Emotion Regulation and Punishment in Public Goods Experiments

A thesis submitted by

Xinxin Lyu

In partial fulfillment of the requirements for the degree of

Master of Science

in

Economics

Tufts University

Graduation month and year: May, 2016

Advisor: Laura Gee

Abstract

This paper reports experiment results of a 3x2 Voluntary Contribution Mechanism (VCM) game with punishment played in partner-matching groups. The first dimension is groups' aggressiveness levels (homogeneous high, homogeneous low, and heterogeneous aggressive groups) and the second dimension is treatments with or without emotion regulation. We ask two questions: whether emotion regulation can increase people's punishment efficiency and increase group welfare; and how it takes effects in groups of different aggressiveness levels. We find that emotion regulation reduces punishments and increases groups' emotional and monetary welfares for all types of groups. We also find that people of different aggressiveness types contribute differently.

Acknowledgement

I would like express my sincere gratitude to my advisor, Professor Laura Gee, for her patient and detailed guidance throughout my two years master's study. She helped me explore my interests in behavioral economics and offered me the precious opportunity to conduct a well-designed behavioral experiment. There was a time conflict between her behavioral economics course and my required course. In order to help me learn more about the relevant topics, she offered to record all her presentations and send the audio files to me.

I would like to thank my readers, Professor David Garman and Professor Heather Urry. They offered tremendous help to me. Professor Garman was in charge of the administrative work when Professor Gee was on leave in the fall semester. Professor Urry gave me many useful advices from psychology perspective. I would also like to thank Professor Kyle Emerick. Working as his research assistant practiced my Stata skills and advanced my understanding in econometrics. Finally, I want to express my greatest gratitude to all my friends and families for their emotional support and love.

Table of Contents

Introduction

In public economics, a long lasting research interest is how to effectively deter free riders. Free riding on public goods occurs because everyone can benefit from the public goods whether or not they personally contribute. The Voluntary Contribution Mechanism (VCM) game with punishment (Fehr and Gächter, 1999) is designed to study this specific question. Punishment seems to be an effective way to increase contributions to the public goods. Its positive effect has been verified by many studies (Fudenberg and Pathak, 2010; Masclet et al.,2003; Spraggon et al., 2015; Masclet and Villeval, 2008; Nikiforakis and Normann, 2008; Decker et al., 2003; Andreoni and Gee, 2012; Ouss and Peysakhovich, 2013; etc).

There are two stages in a VCM game with punishment. In stage 1, subjects are asked to decide how many tokens they want to contribute to the public goods. Each token kept by a subject gives a higher payoff to only that subject. While each token contributed to the public goods gives a lower payoff to all the group members. In stage 2, subjects see other group members' contributions to the public goods and decide how much they want to punish their other group members by reducing their payoffs. Traditionally, it costs $1 to reduce the payoff of another group member by $3. Clearly, from a group welfare perspective, this kind of punishment is highly costly or even detrimental if the benefit of increased public contributions is outweighed by the loss from punishing. There are indeed more reports of decreased group welfare as measured by net earnings after paying punishment costs (Dreber et al., 2008; Gächter and Fehr, 1999; Fehr and Gächter, 1999;

Gächter et al., 2008; Egas and Riedl, 2008) than increased ones (Masclet et al., 2003; Denant-Boemont et al., 2007; Nikiforakis and Normann, 2008).

Why might welfare be reduced? One theory may be that there exists "over-punishment" in the experiment. We define over-punishment as an excessive usage of punishment, when its overall cost has exceeded the benefit of raised public contributions. The rise of over-punishment is related to subjects' emotions (Bosman and Van Winden, 2002; Joffily et al., 2014; Masclet et al., 2003; Casari and Luini, 2012; Müller and Duersch, 2013; Spraggon et al., 2015;Drouvelis and Grosskopf, 2014;Hopfensitz and Reuben, 2009).

If emotions cause over-punishment, then regulating subjects' emotions should increase punishments' efficiency and decrease excessive punishment. Few papers have explored this area. One relevant paper is by Dickinson and Masclet (2015) who study emotion venting's effect on group welfare. Subjects are asked to release their negative emotions in different ways before making punishment decision. They find that optimal punishment requires a moderate level of emotional arousal and that too much emotional reduction actually decreases the punishment's efficiency. A natural follow up question is whether people can adjust to an appropriate emotional level once they are told to do so.

In this paper, we instruct people to regulate their emotions and adopt a neutral attitude in their decision-making. Moreover, we want to know whether emotion regulation is more beneficial to certain types of people. Some people may be more prone to emotional arousal than others, and thus may benefit more from emotion regulation. For example, a high school football player is more likely to over punish his lazy teammates compared with a graduate student who faces his silent classmates in a study group.

As a result, we design a 3x2 experiment, where the first dimension is the group's aggressiveness and the second dimension is whether the group has emotion regulation. We seat subjects in homogeneous high, homogeneous low, or heterogeneous aggressiveness groups. All subjects play the VCM game for 10 periods with same group members (partner-matching). Subjects are asked to adopt a neutral attitude in the regulation treatment groups or to follow their spontaneous emotions in the no-regulation treatments. We find that emotion regulation increases group welfare materially and emotionally. High and low aggressive subjects do not punish differently and both reduce their punishment under emotion regulation. Emotion regulation benefits equally either homogeneous high or homogeneous low groups. We also find that high aggressive subjects contribute more than low aggressive subjects.

Literature Review

The Voluntary Contribution Mechanism game (VCM), also known as public goods game, is designed to study the conflict of interest between the individuals and the group. Former studies have shown that punishments can effectively deter free-riders in VCM games. Fehr and Gächter (1999) first introduced the punishment stage and find that punishment increases the public contributions significantly regardless of whether the game is played in a perfect stranger or a partner-matching setting. Later studies alter the punishment mechanism[1] in different ways and mostly establish a positive correlation between the group public contributions and the availability of punishment options.

Despite a clear rise in public contributions, punishment's effects on net earnings are more ambiguous. Existing studies report more detrimental effect of punishment than beneficial ones. For example, Fehr and Gächter (1999) report that compared to groups without punishment options, there is an initial payoff loss for both stranger- and partner-matching groups. A relative payoff's gain happens in later periods, but the group average still suffers from welfare loss pooling all decision periods together.

Several reasons might cause the frequent reports of welfare loss.

[1] Some influential changes include: unobservable punishment, where the punished can only see their received punishments in the final period (Fudenberg and Pathak, 2010), non-monetary punishment (Masclet et al., 2003, Spraggon et al., 2015), alterations of the cost-to-punishment ratio (Masclet and Villeval, 2008; Nikiforakis and Normann, 2008) , the existence of a third party punisher (Decker et al., 2003; Andreoni and Gee, 2012; Ouss and Peysakhovich, 2013; Ouss and Peysakhovich, 2013), etc.

First, the experiment length is not long enough to allow the welfare gain to offset the initial severe punishment cost. Based on this point, Gächter et al.(2008) extend the experiment from 10 periods to 50 periods and find that the punishment in a partner-matching group has a long-term positive effect on group's final earnings. Nikiforakis and Normann (2008) change the exchange rate of punishment among partner-matching VCM game and find that the more severe the punishment is, the earlier will the group experience welfare improvement.

Second, since punishment decisions are made at the same time, there might be unnecessarily repeated punishment points assigned to the same free-rider from different group members in one period. Following this logic, there are a number of studies looking for alternative centralized punishers to coordinate the punishment. Relevant studies in this area include Decker et al.(2003) and Andreoni and Gee (2012).

Last but not the least, over-punishment exists. People do not use punishment homogeneously to promote public contributions. According to Casari and Luini (2012), there are three incentives for people to punish: instrumental punishment, reciprocity punishment and expressive punishment. Instrumental punishment means people use punishment to achieve a certain goal without much personal feelings, which should be the only motivation for a payoff maximizing person to choose punishment. The other two punishments involve emotional motivations. Reciprocity punishment means people choose to punish for retaliation. Expressive punishment means that people use punishment as a way to show their feelings or disapproval. Recent studies provide evidence to support the latter two theories and these "irrational punishment" could be the

cause of excessive punishment. Literatures related to reciprocity punishment can be seen from experiments that add retaliating stage/counter-punishment stage (Nikiforakis 2008; Casari and Luini 2012; Hopfensitz and Reuben 2009). Evidence about expressive punishment's existence can be found in Masclet et al.(2003) which shows that even without monetary reduction to free-riders, subjects still engage in punishing. Given the fact that non-monetary punishment can neither impose monetary penalty nor change group members' net earnings distribution, the authors suggest that punishment is also a method of expressing disapproval to free-riders. Müller and Duersch (2013) find that subjects will even pay for the right to execute the punishment to free-riders personally. Casari and Luini (2012) find that punishment of a non-cooperator by another subject is not a substitute good for one's own punishment. Studies in other experiment settings, such as the power to take game (Müller and Duersch, 2013; Galeotti, Fabio 2015) and the prisoners' dilemma game (Dreber et al., 2008) also shows that subjects use punishments to the point of sacrificing their economic gains.

