



HAL
open science

Emotions, sanctions and cooperation

Mateus Joffily, David Masclet, Charles Noussair, Marie Claire Villeval

► **To cite this version:**

Mateus Joffily, David Masclet, Charles Noussair, Marie Claire Villeval. Emotions, sanctions and cooperation. *Southern Economic Journal*, 2014, 80 (4), pp. 1002-1027. 10.4284/0038-4038-2012.067 . halshs-00862782

HAL Id: halshs-00862782

<https://shs.hal.science/halshs-00862782v1>

Submitted on 19 Jul 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Emotions, Sanctions and Cooperation

Mateus Joffily¹, David Masclet², Charles N. Noussair³, Marie Claire Villeval^{4*}

March 2013

Abstract: We use skin conductance responses and self-reports of hedonic valence to study the emotional basis of cooperation and punishment in a social dilemma. We show that the availability of sanctions activates a "virtuous emotional circle" that accompanies cooperation. Emotional reaction to free-riding leads cooperators to apply sanctions. In response, and in addition to the monetary consequences of receiving sanctions, the negative emotions experienced by the free-riders when punished lead them to increase their subsequent level of cooperation. The outcome is an increased level of contribution that becomes the new norm. Therefore, emotions sustain both the use of altruistic punishment and cooperation.

Keywords: Emotions, Sanctions, Cooperation, Skin Conductance Responses, Experiment

¹ University of Lyon, Lyon, F-69007; CNRS, GATE, 93, Chemin des Mouilles, F-69130, Ecully, France. E-mail: joffily@gate.cnrs.fr.

² CNRS, CREM, 7 Place Hoche, 35065 Rennes, France, CIRANO, Montreal. Email: david.mascllet@univ-rennes1.fr.

³ Department of Economics, Tilburg University, PO Box 90153, 5000 LE Tilburg, The Netherlands. E-mail: C.N.Noussair@uvt.nl.

⁴ University of Lyon, Lyon, F-69007; CNRS, GATE, 93, Chemin des Mouilles, F-69130, Ecully, France. E-mail: villeval@gate.cnrs.fr.

* All authors contributed equally to this work. They thank S. Ferriol for programming assistance and K. Straznicka and J. Rosaz for valuable research assistance. We thank Ernst Fehr, Astrid Hopfensitz, Andrew Oswald, Frans van Winden, Lise Vesterlund, and participants at seminars at the University of Warwick, at the University of Pittsburgh, at Tilburg University, and at the ASFEE workshop on "How can Neurosciences inform Economics?" Grenoble. A grant from the Rhone-Alpes Region (CIBLE program) is gratefully acknowledged.

1. INTRODUCTION

Understanding how cooperation arises in social dilemmas is a longstanding focus of research in social sciences, and particularly in experimental economics. The most widely used arena for the investigation of cooperation in social dilemmas is the voluntary contributions mechanism (VCM, hereafter) that permits measurement of the extent of self- versus group-interested behavior (e.g., Marwell and Ames, 1979). In this setting, a group of individuals are each presented with an opportunity to allocate an endowment between two uses. The first use is a private account, which benefits only the individual. The second is a group account, which benefits all group members. The payoffs are specified so that it is a dominant strategy for each individual to place his entire endowment in his private account, but attaining the social optimum requires all individuals to allocate their full endowment to the group account.

Some studies have identified modifications to the VCM game that may potentially increase cooperation. In particular, a number of scholars have argued that the existence of punishment opportunities aids in sustaining cooperation in social dilemmas (Homans, 1961; Coleman, 1990; Elster, 1998; Bowles and Gintis, 2001). In a VCM with opportunity of punishment, a second stage is added to the game in which each subject, after observing each group members' contribution, has an opportunity to reduce the earnings of others, at a cost to himself (Fehr and Gächter, 2000, 2002). In the unique subgame perfect equilibrium to the game, individuals do not punish in the second stage and do not contribute in the first stage. Over the last 15 years, a huge literature has been devoted to explaining the main overall patterns observed in VCM experiments that contrast with theoretical predictions. The first pattern is that in absence of punishment opportunities, initial contributions are substantial but decline as the game is repeated and cooperation converges to a negligible level in the long run (Andreoni, 1988; Ledyard, 1995). The second pattern is that when punishment is available, participants exhibit a willingness to punish others, and the punishment opportunity increases contributions markedly (Yamagishi, 1986; Fehr and Gächter, 2000, 2002; Masclet *et al.*, 2003; Carpenter, 2007; Noussair and Tucker, 2005).¹

¹ If altruistic sanctions increase cooperation, a debate has developed with regard to their impact on efficiency. Some authors emphasize the waste of resources from costly sanctions (Houser *et al.*, 2008; Milinski and Rockenbach, 2008; Herrmann *et al.*, 2008). Others show that in the long run the availability of costly punishment increases surplus through its strong positive effect on cooperation (Gächter *et al.*, 2008). This is especially the

Several hypotheses have been proposed to explain why subjects over-contribute in comparison with the Nash equilibrium even in absence of punishment opportunities. These explanations rely on strategic motives (Kreps *et al.*, 1982) non-strategic motives that involve notions of fairness (Rabin, 1993; Falk and Fischbacher, 2006) or moral principles (e.g., Croson, 2007). Studies investigating the rationale behind altruistic sanctions also evoke both strategic and non-strategic motives. Negative reciprocity leads people to react to unfair intentions by sacrificing a part of their payoffs to punish others, even when there are no reputation gains from doing so (Rabin, 1993). The role of distributional concerns, leading to the use of punishment to reduce payoff inequality aversion has been also suggested (Fehr and Schmidt, 1999).

