suffered destruction and the destruction inflicted on the partner.

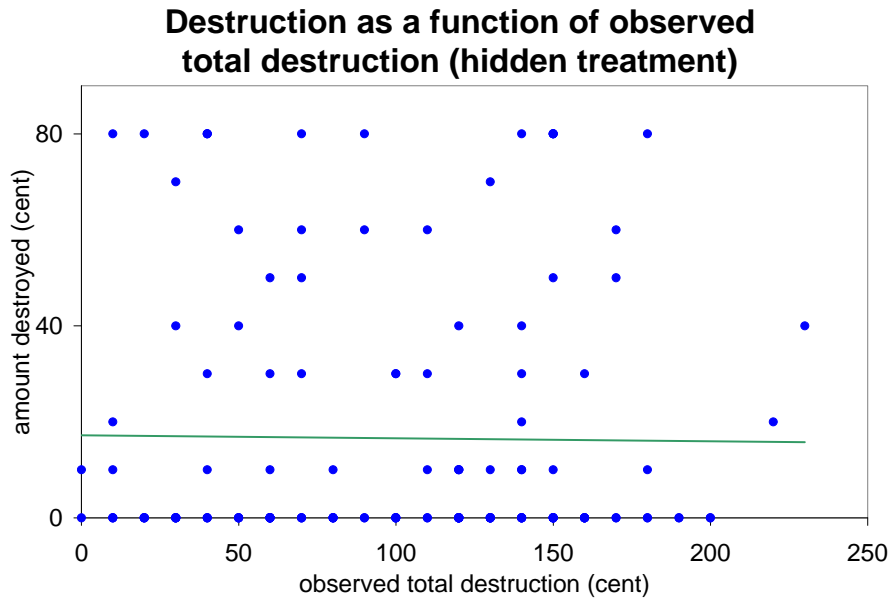


Figure 4

4. Discussion

We introduce the joy-of-destruction game. In this game, players can burn one each other's money, but we have removed all conventional reasons to do so. No material gain can be achieved, no wrongdoing be punished, no inequality reduced. Nevertheless, we observe a substantial incidence of nasty behavior in our hidden treatment, where spiteful acts could be covered by random destruction. When destruction cannot be hidden, it rapidly goes away, but the difference between the treatments shows that this is due to fear of retaliation, not kindness.

We interpret the high destruction rates in the hidden treatment as indication of an emotional pleasure of being nasty. Other explanations are of course possible. An obvious candidate is the experimenter demand effect. Subjects may destroy money just because they feel the experimenter expects them to do so. In our design, we have taken great care to minimize such effects, by nesting the joy-of-destruction game into a bigger task unrelated to social preferences. Further, it seems implausible that the experimenter demand effect would work differently across treatments, so it could not explain the substantial treatment differences we observe right from the start of the experiment.

Individuals burning money could be motivated by comparative payoffs. They could wish to increase their relative standing compared to the opponent. However, proponents of fairness utility models (e.g. Fehr and Schmidt (1999), Bolton and Ockenfels (2000)) usually find that people are inequality averse, disliking unfavorable as well as favorable inequity. We also observe no effect of the own endowment, which renders this explanation incomplete.

Subjects destroying money might be driven by a motive of "pre-emptive retaliation". They may expect the other person to destroy money and "respond" by doing so themselves. If this is so, it would at least imply that individuals expect a pleasure of being nasty in others.

Perhaps our results raise more questions than they answer. The most pressing one, though beyond the scope of this paper, is how to reconcile our results with the abundance of phenomena that prove the “good” side of human nature. Is it the same people, who cooperate in dilemma games, but hurt others in our game? Or are there “good” and “bad” people among our subjects, such that the good people cooperate in other games and just do not destroy in our game, while the bad people hurt others in our game and just behave selfishly in other games? Are social preferences context dependent? To answer these questions more work is needed and our experiment can only be a small step.