Where does the excessive punishment come from? One important source seems to be subjects' emotions. Bosman and Van Winden (2002) find that high emotional intensity may drive subjects to act destructively in a power-to-take game. Joffily et al.(2014) find similar results in VCM games. They use skin response and self reported hedonic valence[2] to measure subjects' emotional change. They find that observing free-riding triggers emotional arousal and negative emotional feelings. The amount of punishment points is positively related with the punishers' psychological arousal level. Punishers who assign

[2] 10-point Likert Scale of self-reported feelings , 0 states as "extremely unpleasant" and 10 states as "extremely pleasant". 5 as a neutral feeling.

more punishment points report higher positive feelings in stage 2. So it appears that over-punishment can compensate the punishers' negative emotional feelings towards free-riders and that some punishers enjoy punishing. Müller and Duersch (2013) find subjects report higher happiness when they can punish free-riders by themselves. Casari and Luini (2012) show that subjects lack strategic thinking in punishment usage and tend to use punishment to express their disagreements.

On one hand, punishment clearly has benefits for the person doing the punishing. On the other hand, people who received punishment or expect to be punished may punish others as retaliation. Joffily et al.(2014) report that the punished experience high psychological arousal with negative emotional feelings. Many experiments provide evidences that the punished will retaliate if given the opportunity (Hopfensitz and Reuben, 2009; Nikiforakis and Normann, 2008; Engelmann and Nikiforakis, 2014).

Among all the negative emotions that might increase punishment unnecessarily, anger's role is especially emphasized. Hopfensitz and Reuben (2009) and Ben-Shakhar et al.(2007) show higher anger intensity is positively related to higher punishment and anger can cause a loss of resource. Drouvelis and Grosskopf (2014) use short video to induce subjects' happiness or anger before they play the public goods game and find that happy subjects tend to contribute more and punish less than angry subjects.

To our knowledge, only one paper tries to improve group welfare by controlling subjects' emotions. Dickinson and Masclet (2015) design three treatments to help subjects release their negative emotion before making punishment decisions. The first treatment adds a cooling down period. The second treatment adds a self-reporting of emotional states to

the cooling down period. The third treatment adds virtual punishment points to the cooling down period and the self-reporting. Results show that subjects have the highest net earnings in the cooling down only treatment. Excessive emotion ventings in the second and third treatments reduce contribution levels and decrease the punishment's efficiency.

Unlike Dickinson and Masclet (2015) whose treatments let subjects vent their naturally occurring emotions, in this paper, we ask subjects to change their emotions through a regulation exercise. We would expect that people can adjust their emotions to an appropriate level under clear instructions. Emotion regulation can increase punishment efficiency and improve group welfare better than emotion venting. The policy implication of our experiment might be: if punishment's efficiency can be improved when people consciously regulate their emotions, then emotion regulation should be implemented with the availability of punishment. We adopt cognitive reappraisal as our regulation strategy. In the emotion regulation treatments, we instruct subjects to stay calm and to reinterpret other subjects' selfish investment decision as "just trying to do as well as he or she could". Studies in Psychology have found that cognitive reappraisal is an effective way of regulating emotions (Xiao and Houser, 2005; Mauss et al., 2007).

Another question we try to answer is whether emotion regulation is more beneficial for certain sub-populations than others. We recruit subjects based on their aggressiveness level[3] and seat them in groups with members of same type. Intuitively, low aggressive

[3] The determinant of high or low aggressiveness is based on subjects' self reported Buss-Perry aggression questionnaire score and their relevant ranking in Tufts score survey pool. Details refer to Experiment design, pre-screen and recruitment. In the paper, we use Buss-Perry aggression questionnaire (Buss and Perry, 1992) to measure people's aggressiveness. The questionnaire's score has been found highly

people are less likely to use emotions in their decision making as frequent or severely as high aggressive people.

associated with aggressive behaviors in video games (Anderson et al., 2004) and economic decision makings (Zhang et al., 2013). The questionnaire's score has been found highly associated with aggressive behaviors in video games (Anderson et al., 2004) and economic decision makings (Zhang et al., 2013).

Experiment Design

As stated before, our experiment has a 3x2 design, where we have homogeneous high/homogeneous low/heterogeneous aggressive subjects play VCM games with or without emotion regulation. There are 2 parts in our experiment design. First, we pre-screen potential subjects and recruit them in homogeneous or heterogeneous groups. Second, we randomly assign emotion regulation to different groups and want to see whether the treatment has similar effects on different groups. In the "prescreen and recruitment" section, we present the classification of high and low aggressive people and the distribution of people's aggressiveness in our collected data. In the "lab experiment" section, we will explain our treatment and how the experiment is conducted.

Prescreen and Recruitment

In our experiment advertisement, all interested individuals are directed to fill in an online survey which contains a consent form and an information-collecting survey. In the survey, individuals need to answer 29 Buss-Perry Aggression Scale questions (Buss and Perry, 1992) and report their gender and age. We will calculate their aggressiveness score based on their response and only recruit people who have a valid Buss-Perry Aggression score.

The Buss-Perry questionnaire measures individuals' aggressiveness from 4 dimensions, namely physical aggression, verbal aggression, anger and hostility. Each subject's aggressiveness score is computed as the average score of these four dimensions, ranging

from 1 to 7.[4] In order to have a valid score, subjects cannot have 2 or more null answers in either section.

We use 2.58 as the cut-off value for high- or low-aggressive types, which is the mean value of 223 responses collected in 2014 before our experiment started on September 1st 2015.[5] Anyone with a score above 2.58 is defined as a high aggressive person.

All subjects are recruited based on their aggressiveness type and randomly assigned to a heterogeneous group or a homogeneous group of their type (high/ low). Subjects do not know their aggressiveness type or other group members' types.

In total, 168 subjects participated in the experiment. Among the 42 groups we collected, 17 are homogeneous low aggressive, 16 are homogeneous high aggressive, and 9 are heterogeneous groups. Table A1 in appendix shows the individual aggressiveness scores in different treatment arms. T-test of individual aggressiveness scores shows that there is a significant difference between high and low aggressive groups (p= .00) and no significant difference for groups with or without emotional (p=.14 for high; p=.86 for low; p=.45 for heterogeneous). The groups are balanced.[6]

4 1 means "extremely uncharacteristic of me"; 7 means "extremely characteristic of me". People can also choose "prefer not to say" or "don't know", which we will calculate as a null score and their score in this section will be computed as the mean of all answered questions.. We rule out individuals who skip or provide null answers for more than 2 questions in either dimension.

5 From September 1st 2015 to Jan 21st 2016(the last date we recruit eligible subjects), we collect 265 additional survey responses.

6 We also find that the groups are balanced in terms of gender ratio: Among all subjects in homogeneous groups, 55.7% are females. High aggressive groups have50 % females while low aggressive groups have 61.2%.

Lab Experiment

In each session, the instructions are randomly assigned to either a treatment with or without emotion regulation. Subjects play a VCM game repeatedly for 10 periods in the same group of four people. (N=4; T=10; Partner-matching). Although the group composition won't change, every group member will be assigned a new identifier number in each period. So it is impossible to track a particular individual across periods. At the end of the experiment, 1 period is randomly selected as the period-that-counts and these actual earnings will be their final earnings for the experiment in addition to a $5 show up fee.

Experiment Procedure

Both treatments with or without emotion regulation follow the same experimental procedures.

At the beginning of the experiment, all subjects are asked to rate their current moods[7], as shown in Figure A1. Then the experimenter reads the instructions for the first 5 periods. In addition to explaining how the VCM game works to the participants, the instructions include information on how to approach the task. This information on "how to approach the task" varies between the treatments with and without emotion regulation, as will be detailed in a moment.

7 The eleven moods include: anxiety, loneliness, self-confidence, sadness, feeling of being rejected, happiness, feeling of being judged, anger, feeling of being loved, embarrassment, feeling of being active and alerted. The rating is 7-likewise scaled, from "not at all" to "very much".

Before the first 5 periods formally begin, subjects need to complete a comprehension test to make sure that they understand how the VCM game is played and how to approach the task. The first 5 periods are the classical VCM game with punishment.

Upon completion of the 5^{th} period, instructions for the next 5 periods are given out. In periods 6-10, subjects play the same VCM game with punishment. The only difference is that at the end of each period, they are also allowed to report their positive and negative emotion on a 5*5 grid as shown in Figure A2. The self-reported emotions can be used to measure a group's emotional welfare, which has been understudied in the former welfare discussion; however, we only implemented this task in periods 6-10 in case the self-report vents subjects' emotion or causes extra emotion change,. After all 10 periods are played, subjects complete the same self-report mood rating(as Figure A1) they did at the beginning of the session.

The whole experiment lasts around 1 hour.

Emotion Regulation

We have 6 treatments. In the treatments without emotion regulation, subjects are instructed to approach the task by responding naturally with whatever thoughts and emotions occur to them spontaneously. In the treatments with emotion regulation, subjects are asked to adopt a neutral, analytical, and objective attitude in order to stay calm. Instructions of how to approach the task will appear in each period before and during subjects make their punishment decisions, as shown in Figure A3 and Figure A4.

VCM Game Setting

The VCM game follows the setting of Fehr and G ächter (1999). In each period, there are 2 stages. In the first stage, each subject is endowed with E (e.g. 5) tokens and decides how much they would like to contribute to the public goods g_i. Each token they keep pays a higher return α (e.g. \$1.5) to that individual only; while each token contributed to the public goods pays a lower return β (e.g.\$1) to the individual and all the group members. At the end of Stage 1, subjects see how much the other members of the group contributed to the public goods and everyone's Stage 1's earning. The earning is calculated as:

$$\pi_i^1 = \alpha * (E - g_i) + \beta * \sum_{i=1}^{n} g_i$$
$$0 \le g_i \le E$$
$$\beta < \alpha < n\beta$$

where π_i^1 denotes individual i's outcome at first stage, $E - g_i$ is the token number that individual i keeps as private goods, g_i is the token number that individual j contribute to the public goods. In our experiment, the first stage profit is calculated as:

$$\pi_i^1 = 1.5 * (5 - g_i) + 1 * \sum_{i=1}^{4} g_i$$

The first stage earnings will be between 0 and 22.5(max is keeping all 5 tokens and other 3 members contribute 15 to the public goods).