The present paper investigates another possible effect sustaining both cooperation and punishment behavior: the role played by emotions individuals experience when making decisions, and when learning the decisions of their group members. This interpretation in terms of emotions is complementary to the other motives mentioned above. For instance, one can plausibly argue that individuals may be willing to punish those whose decisions generated unfair payoff differences, because such interpersonal comparisons evoke strong negative emotions. Indeed, emotions might lead to punishment in response to selfish actions, even in cases when individuals cannot reduce inequality (e.g. Masclet and Villeval, 2008). Studying the role of emotions in decision-making may contribute to a better understanding of human behavior (Damasio, 1994), particularly in economics (Loewenstein, 1996; Elster, 1998; Bosman and van Winden, 2002; de Quervain *et al.*, 2004; Ben-Shakhar *et al.*, 2007; Glimcher *et al.*, 2009; Hopfensitz and Reuben, 2009; Coricelli *et al.*, 2010; Meshulam *et al.*, 2012).

Specifically, we study the role of emotions in the emergence and maintenance of cooperation. We measure emotions at different stages of a standard VCM game with and without sanction opportunities. Emotions may play a role in a number of ways in the game we study. First, emotions may support cooperation. This may be the case if contributing to the public good raises positive feelings such as warm-glow. In a sense, individuals would experience a kind of

case when punishment can operate, in conjunction with reputation, in a setting in which the same players interact repeatedly (Milinski and Rockenbach., 2006; Rand *et al.*, 2009).

trade-off between the monetary incentive to free ride and the pleasure of doing good.² Furthermore, some individuals may be willing to contribute to the public good to avoid disutility that arises from deviating from their "morally ideal contribution level" (Andreoni, 1990; Nyborg, 2000; Brekke *et al.*, 2003; Rege and Telle, 2004; Croson 2007).³ Second, emotions may also play a role when observing others' contributions. If the free riding of other group members generates negative emotions, the experience of such feelings may increase the cooperators' willingness to pay to punish. Third, the decision to punish may also be affected by the anticipation of emotions. In particular, the punisher may incur an emotional gain from punishing deviators, which may compensate for the monetary cost of punishment (de Quervain *et al.*, 2004). Finally, negative emotions experienced when being punished, coupled with the monetary impact of sanctions, may potentially affect the decision to change one's contribution in subsequent periods.

While emotional processes have already been evoked as a potential determinant in the decision to cooperate and/or punish in the context of a VCM (see Fehr and Gächter, 2000), direct evidence in the literature remains sparse. This is probably due to the difficulty of measuring emotions in such a context. Most previous studies have focused on two-person interactions. For example, functional Magnetic Resonance Imaging (fMRI) studies in neuroeconomics have revealed that the striatum, a brain area whose level of activation typically correlates with pleasant stimuli and the processing of rewards, exhibits increased activation from mutually cooperative behavior in prisoner's dilemma games (Rilling *et al.*, 2002; 2004). As for the decision to punish, it has been found that when observing opportunistic behavior, anterior insula activation, which is typically associated with aversive stimuli, correlates with an individual's subsequent decision to punish another subject (Sanfey *et al.*, 2003). This is congruent with studies that have identified that anger accompanies the application of costly punishment in two-person interactions (Bosman and van Winden, 2002; Ben Shakhar *et al.*, 2007). These studies measure emotions by means of self-reports and also, in the case of Ben Shakhar *et al.* (2007), by using electrophysiological measures of emotional arousal (see also Meshulam *et al.*, 2012). Moreover, a study using positron emission tomography has revealed that the costly punishment of social norm violators in a trust game

² Frijda (1986) illustrates well the role of emotions in this trade off. In the words of Frijda (1986), emotions appear to be a hypothesis to explain behavior that has neither sufficient nor adequate external purpose reason.

³ In this current study we do not attempt to disentangle these two motives.

also activates the dorsal striatum, in the anticipation of the satisfaction derived from punishing those who abuse one's trust (de Quervain *et al.*, 2004). However, the link between being punished and subsequent contribution decisions remains almost unexplored. To our knowledge, only Hopfensitz and Reuben (2009) have studied the importance of (self-reported) moral emotions (guilt and shame) in the decision to cooperate in a two-period game involving two players.

Our paper differs from these previous studies on several aspects. We investigate the relationship between emotions and behavior at several stages of interaction, including the contribution and punishment decisions in the context of a four-player public good experiment. Our design allows us to measure emotional states at all key moments of the interaction. We can thereby study the dynamic interplay between emotions and cooperation. Specifically, we study the emotional responses related to the level of cooperation an individual exhibits and to the observation of others' cooperation levels. When sanctions are possible, we consider the emotions that accompany the decision to punish, as well as the receipt of punishment. Lastly, we study the relationship between the emotional response to sanctions received and subsequent changes in cooperation. We measure emotional states along two dimensions: *i*) the level of physiological arousal, and *ii*) the hedonic valence of feelings, pleasant vs. unpleasant. The skin conductance response (SCR, hereafter) is used as a metric of the physiological arousal associated with emotional states (Boucsein, 1992; Bradley *et al.* 2001a; Critchley *et al.*, 2000). Unfortunately, SCR is ambiguous with regard to the valence of emotions, that is, whether the emotional state is positive or negative. Therefore, self-reports of experienced emotion, on a Likert scale, are used to measure the valence of the emotional experience.⁴

⁴ Multivariate analyses of verbal reports of emotional stimuli (see for example Russell, 1980) have shown that most of the variance in emotional reports can be explained by two main factors: valence (varying from negative to positive) and arousal (varying from low to high). According to this perspective, judgments of pleasure or displeasure indicate which motivational system is active, and judgments of arousal indicate the intensity of motivational activation (Bradley *et al.*, 2001a). Other theorists (Barrett and Russell, 1999; Reisenzein, 1994) have argued that verbal reports of arousal and the intensity of experienced emotion are not necessarily equivalent measures. Nevertheless, traditional perspectives widely agree that physiological responses play an important role in the generation of emotion (Gross and Barret, 2011). Self-reports of valence and the arousal dimensions have been correlated with autonomic and somatic responses to emotional stimuli (Bradley *et al.*, 2001a). While facial electromyography and heart rate are correlated with the emotional valence, SCR positively correlates with emotional arousal, independently of valence.