References

- ANDERSON, C and L. PUTTERMAN (2005). Do Non-strategic Sanctions Obey the Law of Demand? The Demand for Punishment in the Voluntary Contribution Mechanism. *Games and Economic Behavior*, forthcoming.
- BERG, J., J. DICKHAUT, and K. MCCABE (1995): Trust, Reciprocity and Social History. *Games and Economic Behavior*, 10, 122-142.
- BOLTON, G.E., and A. OCKENFELS (2000): ERC - A Theory of Equity, Reciprocity and Competition. *American Economic Review*, 90:1, 166-193.
- FEHR, E., and S. GÄCHTER (2000): Cooperation and Punishment in Public Goods Experiments, *American Economic Review* 90, 980-994
- FEHR E., G. KIRCHSTEIGER, and A. RIEDL (1993): Does Fairness Prevent Market Clearing? An Experimental Investigation, *Quarterly Journal of Economics* 108: 437-459.
- FEHR, E., and K.M. SCHMIDT (1999): A Theory of Fairness, Competition, and Cooperation. *Quarterly Journal of Economics*, 114, 817-868.
- FORSYTHE, R., J. HOROWITZ, N. SAVIN and M. SEFTON (1994): Fairness in simple bargaining experiments. *Games and Economic Behavior*, 6, 347–369.
- GÄCHTER, S., B. HERRMANN, and C. THÖNI (2004): Trust, voluntary cooperation, and socio-economic background: Survey and experimental evidence, *Journal of Economic Behavior and Organization*, 55, 505-531.
- GÜTH, W., R. SCHMITTBERGER, and B. SCHWARZE (1982): An Experimental Analysis of Ultimatum Bargaining. *Journal of Economic Behavior and Organization*, 3, 367-388.
- HOFFMAN, E., K. MCCABE, K. SHACHAT, and V.L. SMITH (1994): Preferences, Property Rights, and Anonymity in Bargaining Games. *Games and Economic Behavior*, 7, 346-380.
- ZIZZO, D., and A.J. OSWALD, (2001): Are People Willing to Pay to Reduce Others’ Incomes? *Annales d’Economie et de Statistique* 63-64, 39-62.

Appendix: Instructions for participants

Thank you for participating in this experiment.

The experiment will last approximately 2 hours.

You will be paid after the experiment. No other participant will know how much you earned. You will be paid € 2,00 for your participation plus any additional earnings that you have in the experiment.

Please, do not talk during the experiment.

The experiment consists of 10 rounds of decision making.

Before the first round, you will be randomly matched with another participant. You will not learn the identity of the participant you are matched to, but you know that this will be the same person for all 10 rounds.

In each round, you first randomly draw a card with one of the three numbers, 1, 2, or 3.

The number on the card you draw determines the number of tasks that you will have in this round. To complete each task you must rate an advertisement on five different aspects, according to the scales on the evaluation sheet. The advertisements are presented to you on separate pieces of paper. You will only receive as many ads as the number of tasks that you have drawn. For each completed task you will receive € 1.20.

After you have completed your task(s), you have the opportunity to destroy up to € 0.80 of the income of your matched player, but you do not have to destroy any part. Your destruction choice must be a multiple of 10 Cents from the set {0, 10, 20, 30, 40, 50, 60, 70, 80}. Please, indicate your destruction choice by marking the appropriate box on your decision sheet.

[Hidden treatment only: Next, you randomly draw a card out of a deck of 17 cards. The cards are labeled with 17 distinct colors: *Dark Blue, Dark Brown, Dark Green, Dark Orange, Dark Pink, Dark Red, Dark Violet, Dark Yellow, Light Blue, Light Brown, Light Green, Light Orange, Light Pink, Light Red, Light Violet, Light Yellow, and White.*

Each color is assigned to a number between 0 and 160 in steps of 10, i.e. 0, 10, ..., 150, 160. This assignment has been made randomly before the experiment for each of the rounds separately and is kept secret. This means that before a round has ended, you will have no idea which color is assigned to which number. After the experiment, you can check the complete assignment of colors to numbers for each round, when we reveal the color-number-assignment-sheet that is taped to the blackboard.

The number assigned to the color that you draw determines what amount of your payoff that is automatically destroyed. So, if you draw the color assigned to the number 10, then 10 Cents of your payoff are automatically destroyed; if your color is assigned to 20, then 20 Cents of your payoff are automatically destroyed; and so on. Only if your color is assigned to 0, no part of your payoff is automatically destroyed.]

Once all participants have made their decisions for a round, you will be informed which part of your income was destroyed and what amount is left as your net earnings for that round.

[Hidden treatment only: Note, however, that you will not be informed which part of the income destruction is due to the card you drew yourself and which part is due to the destruction choice made by the participant who is matched to you.]

This same procedure will be repeated another 9 rounds.

Before the first round and before the sixth round, a short questionnaire will be handed out to you, which we ask you to please complete.

After the last round, we ask you to remain seated. You will receive a final questionnaire, which we also ask you to please complete.

Meanwhile, we will call the participants out of the lab one by one and pay them their total earnings in cash.