Having seen other group members' contribution, in Stage 2 subjects are given the opportunity to purchase as many "punishment points" as they can afford with their stage

1earnings. These "punishment points" will cost the punisher a given amount ρ (e.g. \$1) but will cause a much greater loss ρ' (e.g. \$3) to the subject they choose to punish. A subject's final earning for one period will be Stage 1's earning minus costs of purchasing punishment points minus cost of punishments levied on the player. The formula is:

$$\pi_i^2 = \pi_i^1 - \rho \sum_{j \neq i} P_{ij} - \rho' * \sum_{k \neq i} P_{ki}$$

where π_i^2 denotes individual i's Stage 1's earning, P_{ij} is the punishment point individual i purchased and assigned to individual j. In our experiment, the net earning is calculated as:

$$\pi_i^2 = \pi_i^1 - \sum_{j \neq i} P_{ij} - 3 * \sum_{k \neq i} P_{ki}$$

Under this experiment setting, the socially optimal equilibrium is when everybody invests all 5 tokens in the public goods. In contrast, assuming self-interest agents and backwards induction, the Nash equilibrium is when everybody keeps all five tokens to him/herself. In this Nash equilibrium, each subject is maximizing their payoff function by completely free-riding. In order to understand this, first, consider a one-shot game where the group only plays the VCM game once. At stage 2, as each punishment point will cause \$1 cost, nobody will assign punishment points to other group members. Knowing that nobody will assign punishment points at stage 2, each individual's profit maximizing decision should be $g_i = 0$ and keep all tokens. In our game, subjects know that the game will play 10 periods. In the last period, for the same reason as in a one-shot game, a self-interest subject will punish 0 at stage 2 and invest 0 in the public goods. Using logic of backward

induction, all subjects will play the game as a one-shot game for all 10 periods. So the Nash Equilibrium is that all subjects will invest 0 in the public goods for all 10 periods. The addition of emotion regulation does not change the payoff function, nor does the groups' aggressiveness level. Nash equilibrium remains $g_i = 0$ and $p_{ij} = 0$ for all subjects i and j.

Theory and Relevant Predictions

Our experiment has two main hypotheses: first, we assume people of different aggressiveness levels act differently in VCM games. We speculate that more anger-prone people are more likely to use over-punishment and thus have lower net earnings at stage 2. We did not expect that high and low aggressive people contribute differently. However, as will be shown in the result part, we find the two aggressiveness types' difference lies in the public contributions. Second, we assume emotion regulation can help subjects adjust their emotions to an appropriate level and use punishment more effectively. We expect that high aggressive people will benefit more from the emotion regulation. We don't expect similar effects among low aggressive people since low aggressive people won't be as affected by their emotions in their punishment decisions.

Regression Models

The main regression model is:

$$Y_{it} = \beta_0 + \beta_1 high_i + \beta_2 reg_g + \beta_3 high_i * reg_g + X'\beta + \varepsilon_{it} \quad (1)$$

where Y_{it} denotes our main dependent variables: punishment, public contributions, net earnings and emotion outcome. *high* denotes a dummy variable for high aggressive

individuals, *reg* denotes a dummy variable for treatments with emotion regulation. X' denotes all control variables we included in the regression.

Results

Three Main Dependent variables

Given our data's nature, we are going to present the experimental results in two parts. In part 1, we only use homogeneous group data, meaning only those groups with either all high aggressive types or all low aggressive types. We first use group-level analysis to detect patterns and then use individual level data to support the findings. In part 2, we add the heterogeneous groups and see how the result changes.

Table 1 Descriptive Data at Group Level

	Group Punishment Points	Group Public Contribution	Average Net Earnings per Group member	Count
High Aggressiveness With Regulation	1.20 (1.885)	16.26 (4.260)	16.46 (3.805)	80
High Aggressiveness Without Regulation	2.31 (3.275)	16.18 (3.553)	15.30 (4.364)	80
Low Aggressiveness With Regulation	0.79 (1.917)	13.27 (5.041)	15.00 (3.780)	90
Low Aggressiveness Without Regulation	2.34 (2.788)	14.09 (4.902)	13.97 (4.447)	80
Heterogeneous Group With Regulation	1.88 (2.396)	15.48 (3.781)	15.30 (4.128)	50
Heterogeneous Group Without Regulation	5.25 (8.261)	15.28 (4.391)	11.80 (8.262)	40

Mean value and Standard error in parenthesis. "Group punishment points" is the mean value of punishment points assigned by all four group members in one period, ranging from 0 to 80. The maximum is calculated as the possible highest Stage 1's earnings for a group. In our data, the maximum for each treatment is (from top to bottom) 8; 14; 16; 13; 9; 29. "Group public contribution" is the mean value of total tokens assigned in the public goods in one period, ranging from 0 to 20. In our data, the minimum for each treatment is (from top to bottom): 4; 9; 4; 3; 8; 6. "Average net earnings per group member" is the mean value of group total net earnings in one period divided by 4. The maximum is 20. In our data, the minimum for each treatment is (from top to bottom):6; -.25; -1; .75; 6.625; -12.125.

Part1. Homogenous Groups

In total, there are 33 homogeneous groups. We have 9 groups of low aggressive with emotion regulation (denoted as LowReg) and 8 groups in each of the other three treatment arms (denoted as HighReg, HighNoreg, LowNoreg).

Punishment

Without regulation, high and low aggressive groups punish similar amounts. Pooling all decision periods, the average punishment assigned per period in HighNoreg is 2.313, compared to 2.337 in LowNoreg(see table 1). A Mann-Whitney test also shows that punsishments do not differ from each other by aggressiveness types (p=.60).

Emotion regulation decreases punishment for both high and low aggressive groups. On average when we add regulation, the HighReg assign 1.2 punishment points which is greater than the 0.789 punishment points assigned by the LowReg groups (p=0.16). Low aggressive groups decrease their punishment more than high aggressive and Mann-Whitney test shows that the punishment drop of each group is only significant for low aggressive groups. (Mann-whitney test: high aggressive p=.83; low aggressive: p=.03)[8]. Also, more periods without any punishments occur when we add regulation. Periods where 5 or more punishment points were assigned happen more frequently in non-regulated groups and the frequency decreases a lot once emotion regulation is implemented (see Appendix Table A2).

[8] The result is robust for first 5 periods. Pooling first 5 periods, the punishment drop is still only significant for low aggressive groups. (Mann-whitney test: high aggressive p=.63; low aggressive: p=.08)[8].

Figure 1 presents how the group total punishment points evolve across time. In all 10 periods, the non-regulated groups have higher punishment levels than the regulated groups. More punishments appear in first 5 periods than in last 5 periods.

Figure 1 Mean of Group Punishment Points Purchased in Each Treatment across 10 Periods