Our main findings are that lower arousal and more positive valence are associated with higher contributions. Learning that one has made lower contributions than others triggers higher arousal especially in the sanction treatment. Regarding the punishment decision, we find that for a given deviation in contribution levels, greater physiological arousal and experiencing more negative emotions when learning others' free riding increase the likelihood of punishing others. We also observe some emotional satisfaction from sanctioning free-riders, the anticipation of which may influence the sanctioning decision. We find that greater punishment received leads to higher arousal and more negative emotions, which in turn leads to a greater increase in contributions in the subsequent period. These patterns characterize a "virtuous emotional circle" showing that emotions play an important role in the emergence of cooperation.

The remainder of the paper is organized as follows. Section 2 presents our experimental design. Section 3 develops the theoretical predictions of the model. The results are detailed in section 4. Finally, section 5 discusses the results and concludes.

2. THE EXPERIMENT

2.1. Experimental design

Our experiment consists of two treatments of a public good game, the Baseline treatment and the Sanction treatment. The design is largely based on Fehr and Gächter (2000). Only one treatment was in effect in each session, and no subject participated in more than one session. The twelve participants that attended each session were assigned to groups of four with fixed membership. The four members of each group interacted repeatedly but anonymously for 20 periods under a "partner" matching protocol.

In each period of the Baseline treatment (see instructions in Appendix A), each participant received an initial endowment of 20 ECU (*Experimental Currency Units*, 100 ECU = 2 Euro). Players then simultaneously chose an amount between 0 and 20 ECU to contribute to their group account, and the remainder would be placed in their private account. Within each group, contributions were totaled, multiplied by 1.6, and redistributed equally among the members of the group. This means that each group member received 40% of the total amount the group assigned to the group account, in addition to the amount he assigned to his own private account. The payoff function in this treatment is:

$$\pi_i^1 = \left(20 - c_i + 0.4 * \sum_{j=1}^n c_j \right) \quad (1)$$

with c_i , the individual contribution to the group account. At the end of each period, participants were informed on the total amount contributed to the public account and on the individual contributions of each group member that were listed in random order. The identification number of each member was rematched at each new period.

In the Sanction treatment, a second stage was added to each period. In this second stage participants could punish the other members of their group. After being informed of the contribution of each of the three group members, a participant could assign between 0 and 10 punishment points to each member. Sanctioning was costly to both the punisher and his target. Each point assigned to an individual reduced the recipient's earnings by 10 percent, with a maximum possible reduction of 100 percent. The cost of punishment points for the participant who assigned them was discrete, but convex in the number of points. The cost function for the assignment of punishment is given in Table 1.

Table 1. The cost of sanctions to the punisher and to the target

Number of punishment points	1	2	3	4	5	6	7	8	9	10
Cost to the punisher	1	2	4	6	9	12	16	20	25	30
Percentage reduction of target's payoff	10	20	30	40	50	60	70	80	90	100

The payoff function at the end of the second stage in this treatment can be written as follows:

$$\pi_i^2 = \left(20 - c_i + 0.4 * \sum_{j=1}^n c_j \right) * \frac{\max \left\{ 0, 10 - \sum_{j \neq i} P_{ji} \right\}}{10} - \sum_{j \neq i} C(P_{ij}) \quad (2)$$

where P_{ij} is the number of points assigned by i to j , and $C(P_{ij})$ is the cost to i of assigning the points to j . Sanctioning decisions were made simultaneously and anonymously, so that it was not possible for a participant to identify who had punished him. Indeed, the rematching of identification numbers at each new period aimed at avoiding retaliation between two subjects.

2.2. Skin Conductance Response acquisition and valence elicitation

In this study, we measure emotions using both skin conductance responses and hedonic valence elicitation. We interpret skin conductance responses as a measure of physiological arousal associated with an emotional state. SCR was continuously and simultaneously

recorded for each of the 12 participants present until the end of the session. The sessions took place in a noiseless laboratory with a stable temperature set to 21° C.⁵ Skin conductance was analyzed in response to four events of interest. (a) The first is during the time interval in which the contribution decision is made. Each participant was allowed to make her contribution decisions at her own pace. After the last participant had made his decision, an 8 second interval was imposed before all participants were requested to report the valence of their feelings regarding their own contribution decision. (b) The second event of interest is the time of receipt of information about each other group member's contribution to the public account. This information was displayed on the screen for 8 seconds, and participants were simultaneously requested to report the valence of their feelings toward each other player's contribution. There were two more events of interest in the Sanction treatment, during which SCR was analyzed. The first of these was (c), the time interval during which participants decided whether and how much to punish other members of their group. The punishment decision was also self-paced. An interval of 8 seconds was imposed between the decision time of the last participant and the request to all participants to report their feelings regarding their own punishment decisions. Finally, (d) the punishment received from others was displayed on each participant's screen for 8 seconds, and it was followed by a request to participants to report the valence of their feelings toward the total sanction they received.

The subjects were required to report how they felt on a 10-point Likert scale of hedonic valence, which ranged from “extremely unpleasant” to “extremely pleasant”. A level of 5 is interpreted as a neutral sentiment, since 0 is the most unpleasant, and 10 is the most pleasant possible evaluation. As indicated above, this reporting occurred just after each event (a)–(d). When evaluating the contribution of others at time (b), subjects were asked to evaluate separately each of the three other players' contributions. Similarly, at time (c), a separate evaluation was requested for each assignment of points to the three other group members.