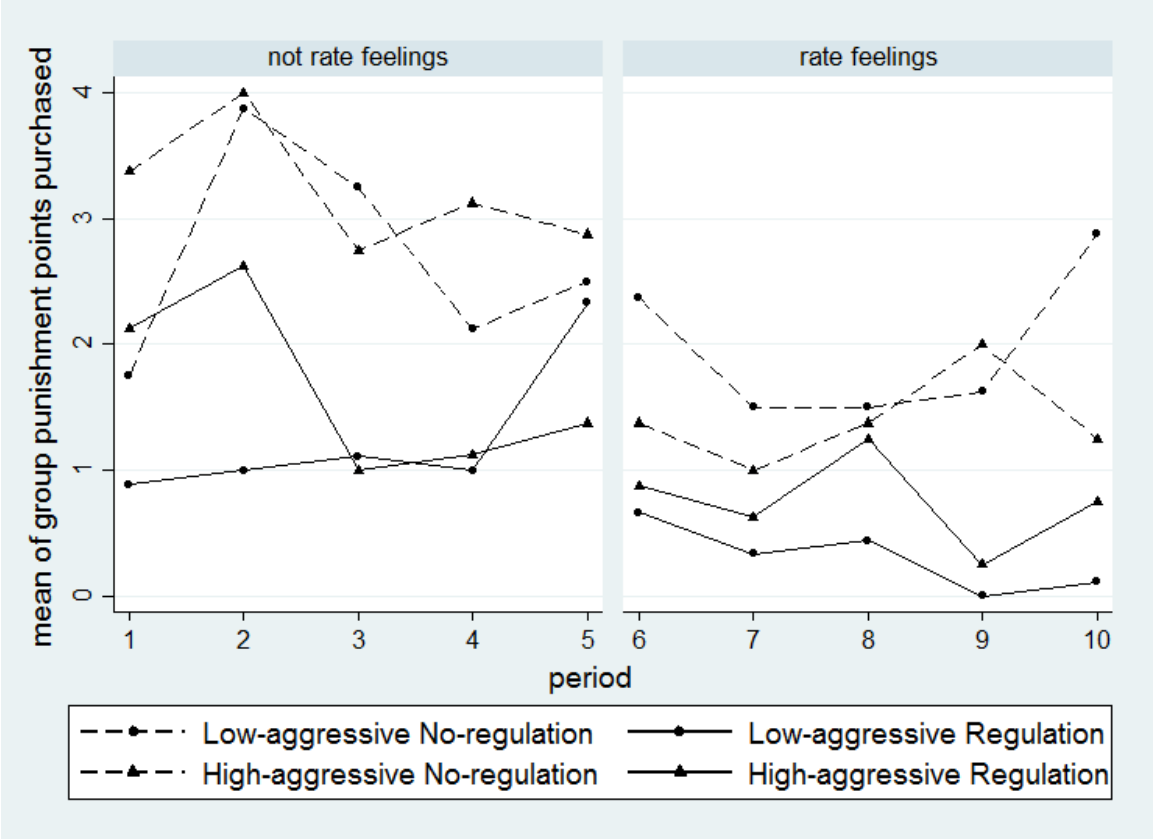


Table 2 Regression Results for Punishment in Homogeneous Groups

	(1) Group Punishment	(2) Individual Punishment
High Aggressive Group	0.168 (0.391)	0.484 (1.075)
Interaction Term of High *Regulation	0.520 (0.539)	1.066 (1.272)
With Emotion regulation	-1.624*** (0.376)	-2.284*** (0.658)
Period	-0.160*** (0.047)	-0.244*** (0.074)
Group Total Public Contribution	-0.092*** (0.030)	-0.149** (0.059)
Constant	4.517*** (0.550)	3.821*** (0.358)
Effective Observations	330 (33 Groups*10 Periods)	243 Minus left Censored, among 1320
R-squared	0.132	-
LowNoreg Dependent Variable's Average	2.337	-

Standard errors in parentheses, column(2) standard error clustered at group level
*** p<0.01, ** p<0.05, * p<0.1

Table 2 column (1) shows the OLS regression results for punishment. The dependent variable is group total punishment points assigned per period. The model follows equation (1). We control group public contributions and period. The inclusion of group public contributions is to account for the idea that punishments should be lower in a group with lots of public contributions. The inclusion of a discrete period variable (ranging from 1 to 10) is to account for people's learning over time, as the game is played in fixed groups.

From column (1), we see that after controlling each group's total public contributions, being in a regulated group significantly decreases group total punishment by 1.624 points. Being in a high aggressive group increases the amount of punishment.

Regression at individual level is consistent with the group level results, as shown in column (2). The dependent variable is each individual's purchased punishment points per period. To account for the large proportion (81.6%) of zeros in the dependent variable, we use Tobit[9] model to adjust the regression and set the censored cutoff value as 0. Standard error is clustered for each group. Individual level result is consistent with our finding. The emotion regulation has a significant negative coefficient(-2.284) on punishment. The coefficients of high aggressive dummy and the interaction term remain positive, and their addition cannot offset the negative impact brought by emotion regulation. Though the result is not statistically significant, it still suggests that high aggressive subjects may not decrease their punishment as much as low aggressive subjects under emotion regulation.

Result 1a: High and low aggressive groups do not punish differently in the public goods game without emotion regulation.

Result 1b: Emotion regulation decreases group's punishment for both high and low aggressive groups. It decreases low aggressive groups' punishment more significantly.

[9] Tobit model is used to account for the large amount of zero punishment in data set and avoid down-biased slop coefficient estimates or upwards-biased intercept estimate (Amemiya 1973). Using Tobit model, we define all non-zero punishment points as a latent dependent variable and include the probability of dependent variable above zero in regression's estimate. It has been proven to solve the inconsistency of estimators (Amemiya 1984).

Support for result 1b is found in the distribution of total punishment points assigned by each individual in four treatment arms (Appendix Figure A5). Emotion regulation decreases "extreme" punishers and increases non-punishers who never punish. "Extreme" punishers who assign over 30 punishment points only appear in non-regulated groups. In terms of non-punishers, 21(out of 36) people in LowReg never punish compared to 13 (out of 32) people in LowNoreg. The number is also increased for high aggressive people, from 11 (out of 32) in HighNoreg to 13 (out of 32) in HighReg. More low aggressive subjects become non-punishers under emotion regulation, which might partially explain why the emotion regulation decreases low aggressive groups' punishment more significantly.

Public Contribution

When it comes to the public contributions, we do not expect to see significant differences between high and low aggressive groups; however, the experiment shows that these two types of groups contribute differently. As shown in Table 1, the public contribution's mean of four treatments is 16.18 (HighNoreg), 16.26 (HighReg), 14.09 (LowNoreg), and 13.27 (LowReg). Although Mann-Whitney tests of each group pooling 10 periods does not show any significant difference among 4 treatment arms, there is evidence suggesting that high and low aggressive groups' contributions are different over time. Figure 2 shows how the group public contributions evolve over time. All four treatments start at the same level of public contribution in first period. High aggressive groups have a rising trend and end up in relatively high public contribution levels; while low aggressive groups tend to stay at their original level in their first 5 periods. In the last 5 periods, non-regulated groups' pubic contributions rise up but regulated groups have a clear declining trend. Considering that the first period should be the only period that won't be influenced by previous punishments, we use first period data to check the difference of public contributions in four treatment arms and do not find any significant differences. Pooling the first 2 periods, the difference remains insignificant. Given the fact that high and low groups do not have different punishment behavior, the lack of significance in early periods suggests that high and low aggressive groups' difference may result from their difference in responses to punishment.

Figure 2 Mean of Group Public Contribution in Each Treatment across 10 Periods

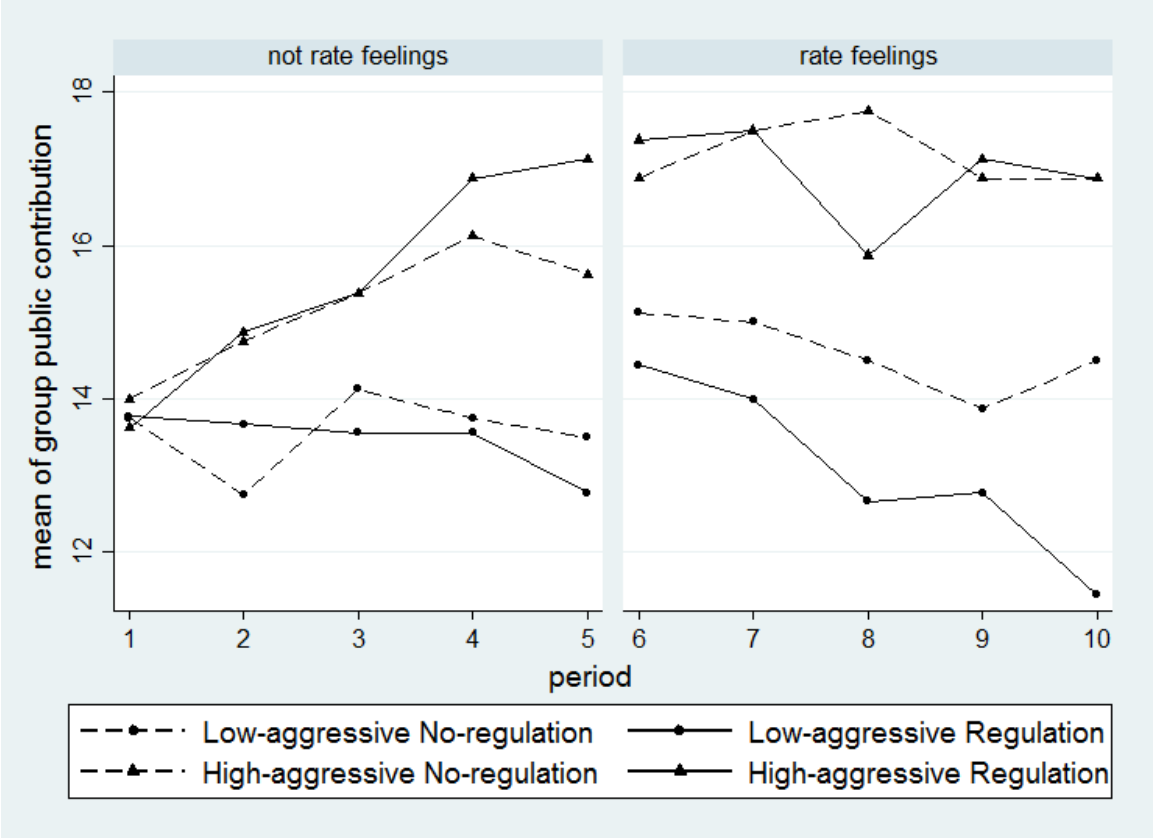


Table 3 Regression Results for Public Contribution in Homogeneous Groups

	(1) Group	(2) Individual
High Aggressive Group	2.088*** (0.709)	0.523*** (0.124)
Interaction Term of High *Regulation	0.908 (0.989)	0.228 (0.169)
With Emotion regulation	-0.821 (0.689)	-0.207 (0.131)
Period	0.143* (0.086)	0.044*** (0.014)
Constant	13.302*** (0.689)	3.278*** (0.126)
Effective Observations	330 (33 Groups*10 Periods)	1,320 (33 groups* 10 periods * 4 members)
R-squared	0.088	0.052
LowNoreg Dependent Variable's Average	14.09	3.52

Standard errors in parentheses, column (2) use GLS regression to account for the fixed group and individual characteristic. *** p<0.01, ** p<0.05, * p<0.1

Table 3 Column (1) presents regression result for public contributions at group level. Regression follows equation (1) and here we only control the period effect. Being in high aggressive groups increases the group public contributions by two dollars (2/14.9=13.3%). The result is statistically significant. Unlike Dickinson and Masclet (2015) who use emotion venting, our emotion regulation does not reduce contribution for both types. Emotion regulation has statistically insignificant coefficients on public contribution. For low aggressive people, the emotion regulation decreases their group total contribution by 0.8 dollar; but for high aggressive groups, it increases the

contribution by 0.1 dollar. Comparing with the baseline level of 14.9, it is a very small effect.

Column (2) presents GLS result at individual level. The dependent variable is each individual's public contribution per period. We use general least square regression to account for the fixed group and individual characteristics. The result is consistent with what we found at group level.

Result 2: High aggressive groups tend to contribute more than low aggressive groups.

Result 3: In homogeneous groups, emotion regulation does not influence contributions to the public good directly.

What causes the two types to contribute differently? At individual level adding their 10 periods' contributions, low aggressive subjects have a wider range of public contributions than high aggressive subjects (see Appendix Figure A6). Only in the low aggressive groups are there subjects who never contribute across the 10 periods. We speculate that low aggressive may be less sensitive to punishments they received or less willing to change their contribution levels. As subjects punish less in later periods, low aggressive groups decrease their contributions.

Group Welfare

Lastly, we use net earnings at stage 2 as a measurement of group welfare. Pooling all 10 periods, the average individual's net earnings in each period is higher for the emotion regulation treatments (see Table 1).

Figure 3 shows groups' net earnings' evolution across 10 periods. For most of the periods, the net earnings of regulated groups are higher than non-regulated groups. A Mann-Whitney test does not show any significant difference of each treatments pooling all 10 periods ($p>.1$), but regression result in Table 4 column (1) shows that emotion regulation has a significant positive effect on group welfare. On average, emotion regulation increases each group member's net earnings by one dollar, holding other factors constant. Being in high aggressive groups increases the net earnings significantly as well, which seems related to the high contributions made in high aggressive groups. Column (2) presents GLS regression result for individual's public contributions and net earnings. The regression results do not change much from group level's results.

Result 4: Emotion regulation increases group net earnings for both high and low aggressive groups.

Result 5: High aggressive groups have higher net earnings compared to low aggressive groups.

Figure 3 Mean of Individual Net Earnings in Each Treatment across 10 Periods

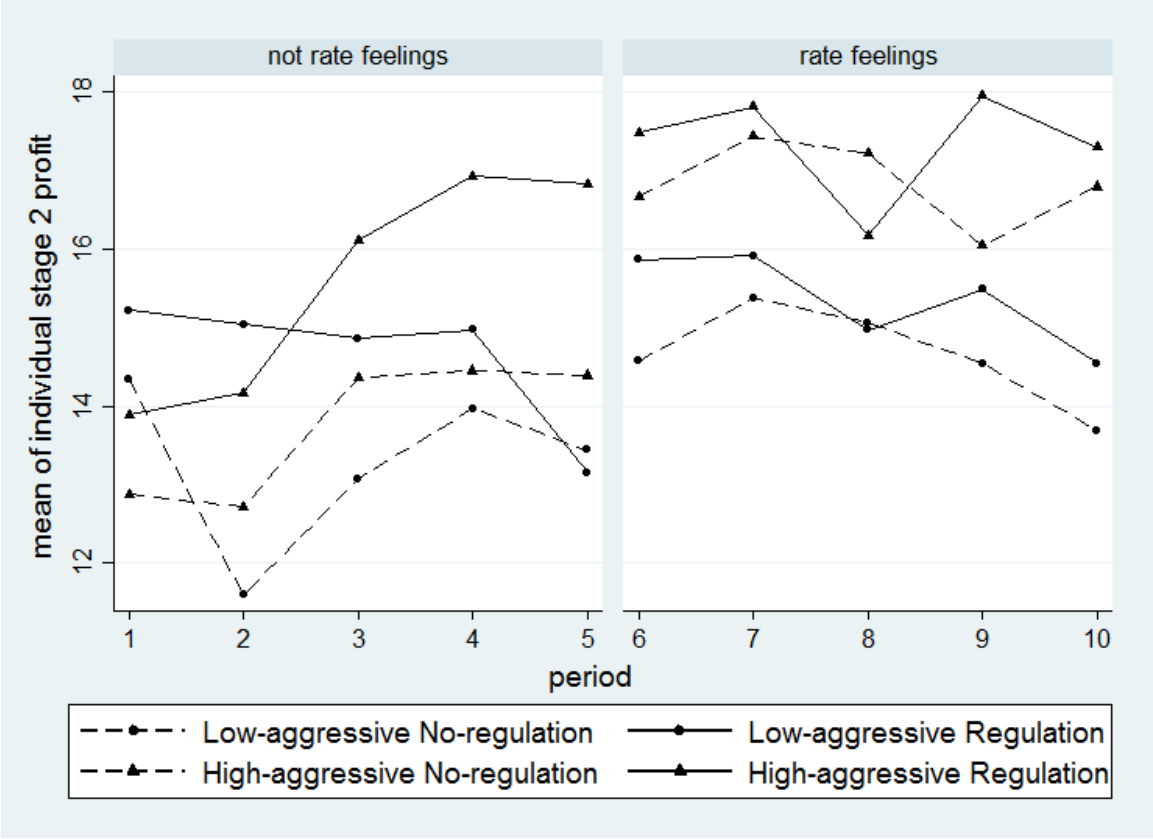


Table 4 Regression Results for Net Earnings in Homogeneous Groups

	(1) Average Net Earnings	(2) Individual Net Earnings
High Aggressive Group	1.330**	1.370***
	(0.638)	(0.384)
Interaction Term of High *Regulation	0.132	0.150
	(0.890)	(0.490)
With Emotion regulation	1.036*	1.005***
	(0.620)	(0.353)
Period	0.262***	0.251***
	(0.077)	(0.042)
Constant	12.526***	12.578***
	(0.620)	(0.363)
Effective Observations	330	1,320
	(33 Groups*10 Periods)	(33 groups* 10 periods * 4 members)
R-squared	0.077	0.066
LowNoreg Dependent Variable's Average	13.97	-

Standard errors in parentheses, column (2) use GLS regression to account for the fixed group and individual characteristic. *** p<0.01, ** p<0.05, * p<0.1

Part 2. Individual Level for all groups

In Part 3, we add all 9 heterogeneous groups in data analysis. Our main goal is to check whether people behave differently in heterogeneous groups.

To account for the heterogeneity of group members' aggressiveness levels, we add a dummy variable for all homogeneous groups. Table 5 presents the adjusted regression results for all three main dependent variables. The results for punishments are consistent with the group level results. Being in homogeneous groups seems to decrease the punishment level, but the result is not significant. For the public contributions, as suggested in Column (2), emotion regulation decreases low aggressive subjects' public contributions by .269 tokens, but increases high aggressive subjects' contributions by .379 tokens on average, holding other factors constant. There is a significant difference in public contributions between two types. Column (3) shows results for net earnings. Being in homogeneous groups increases the net earnings by 1.261 dollars. Though people in homogeneous and heterogeneous groups do not show significant differences in their punishments and contributions, the net earnings are improved when subjects play with similar aggressive type members.

Result 6: People in homogeneous and heterogeneous groups do not punish or contribute differently. Homogeneous groups have higher net earnings compared to heterogeneous groups.

Table 5 Adjusted Regressions for All Individual Data

	(1) Punishment	(2) Public Contribution	(3) Net Earnings
High Aggressive People	-0.031	0.321***	0.948***
	(1.186)	(0.109)	(0.322)
Interaction Term of High *Regulation	1.702	0.327**	-0.070
	(1.718)	(0.150)	(0.483)
With Emotion Regulation	-2.085**	-0.269**	0.778**
	(0.852)	(0.118)	(0.335)
Homogeneous Group	-1.227	0.044	1.261***
	(1.501)	(0.088)	(0.398)
Period	-0.218***	0.045***	0.250***
	(0.072)	(0.013)	(0.041)
Group Total Public Contribution	-0.177**		
	(0.072)		
Constant	5.162***	3.349***	11.633***
	(1.066)	(0.133)	(0.494)
Effective observations	323	1,680	1,680
	Minus left censored	(42 groups*10 periods * 4 members)	
R-squared	-	0.035	0.037

Standard errors in parentheses, column (1) uses Tobit model, standard error clustered for each group. column(2) and (3) use GLS regression. *** p<0.01, ** p<0.05, * p<0.1

Emotion Outcome

What happened to groups' emotions? Did we actually help people reduce their negative emotions? Since we collected people's emotions for the last 5 periods, Table 6 shows for all available data and homogeneous groups' data, how the positive and negative emotions change across four treatment arms. Column (1) and (3) show that given each groups' public contribution level, positive feelings' absolute value is only significantly related with group total public contributions. Column (2) and (4) show that people's negative feelings are decreased under emotion regulation and are negatively related with group public contributions. All the results fit our expectation. In general, subjects report higher positive feelings than negative ones.

Result 7: Emotion regulation decreases people's negative emotions and does not influence their positive emotions.