SCR were analyzed within specific time windows corresponding to events (a)–(d), as indicated in Figure 1. The analysis windows started 1 second after the onset of each event, to account

⁵ SCR was recorded with a BIOPAC MP150W system and two TEL100C telemetry modules (BIOPAC Systems, EU). Two Ag/AgCl electrodes filled with 0.5% saline in a neutral base paste were placed on the subject's distal phalanges of the middle and the index fingers of the non-dominant hand, after the attachment site had been cleaned with a neutral soap (Dawson *et al.*, 2000). A constant voltage of 0.5V was applied between the electrodes. For its analysis, the skin conductance signal was amplified (x2000) and low-pass filtered (30Hz) before being sampled at 125Hz.

for the SCR latency. The analysis window's duration varied by event. Events (a) and (c) had variable analysis window durations since decisions were self-paced. These are labeled as T1 and T2, respectively, in Figure 1. Events (b) and (d) had fixed durations of 7 and 2 seconds, respectively, to comply with the amount of information displayed on the screen at each moment. The data analysis is based on the SCR magnitude.⁶

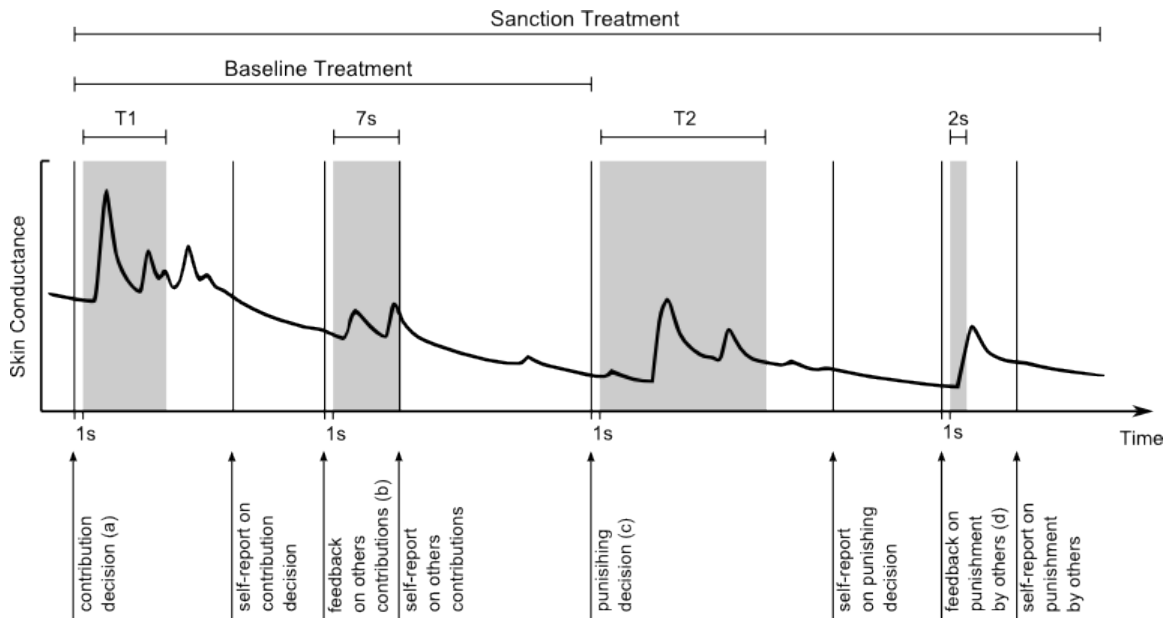


Fig. 1. Timing of events in the Baseline and Sanction treatments within a period

2.3. Procedures

The experiment consisted of four sessions. These sessions were conducted at the *Groupe d'Analyse et de Theorie Economique* (GATE), Lyon, France. 48 subjects (of whom 41.67% were males) were recruited by means of the ORSEE software (Greiner, 2004) from undergraduate courses in the local business and engineering schools. The experiment was computerized using the REGATE program (Zeiliger, 2000). Because there were four

⁶ The raw data were preprocessed as follows. The skin conductance signal was low-pass filtered at 0.5Hz offline, using a 5th order Butterworth low-pass digital filter. The onset and peak of the SCR were automatically detected when the first derivative of the filtered signal changed sign with a routine written in Matlab (The MathWorks Inc., USA). Onsets were identified by a negative to positive zero crossing, while peaks were identified by a positive to negative zero crossing. The SCR amplitude was calculated as the difference between the signal amplitude at the peak and the onset times. It was thresholded at 0.02 μ S. The whole signal was visually inspected prior to further analysis and ectopic response was removed. The SCR magnitude analysis takes into account all periods and assigns zero amplitude to those without a measurable response (see Dawson *et al.*, 2000).

individuals in each group, and group membership remained the same for the entire experiment, there were twelve independent observations at the group level (six per treatment).⁷

Upon arrival, each subject drew a tag from an opaque bag indicating his designated computer. Next, after washing his hands with a neutral soap, each subject was allowed to enter the laboratory and sit in front of his computer. Then, an assistant put electrodes on the non-dominant hand. Participants were required to keep the electrodes on until the end of the session. After checking the quality of the signal recording, the instructions for the experiment were distributed and read aloud. These instructions remained at the disposal of the subjects throughout the session. Understanding of the rules of the game was checked with a questionnaire. Subjects' questions were answered in private. At the end of the session, the participants were required to complete a demographic questionnaire. They were informed at the beginning of their session that a person who was not aware of the content of the experiment would disburse their earnings. On average, a session lasted 75 minutes and participants earned €15.45 (S.D.=2.54), including a €5 show-up fee.