Table 6 Feeling Rating Regression

	(1) Positive	(2) Negative	(3) Positive	(4) Negative
High Aggressive Person	0.161	0.056	0.195	0.043
	(0.118)	(0.105)	(0.137)	(0.120)
With Emotion regulation	0.066	-0.494***	0.067	-0.503***
	(0.121)	(0.108)	(0.132)	(0.116)
Interaction Term of High*Regulation	0.077	0.175	0.010	0.261
	(0.168)	(0.149)	(0.188)	(0.165)
Group Total Public Contribution	0.125***	-0.102***	0.118***	-0.104***
	(0.009)	(0.008)	(0.010)	(0.009)
Constant	0.235	2.634***	0.354**	2.660***
	(0.158)	(0.141)	(0.174)	(0.153)
Observations	820	820	660	660
R-squared	0.229	0.186	0.217	0.194
Data	All	All	Homogeneous	Homogeneous
LowNoreg Average Score	2.108	1.108	2.081	1.138

Conclusion

First, our experiment finds that emotion regulation increases group welfare for both high aggressive and low aggressive subjects. Emotion regulation helps decrease negative emotions and increase net earnings. The fact that it increases earnings for both aggressiveness types suggests that emotion regulation should be provided alongside with punishments as a means to improve group welfare.

Second, emotion regulation decreases punishment levels and do not influence public contributions to the public goods significantly for homogeneous groups. Unlike emotion venting in Dickinson and Masclet (2015), emotion regulation decreases punishment without lowering punishment's efficiency in homogeneous groups. Analysis of all available data shows that the regulation decreases low aggressive people's public contribution but increases high aggressive people's contribution. A possible explanation is that low aggressive people are less responsive to punishments than high aggressive people, so they need higher levels of punishment to increase public contributions. The decreased punishment caused by emotion regulation further discourages low aggressive people's contributions. For high aggressive people, the emotion regulation reduces excessive punishment and improves the punishment efficiency; however, we don't see a significant difference in public contributions between HighReg and HighNoreg for homogeneous groups. We speculate that high aggressive people may be more likely to over punish in a heterogeneous group. In conclusion, high aggressive people benefit most from emotion regulation.

Third, homogeneous groups increase net earnings more than heterogeneous groups. We don't find evidence suggesting that people punish or contribute differently in heterogeneous groups. The reason for this significant drop of net earnings in heterogeneous groups calls for further exploration. A further experiment design can focus on collecting different combinations of heterogeneous groups and compare the results. We speculate that the largely decreased net earnings come from high aggressive people's over-punishment and low aggressive people's unwillingness to change the contributions.

Our experiment is a VCM game played in fixed groups. In real life, there are many examples of small groups trying to provide a public good: a football team tries to keep a clean locker room; a group of students work on a project together; or a department's faculties make sure all departmental needs are met. Allowing peers to punish each other is a common mechanism to increase the provision of these public goods. For example, a football player cannot protect the messy QB in the locker room; a student can give a bad rating to her lazy group mate; or a faculty member can speak negatively about a bad citizen in his tenure proceedings. Such punishments can be harmful to the group welfare if not used efficiently. It is not hard to imagine that too many arguments and fights may frustrate everybody and destroy the whole group's cooperation. Here, our experiment shows that pairing emotion regulation with punishments can improve both emotional and monetary welfare. Regarding policy implications, it seems to be a smart idea to implement a low-cost message like "adopt a neutral, analytical, and objective attitude in order to stay calm" before the opportunities of excessive punishments, like when being asked to report a colleague or a group mate.

Reference

Amemiya, T. (1973). Regression analysis when the dependent variable is truncated normal. *Econometrica: Journal of the Econometric Society*, pages 997–1016.

Amemiya, T. (1984). Tobit models: A survey. *Journal of econometrics*, 24(1-2):3–61.

Anderson, C. A., Carnagey, N. L., Flanagan, M., Benjamin, A. J., Eubanks, J., and Valentine, J. C. (2004). Violent video games: Specific effects of violent content on aggressive thoughts and behavior. *Advances in experimental social psychology*, 36:200–251.

Andreoni, J. and Gee, L. K. (2012). Gun for hire: delegated enforcement and peer punishment in public goods provision. *Journal of Public Economics*, 96(11):1036–1046.

Ben-Shakhar, G., Bornstein, G., Hopfensitz, A., and van Winden, F. (2007). Reciprocity and emotions in bargaining using physiological and self-report measures. *Journal of economic psychology*, 28(3):314–323.

Bosman, R. and Van Winden, F. (2002). Emotional hazard in a power-to-take experiment. *The Economic Journal*, 112(476):147–169.

Buss, A. H. and Perry, M. (1992). The aggression questionnaire. *Journal of personality and social psychology*, 63(3):452.

Casari, M. and Luini, L. (2012). Peer punishment in teams: expressive or instrumental choice? *Experimental Economics*, 15(2):241–259.

Decker, T., Stiehler, A., and Strobel, M. (2003). A comparison of punishment rules in repeated public good games an experimental study. *Journal of Conflict Resolution*, 47(6):751–772.

Dickinson, D. L. and Masclet, D. (2015). Emotion venting and punishment in public good experiments. *Journal of Public Economics*, 122:55–67.

Dreber, A., Rand, D. G., Fudenberg, D., and Nowak, M. A. (2008). Winners don't punish. *Nature*, 452(7185):348–351.

Drouvelis, M. and Grosskopf, B. (2014). The effects of anger and happiness on pro-social behaviour. Technical report, Mimeo.

Fehr, E. and G ächter, S. (1999). Cooperation and punishment in public goods experiments. *Institute for Empirical Research in Economics working paper*, (10).

Fehr, E. and G ächter, S. (2000). Fairness and retaliation: The economics of reciprocity. *The journal of economic perspectives*, pages 159–181.

Fudenberg, D. and Pathak, P. A. (2010). Unobserved punishment supports cooperation. *Journal of Public Economics*, 94(1):78–86.

Gächter, S., Renner, E., and Sefton, M. (2008). The long-run benefits of punishment. *Science*, 322(5907):1510–1510.

Hopfensitz, A. and Reuben, E. (2009). The importance of emotions for the effectiveness of social punishment*. *The Economic Journal*, 119(540):1534–1559.

Joffily, M., Masclet, D., Noussair, C. N., and Villeval, M. C. (2014). Emotions, sanctions, and cooperation. *Southern Economic Journal*, 80(4):1002–1027.

Masclet, D., Noussair, C., Tucker, S., and Villeval, M.-C. (2003). Monetary and nonmonetary punishment in the voluntary contributions mechanism. *American Economic Review*, pages 366–380.

Masclet, D. and Villeval, M.-C. (2008). Punishment, inequality, and welfare: a public good experiment. *Social Choice and Welfare*, 31(3):475–502.

Mauss, I. B., Cook, C. L., and Gross, J. J. (2007). Automatic emotion regulation during anger provocation. *Journal of Experimental Social Psychology*, 43(5):698–711.

Müller, J. and Duersch, P. (2013). Taking punishment into your own hands: An experiment.

Nikiforakis, N. and Normann, H.-T. (2008). A comparative statics analysis of punishment in public-good experiments. *Experimental Economics*, 11(4):358–369.

Ouss, A. and Peysakhovich, A. (2013). When punishment doesn't pay:'cold glow'and decisions to punish. *Available at SSRN 2247446*.

Spraggon, J. M., Sobarzo, L. A. V., and Stranlund, J. K. (2015). A note on stochastic public revelation and voluntary contributions to public goods. *Economics Letters*, 126:144–146.

Xiao, E. and Houser, D. (2005). Emotion expression in human punishment behavior. *Proceedings of the National Academy of Sciences of the United States of America*, 102(20):7398–7401.

Zhang, Q., Zhang, D., Wang, L., et al. (2013). Is aggressive trait responsible for violence? priming effects of aggressive words and violent movies. *Psychology*, 4(02):96.

Appendix

Table A1 Summary Statistics of Aggressiveness Survey Score in Each Treatment

	mean	min	max	count
High Aggressiveness With Regulation	3.292	2.607	4.724	32
High Aggressiveness Without Regulation	3.472	2.828	4.414	32
Low Aggressiveness With Regulation	2.028	1.069	2.552	36
Low Aggressiveness Without Regulation	2.043	1.414	2.536	32
Heterogeneous Group With Regulation	2.894	2.207	4.552	20
Heterogeneous Group Without Regulation	2.740	1.517	3.500	16

Table A2 Punishment Distribution at Homogeneous Group Level

Group Punishment Points Assigned in One Period	Treatment							
	HighReg		HighNoreg		LowReg		LowNoreg	
	No.	Col %	No.	Col %	No.	Col %	No.	Col %
0 Point	46	57.5%	41	51.2%	58	64.4%	29	36.3%
1 Point	11	13.8%	9	11.3%	16	17.8%	12	15.0%
2 Points	8	10.0%	1	1.3%	9	10.0%	7	8.8%
3 Points	6	7.5%	6	7.5%	3	3.3%	12	15.0%
4 Points	1	1.3%	6	7.5%	3	3.3%	6	7.5%
5 or More Points	8	10.0%	17	21.3%	1	1.1%	14	17.5%
Total	80	100%	80	100%	90	100%	80	100%