3. PREDICTIONS

3.1. Standard predictions for rational and selfish players

In this section, we derive the predictions from our treatments under the assumptions of common knowledge of rationality and selfishness. In the Baseline treatment, if the game is played one-shot, there is a dominant strategy to contribute zero. Since each ECU kept in his private account yields an individual 1 ECU, whereas his return from each ECU contributed to the project is 0.4 ECU, a player's dominant strategy is to allocate his entire endowment to his group account. If the game is finitely repeated, the only subgame perfect equilibrium of the game is also for all players to contribute zero in each period. However, the maximum feasible group payoff would be attained only if each player contributes his full endowment to the group account. This would yield each group member 32 ECU. The average level of contribution to the group account is taken as a measure of cooperation. Classical game-theoretic reasoning produces an unique subgame-perfect equilibrium to the two-stage game of

⁷ The number of twelve participants per session was imposed by the availability of twelve pairs of electrodes to measure SCR. A potential risk of a small number of subjects per session is that some people may know each other. By inviting a large number of participants from different schools, and with the rule 'first registered, first in', we attempted to minimize the likelihood that several participants knew each other.

the Sanction treatment. Since punishment is costly, no group member ever uses the opportunity to punish, for any history of play. Thus, facing a threat that is not credible, each player places her entire endowment in his own private account. If the game is finitely repeated, the same strategy profile played every period, regardless of previous history, is the unique subgame perfect equilibrium.

3.2. Behavioral hypotheses

We can relax some of the above assumptions and assume that individuals may have some non monetary incentives to contribute to the public good. Indeed, the previous literature has shown that initial contributions are substantial in repeated games and decline over time. A first motive generally mentioned to explain why subjects contribute in the initial rounds of the game relies on strategic play based on the idea that players take into account future interactions when choosing both their contribution and punishment decision (Kreps *et al.*, 1982).⁸ Reciprocal actions, choosing to cooperate depending on previously observed decisions of others or on beliefs about their decisions (e.g. Rabin, 1993; Falk and Fischbacher, 2006), are widely documented. Another non strategic motive is based on the idea that agents target a morally ideal contribution (Croson, 2007) and make unconditional commitments to contribute (Nyborg, 2000; Brekke *et al.*, 2003; Figuières *et al.*, 2012).

In this study we consider the emotional substrates of contribution decisions. Our conjecture is that contributing to the public good increases positive feelings such as warm-glow. Therefore, some individuals may experience a trade-off between the monetary incentive to free ride and the pleasure to do good. Similarly, some individuals may incur an emotional cost by contributing less than their morally ideal contribution level. The operation of positive emotions in this manner would serve to increase contribution levels. That this pattern would be observed is our first hypothesis (H1).

H1. *Positive emotions are correlated with greater contributions.*

⁸ This hypothesis relies on "strategic play" based on the idea that players take into account future interactions when choosing both their current contribution and punishment decision (Kreps *et al.*, 1982). The justification of the strategic hypothesis is based on the "crazy player" assumption (Kreps *et al.*, 1982), or equivalently on the lack of common knowledge of rationality. If (rational) players believe that there is a crazy irrational player in the group who contributes positively in period 1, it becomes rational for them to play a trigger strategy in early periods and to mix over the strategy space as the repeated game approaches the final period. Therefore, in early periods they have an incentive for establishing a cooperative reputation, by making a large contribution.

Our second conjecture concerns the role of emotions experienced when observing the contribution of others and when comparing it with one's contribution. Observing that others have made high contributions may induce positive emotions in the anticipation of higher wn payoffs. In contrast, observing others' free riding may induce negative emotions. At the same time, in the Sanction treatment free-riders may experience both anticipatory emotions when evaluating the risk associated with the reaction of others and anticipated emotions when anticipating how bad they may feel if punished.⁹ Therefore, emotional arousal should be higher in the Sanction treatment and other things equal, the valence of free-riders' feelings should be lower than in the Baseline when learning others' cooperation as the risk of receiving sanctions mitigates the pleasure of higher expected payoffs. This is stated in H2.

H2 *Emotional valence in response to learning other group members' contributions increases in the level of others' contributions.) In the Sanction treatment, the emotional arousal in response to learning others contributions is greater and the valence is less positive than in the Baseline.*

Two non-strategic motives are generally evoked in the literature to explain why subjects are willing to sacrifice payoffs to punish others: a reaction to unfair intentions and distributional concerns. These two motives presumably have emotional underpinnings. We conjecture that in the Sanction treatment, individuals who have been more emotionally aroused, and more negatively so, when learning of others' low contributions may be more willing to punish free-riders. We also conjecture that punishing norm violators generates emotions with positive (on the "sweet taste of revenge", see de Quervain *et al.*, 2004). This valence of emotions experienced when sanctioning free riders is possibly all the more positive that the valence at the time of feedback was negative. Consequently, the decision to sanction may result from the trade-off between the monetary cost of punishing a free rider and the substantial emotional benefit of reciprocating negatively to free riding. Our predictions are summarized in H3.

H3. *a) The more individuals were (negatively) emotionally aroused when learning the free riding of others, the more punishment points they assign. b) The more punishment points they assign to free riders, the larger the increase of valence of punishers' emotions compared to the valence of emotions at the time of feedback.*

⁹ As defined by Loewenstein *et al.* (2001): "Anticipatory emotions are immediate visceral reactions to risks and uncertainties. Anticipated emotions are typically not experienced in the immediate present but are expected to be experienced in the future (...) anticipated emotions are a component of the expected consequences of a decision." (p. 268). See Coricelli *et al.*, (2010) for a study of emotions in tax evasion.

Our last conjecture concerns the emotional impact of the receipt of sanctions. We anticipate that being punished induces emotional costs in addition to monetary costs. Previous studies have shown that the individuals who have been punished, even without any monetary consequence, tend to increase their contribution in the next period (Fehr and Gächter, 2000; Masclet *et al.*, 2003). We conjecture that individuals are willing to avoid future punishment, with both its monetary and emotional costs, by increasing their contributions in the subsequent period. This effect is described in H4.

H4. a) *The more severely an individual is punished, the more he is emotionally aroused, as reflected in her SCR.* b) *Being punished is associated with negative emotional valence.* c) *For a given number of punishment points received, the more negative the emotions that are experienced, the more an individual increases her contribution in the subsequent period.*

4. RESULTS

Table 2 displays summary statistics on contributions, sanctioning behavior and emotions, measured by the SCR magnitude and the valence of feelings. These statistics are discussed in the following sub-sections.