Figure A1 Screenshot for Mood Rating

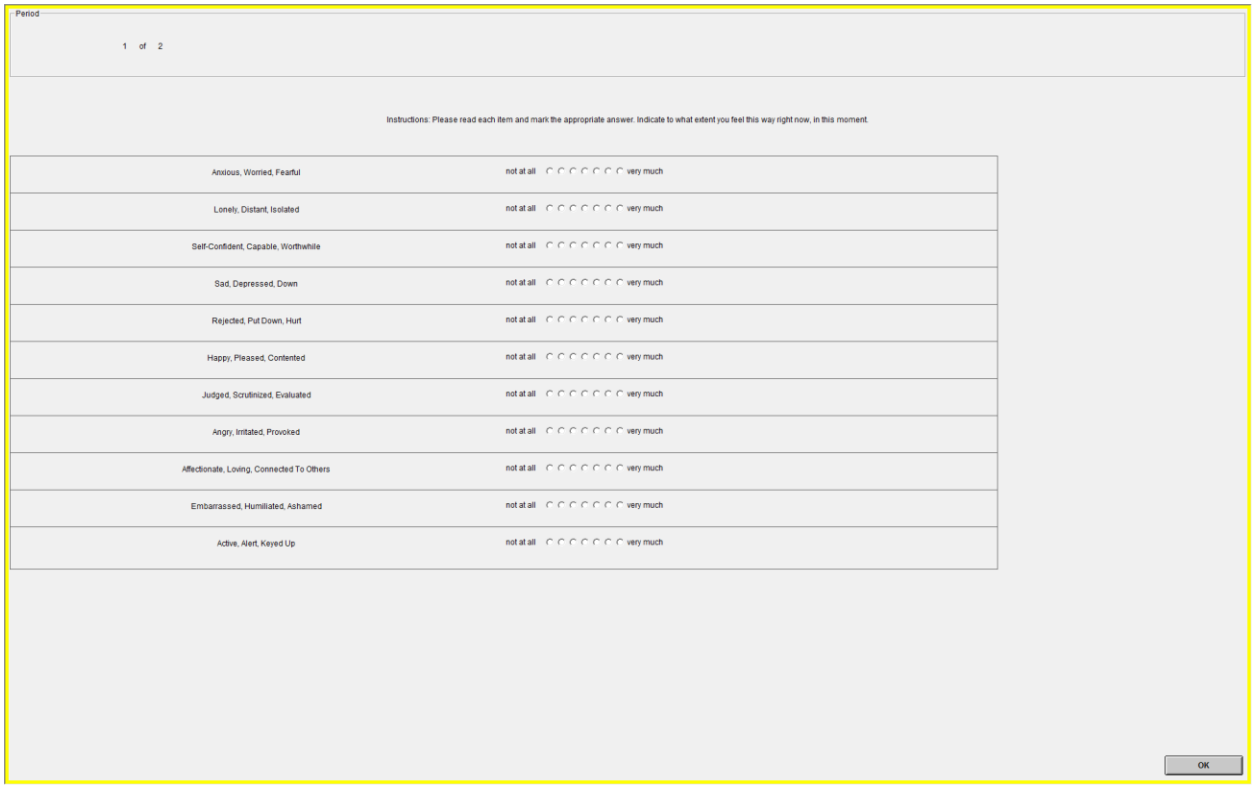


Figure A2 Screenshot for Self-Report Emotion Grid

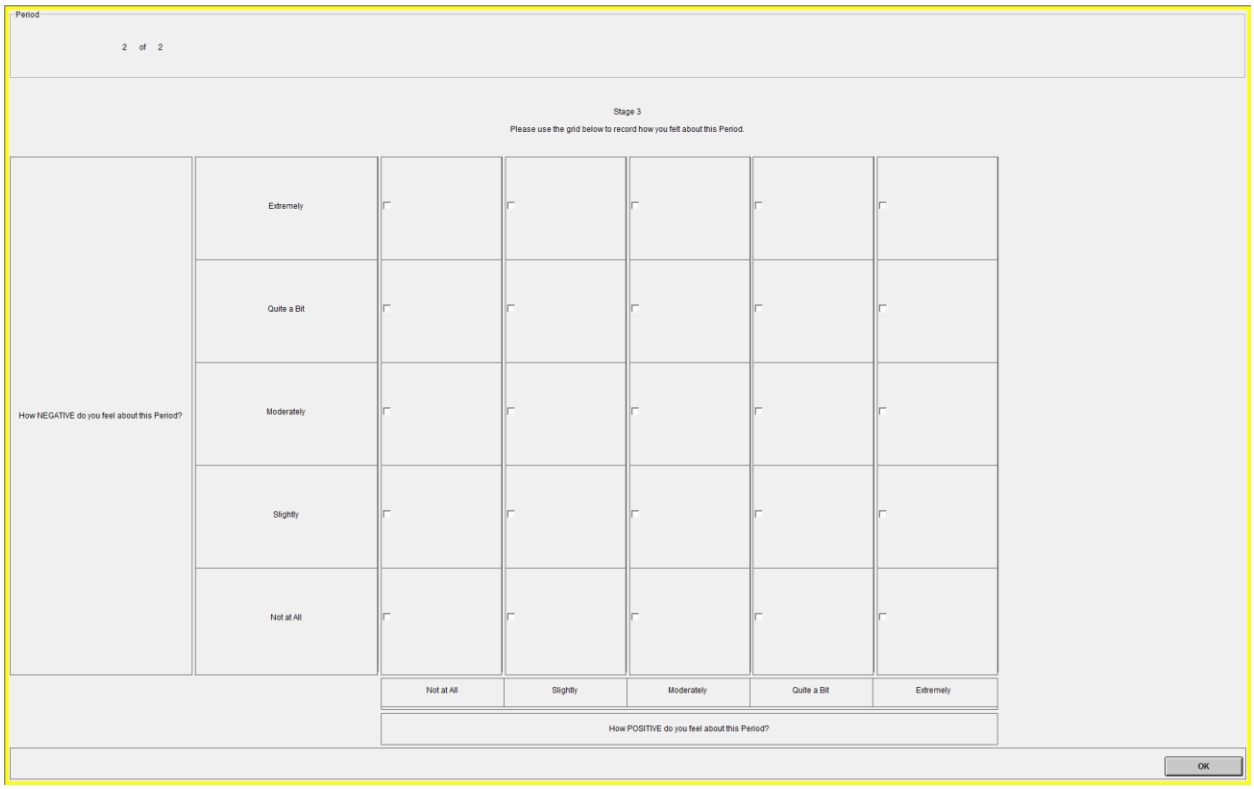
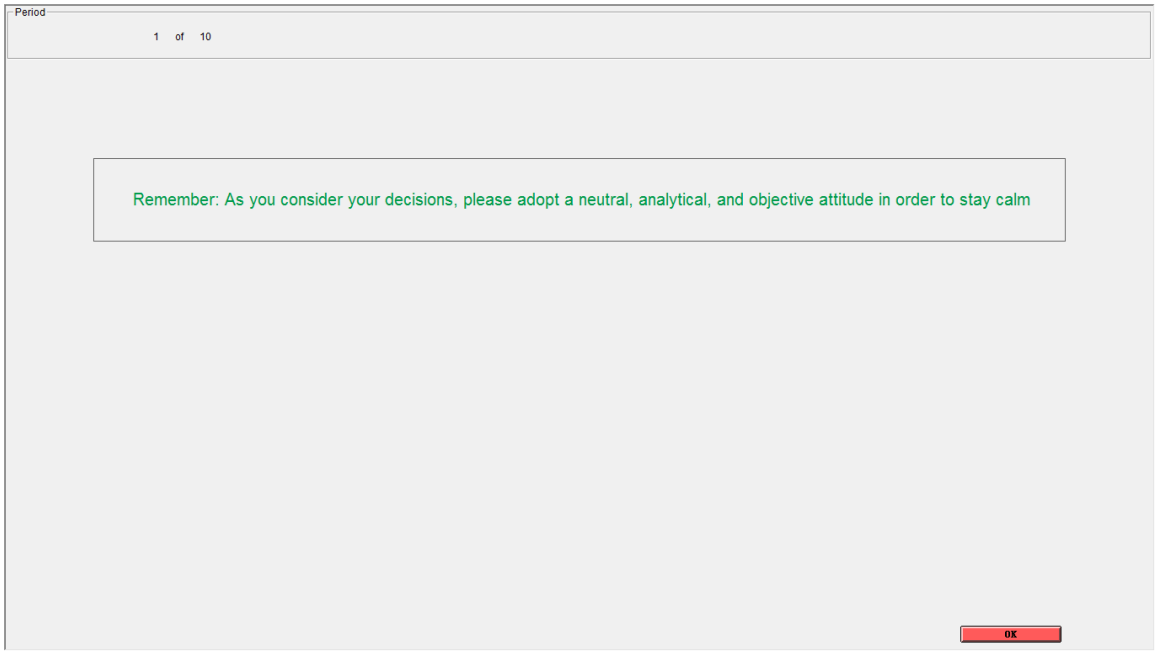


Figure A3-1 Screenshot for Treatments with Emotion Regulation



Note: this is the screenshot before subjects make their punishment decision.