Table 2. Summary statistics

Treatment	Baseline	Sanction
<i>Contribution</i>	9.32 (7.77)	19.17 (2.96)
<i>Punishment points assigned / received</i>	-	0.37 (1.19)
<i>Emotions – SCR</i>		
when contributing	0.26 (0.35)	0.27 (0.39)
when learning others' contributions	0.21 (0.34)	0.25 (0.49)
when assigning sanctions	-	0.35 (0.51)
when assigning no sanction	-	0.26 (0.44)
when receiving sanctions	-	0.21 (0.39)
when receiving no sanction	-	0.14 (0.37)
<i>Emotions – Valence</i>		
when contributing	5.85 (2.48)	7.81 (2.13)
when receiving feedback on others	5.12 (3.30)	8.99 (1.93)
when assigning sanctions	-	6.09 (2.88)
when assigning no sanction	-	8.09 (2.44)
when receiving sanctions	-	3.89 (2.41)
when receiving no sanction	-	8.99 (1.54)

Note: The numbers displayed are mean values. Standard deviations (S.D., hereafter) are in parentheses.

4.1. Contribution behavior and associated emotions

We first consider whether measuring SCR and eliciting the valence of emotions may affect behavior. The same qualitative patterns are observed as in previous studies (Fehr and Gächter, 2000; Masclet *et al.*, 2003; Noussair and Tucker, 2005). The mean level of contribution is significantly higher in the Sanction treatment (19.17) than in the Baseline (9.32) (see Table 2). Figures 2a and 2b display the evolution of the average individual contribution over time in the Baseline and Sanction treatments respectively, by group. A value of 20 ECU corresponds to the social optimum with full cooperation. A value of 0 corresponds to the Nash equilibrium with zero contribution. Figure 2a shows a large dispersion of contributions across groups in the Baseline treatment, with positive levels of cooperation in early periods and a decline over time to levels close to zero by the end of the sessions. In contrast, Figure 2b indicates that in the Sanction treatment mean contributions converge to near the social optimum of full contribution of 20 units after a few periods (with an end-game effect in one group).¹⁰

a. Baseline treatment

b. Sanction treatment

Figures 2. Average contribution by group and by treatment

We examine now the role of emotions in the contribution decisions. Our main findings are summarized in Result 1.

Result 1. *a) Higher contributions are correlated with a more positive self-reported emotional state. b) Greater arousal is correlated with lower contributions.*

Support for Result 1. Our data indicate that the mean self-reported valence increases with the individual contribution level. In the Baseline treatment, the average self-reported valence is

¹⁰ In the Baseline, the mean contribution is 13.04 ECU (S.D.=3.09) in period 1, 8.58 (S.D.=3.95) in period 10, and 3.46 (S.D.=3.55) in period 20. In contrast, in the Sanction treatment, the mean contribution is 13.83 ECU (S.D.=0.43) in period 1, 19.04 (S.D.=0.22) in period 10, and 19.17 (S.D.=0.21) in period 20.

5.40 (S.D.=3.09) for contributions less than 6 ECU (which is very close from the neutral level of 5), 5.87 (S.D.=1.58) for contributions between 6 and 15 ECU and 6.56 (S.D.=2.18) for contributions above 15 ECU. In the Sanction treatment, the corresponding mean values are 7.33 (S.D.=2.16), 6.67 (S.D.=2.14), and 7.90 (S.D.=2.11). In the Sanction treatment, if one excludes as outliers the 6 observations with contributions below 6 ECU, then the more group members cooperate, the more positive the feelings they report. Individuals feel more positive on average in an environment where sanctions are available than in a setting with no possibility of sanctions (see Table 2).

To provide a more formal analysis, we estimated an ordered Probit model with robust standard errors and clustering at the group level, in which the dependent variable is the self-reported valence of emotional state at the time of contribution (N=960) (not reported here but available upon request). The independent variables include the subject's contribution, his contribution interacted with a dummy variable for the Sanction treatment, a time trend, and the subject's gender.¹¹ The coefficient associated with the contribution is borderline significant and positive (coeff.=0.025, $p=0.105$). When this variable is interacted with the Sanction treatment, its coefficient becomes highly significant (coeff.=0.034, $p=0.001$). An interpretation is that in the Sanction treatment, the satisfaction of contributing is less mitigated by the risk of others' free riding than in the Baseline, as subjects learn quickly that the norm is to fully contribute.

In contrast with valence, the SCR magnitude when contributing decreases with the contribution. In the Baseline, the average SCR is 0.298 (S.D.=0.425) for contributions less than 6 ECU, 0.260 (S.D.=0.312) for contributions between 6 and 15 ECU and 0.184 (S.D.=0.253) for contributions above 15. In the Sanction treatment, the corresponding values are 0.588 (0.621), 0.264 (0.368), and 0.266 (0.391). We estimated a random-effects Tobit model in which the dependent variable is the SCR magnitude the subject exhibits when making his contribution decision (N=960).¹² The independent variables are the same than in

¹¹ We control for gender because of the widely held belief that males and females differ in their emotional responses. There is a strong stereotype of females being more emotional than males. Scientific evidence is, however, mixed. Studies based on self-reports show that females are more emotionally responsive than males (Fujita *et al.*, 1991; Bradley *et al.*, 2001b), but this may also result from stereotyping. Controlling for stereotyping, the effect is not robust (Barrett *et al.*, 1998). Studies based on physiological responses to emotional stimuli provide also contradictory results (see McRae *et al.*, 2008). In most of our regressions gender is not significant.

¹² The Tobit specification is used because of the left-censored dependent variable. Subject-specific random effects account for the fact that the same subjects make repeated decisions. We also estimated a Tobit model with

the previous model. The contribution variable receives a significant and negative coefficient (coeff.=-0.007, $p=0.039$), with no additional effect in the Sanction treatment (coeff.<0.001, $p=0.972$). This indicates that SCR magnitude is higher for individuals who cooperate less, perhaps in the anticipation of making higher earnings if others cooperate. Individuals feel, however, less aroused over time (coeff.: -0.014, $p<0.001$), which is likely due to a habituation effect.¹³

Consistent with hypothesis 1 (H1), these findings regarding the valence of feelings suggest that participants in a social dilemma face a trade-off between the monetary incentive to free ride and the satisfaction to contribute. In contrast, the negative link between arousal and contribution could suggest the presence of anticipatory feelings in free riders associated with the expectation of higher earnings depending on others' contributions. The fact that the Sanction treatment has no additional effect on SCR indicates that when deciding to free ride subjects are not more aroused by the expectation of sanctions.

4.2. The emotional impact of observing other group members' contributions

In this section we investigate the emotions experienced at the moment of observing contributions of others and when discovering differences between one's and others' contributions. Our findings are summarized in Result 2, The findings support H2 in most respects.

Result 2. *a) Individuals are more aroused and experience less positive emotions if they learn that others free ride. b) Learning that one contributes more than others affects valence negatively. It triggers arousal in the Sanction treatment. c) Contributing less than others triggers both greater arousal and more positive feelings.*

Support for Result 2. Table 3 analyzes the determinants of the self-reported valence (model (1)) and the SCR magnitude (model (2)) at the time of observing other group member's contribution. Model (1) uses an ordered Probit specification with robust standard errors and clustering at the group level. The independent variables include the other group member's contribution, the value of any positive difference between one's own contribution and another

clustering of standard errors at the group level, but without random effects. The estimates are qualitatively similar to all the results reported here.

¹³ We checked whether SCR was correlated with valence by estimating separate regressions of SCR for subjects reporting negative feelings (a valence lower than 5) and for those reporting neutral or positive feelings. This shows that the negative link between contribution and SCR holds only for the subjects who hold neutral or positive feelings. No link is found for the subjects who report negative feelings.

group member's contribution (i.e., one contributes more than the other), and the absolute value of the negative difference between one's own contribution and another group member's contribution (i.e., one contributes less than the other). Each of these variables is interacted with a dummy variable for the Sanction treatment to measure whether differences from others' contributions induce different emotions when sanctions are possible. Lastly, the regression includes a time trend and the subject's gender. Model (2) reports the results of a random-effects Tobit regression, in which the SCR magnitude registered at the time of receipt of feedback about others' contributions is the dependent variable. The independent variables are the same as in model (1), except that we consider the differences between one's own contribution and the mean contribution of the other group members.

Table 3. Determinants of SCR and valence of emotions when being informed about other group members' contributions

	Self-reported valence (1)	SCR (2)
Contribution of the other group member c_j	0.036*** (0.006)	-
Mean contribution of other group members \bar{c}_{-i}	-	-0.013 ***(0.007)
Value of positive difference from other's contribution ($\max\{0, c_i - c_j\}$)	-0.098*** (0.027)	-
Value of positive difference from other's contribution ($\max\{0, c_i - c_j\}$)*Sanction treatment	-0.141*** (0.045)	-
Absolute value of negative difference from other's contribution ($\max\{0, c_j - c_i\}$)	0.081*** (0.014)	-
Absolute value of negative difference from other's contribution ($\max\{0, c_j - c_i\}$)*Sanction treatment	-0.052** (0.026)	-
Value of own positive difference from the average: ($\max\{0, c_i - \bar{c}_{-i}\}$)	-	0.004 (0.008)
Value of own positive difference from the average: ($\max\{0, c_i - \bar{c}_{-i}\}$)*Sanction treatment	-	0.035* (0.019)
Absolute value of own negative difference from the average: ($\max\{0, \bar{c}_{-i} - c_i\}$)	-	0.021*** (0.008)
Absolute value of own negative difference from the average: ($\max\{0, \bar{c}_{-i} - c_i\}$)*Sanction treatment	-	0.011 (0.015)
Time trend		
Male	-0.032 ***(0.012)	-0.014*** (0.004)
Constant	-0.194 (0.164)	0.118 (0.093)
	-	0.139 (0.126)
N	2880	960
Left-censored obs.	-	414
Log-likelihood / Log pseudo-likelihood	-4084.411	-747.749
Pseudo R ²	0.255	-

Notes: Model (1) is an ordered Probit model with robust standard errors (in parentheses) and clustering at the group level. Model (2) is a random-effects Tobit model. *, **, and *** indicate statistical significance at the 0.10, 0.05, and 0.01 level, respectively

Model (1) shows that the more another group member contributes, the more positive the valence in response, with a greater positive effect in the Sanction treatment. When observing that one contributes more than the other party (a positive difference between one's own contribution and the other's contribution), valence is negatively affected, compared to a situation in which both players contribute the same amount. This effect is much stronger in the Sanction treatment, suggesting that a deviation from the norm is judged worse in this environment, where almost all players contribute their full endowment. On the other hand, observing that the other contributes more than oneself increases the valence of emotions, probably due to the expectation of higher earnings. This effect is, however, reduced in the Sanction treatment. This may result from the risk of sanctions, but more likely arises from the fact that the valence is already very high when the group members contribute the same amount (see Figure 3 that relates the difference between one's own and the other's contributions and the valence of emotions). It should be noted that the coefficient associated with a positive difference is higher, and considerably so in the Sanction treatment, than the coefficient associated with a negative difference.