Figure A3-2 Screenshot for Treatments with Emotion Regulation

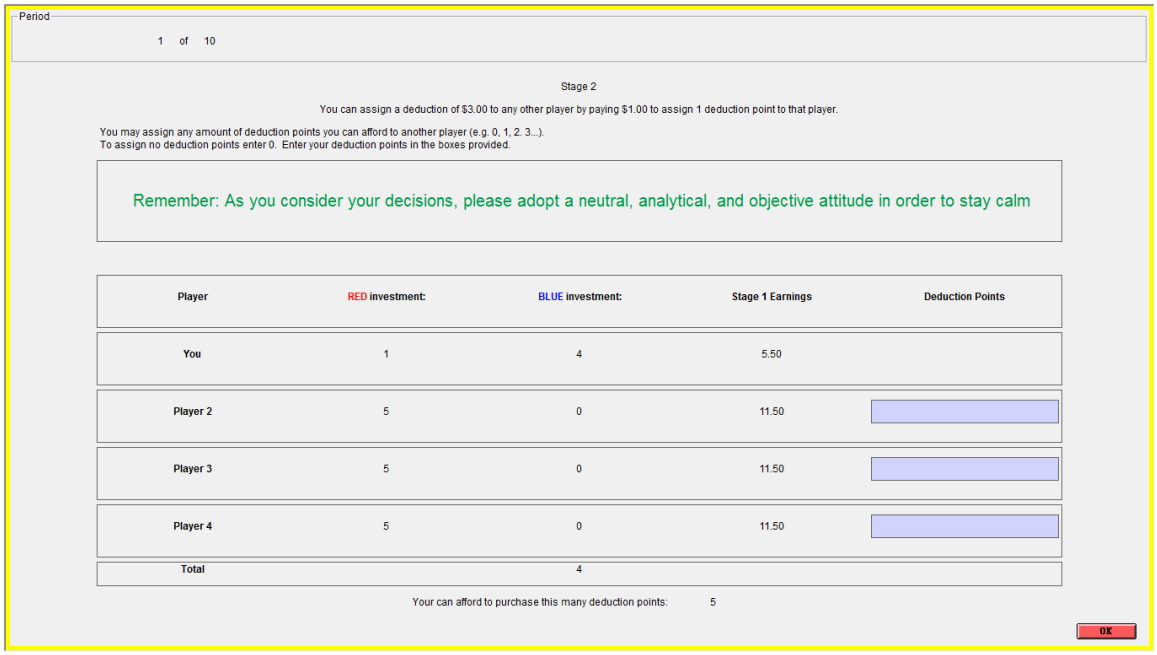
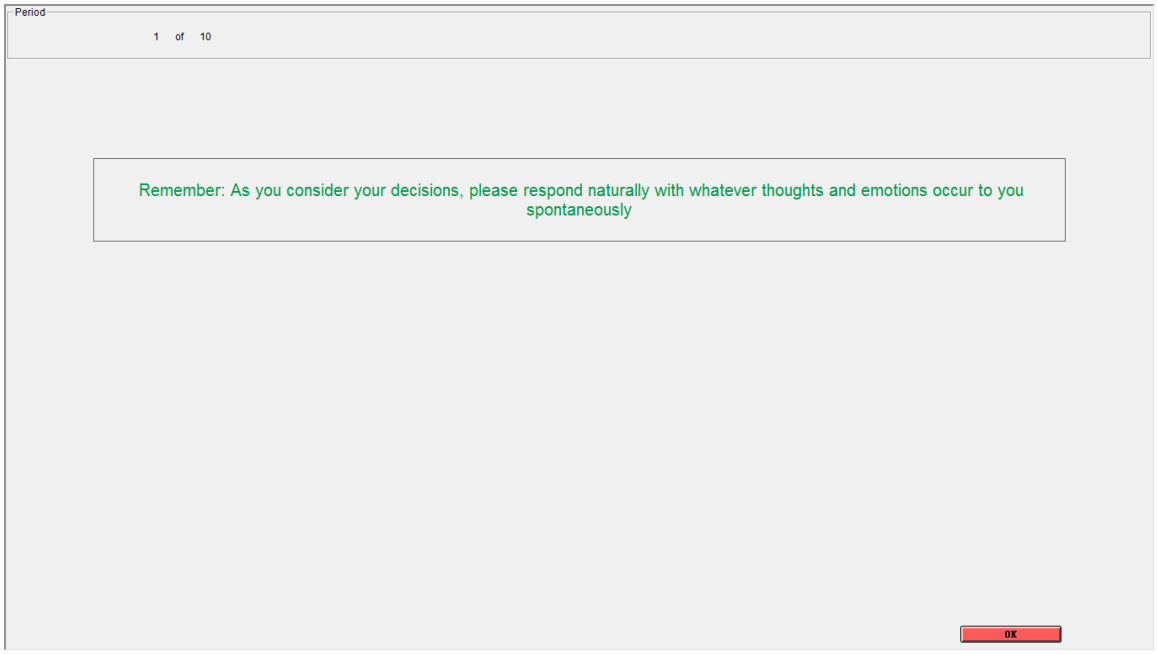


Figure A4-1 Screenshot for Treatments without Emotion Regulation



Note: this is the screenshot before subjects make their punishment decision.

Figure A4-2 Screenshot for Treatments without Emotion Regulation

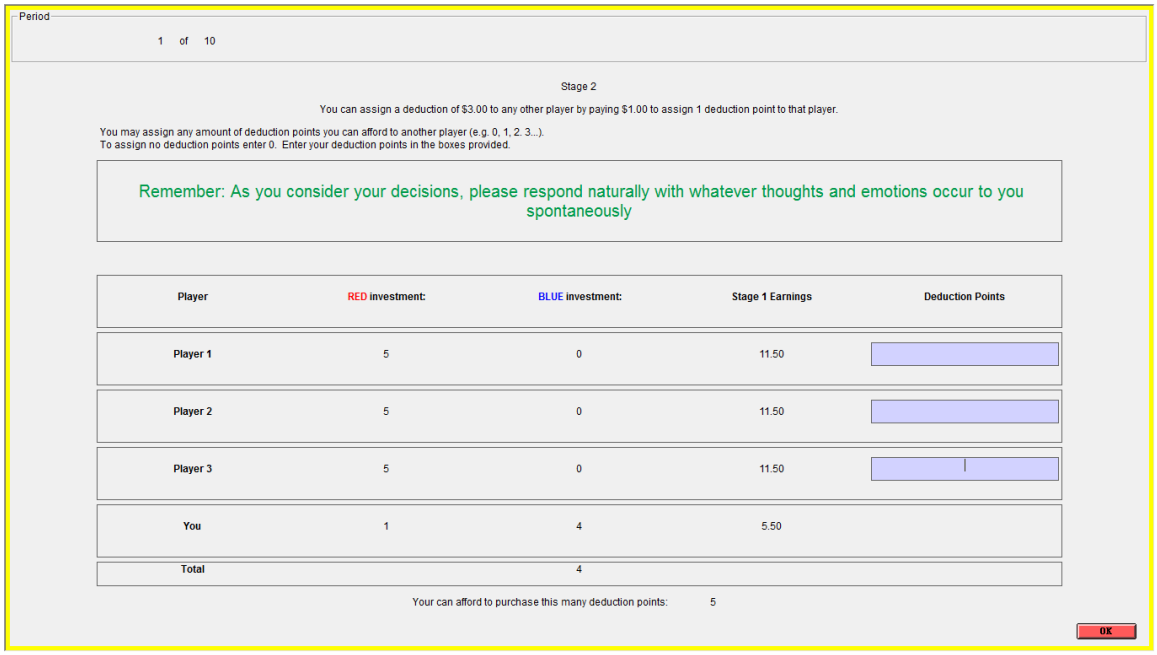


Figure A5 Individual Total Punishment Distribution (10 Periods Added-up) in Different Treatment Arms

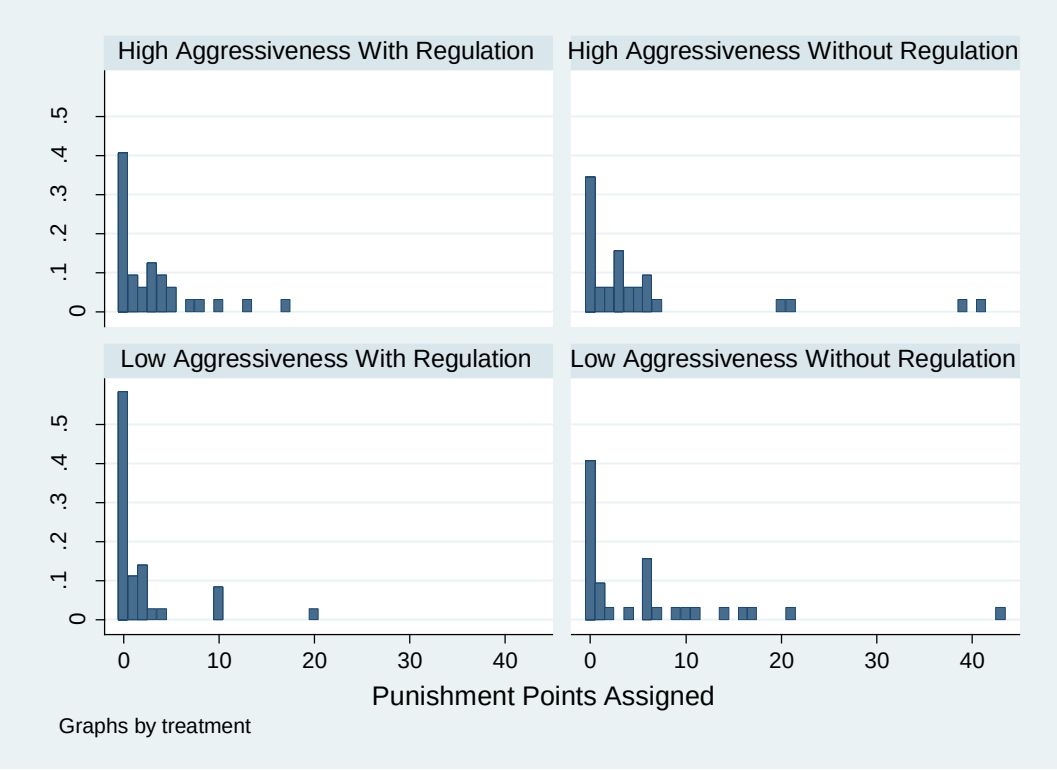


Figure A6 Individual Total Public Contribution (10 Periods Added-up) Distribution in Different Treatment Arms