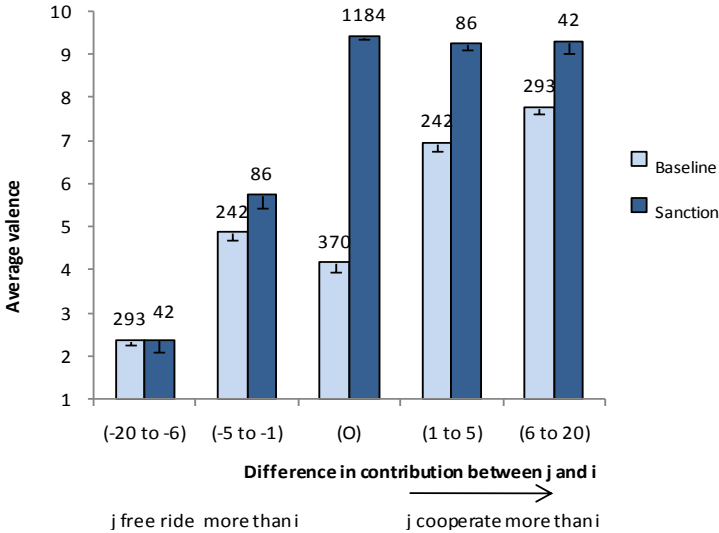


Figure 3. Average self-reported valence of emotions when being informed of another player's contribution, as a function of the difference between the other's contribution and one's own contribution ($c_j - c_i$)

The estimates of the model (2) show that when there is no possibility of sanctions, the subjects are more physiologically aroused when learning that the others contribute low amounts in the absolute and when they contribute less than the average themselves. Indeed, the coefficient of the absolute value of the negative difference from the average is highly significant. In the Sanction treatment, the physiological impact of the others' contribution is offset and there is no additional effect of learning that one contributes less than the average. This suggests that free-riders do not anticipate being punished or, alternatively, that the anticipation of sanctions generates no physiological reaction. In contrast, learning that one contributes more than others produces a significant physiological arousal in the Sanction treatment but not in the Baseline. A possible interpretation is that the cooperators feel anticipatory emotions associated with the perspective of sanctioning the free-riders.

In summary, the cooperators are more emotionally aroused and evaluate more negatively the contribution of free-riders in the Sanction treatment than in the Baseline. In contrast, free-riders are not more emotionally aroused when sanctions are possible. This suggests that the availability of sanctions affects the emotional reaction to the observation of free-riding of the cooperators much more than it affects the emotions of the free-riders.¹⁴

4.3. Emotions and punishment behavior

We now examine the role of emotions in the decision to punish and the emotions experienced when sanctioning others. On the first dimension, our findings are stated in Result 3 and provide strong support for hypothesis 3.

Result 3. *Punishment behavior is predicted by both greater physiological arousal and negatively-signed emotions when learning the contribution of others.*

Support for Result 3. As shown in figure A in Appendix B, there is a strong tendency for i to punish j more, the less that j contributed compared to i , as usually observed in VCM games. There is also a modest tendency for i to punish j more, the more j has contributed compared to i (anti-social punishment). A small difference (where a negative number denotes how much

¹⁴ We have also run separate regressions of the SCR for individuals who report a valence below 5 and for individuals who report a valence at least equal to 5 (available upon request). In the first sub-sample, the SCR increases in the difference between one's contribution and the mean contribution of others when the subject contributes more than the others but only in the Sanction treatment. In the second sub-sample, the SCR increases in the difference between one's contribution and the mean contribution of others when the subject contributes less than the others, but with no additional effect of the sanction treatment.

more the sanctioner contributes than the target), between (-2,+2), triggers on average 0.04 punishment point. The corresponding number of points is 1.04 for the range (-3, -8), 1.56 for the range (-9,-14), and 2.43 for (-15, -20). In the regressions reported in Table 4, we test whether a greater arousal and more negative emotional valence lead to more punishment for a given difference in contributions. Model (1) estimates the probability of sanctioning another group member by means of a random-effects Probit model. Marginal effects are reported into brackets. Model (2) estimates the number of points assigned to the group member j by means of a random-effects Tobit model. In both models, the independent variables include the valence of feelings and SCR when contributing and at the time of learning the contributions of others, the average contribution of others, and the difference between i 's and the average contributions of others in the group. In model (2), we also include two variables capturing the positive difference between j 's and the average contributions and the absolute value of the negative difference between j 's and the average contributions.

Table 4. Emotions and the decision to punish

	Probability of sanctioning another group member (1)	Severity of sanctions on a group member j (2)
Physiological arousal when contributing	-0.103 (0.243) [-0.017]	0.502** (0.241)
Physiological arousal when being informed of others' contributions	0.281* (0.159)[0.047]	0.113 (0.179)
Valence of emotions when contributing	0.076 (0.059) [0.013]	0.116* (0.063)
Valence of emotions when receiving feedback about others' contributions	-0.424*** (0.090) [-0.071]	-0.584*** (0.064)
Average contribution of others (\bar{c}_i)	-0.088* (0.050) [-0.015]	0.054 (0.054)
Difference between i 's and the average contribution of others ($c_i - \bar{c}_i$)	0.047 (0.030) [0.008]	-0.024 (0.031)
Value of positive difference between j 's and the average contribution ($\max\{0, c_j - \bar{c}\}$)	-	0.061 (0.070)
Absolute value of negative difference between j 's and the average contribution ($\max\{0, \bar{c} - c_j\}$)	-	0.088*** (0.033)
Time trend (t)	-0.069*** (0.019)[-0.012]	-0.100*** (0.024)
Male	-0.133 (0.269) [-0.022]	-0.184 (0.430)
Constant	4.311*** (0.806)	0.998 (1.001)

N	480	1440
Left-censored observations	-	1327
Log-likelihood	-140.563	-340.121

Notes: Model (1) is a random-effects Probit model. Marginal effects are in brackets. Model (2) is a random-effects Tobit model. *, **, and *** indicate statistical significance at the 0.10, 0.05, and 0.01 level, respectively.

Model (1) shows that the probability of punishing another group member is predicted by greater physiological arousal and negatively-signed emotions when learning others' contribution. The more negative the emotions experienced when the individual receives information about his group members' contributions, the greater his probability of assigning punishment. Once this and other variables are controlled for, a higher SCR at the time of observing others' contributions also increases the probability of allocating punishment. Model (2) indicates that physiological arousal and the experience of more positive feelings when making one's own contribution increase significantly the severity of punishment assigned. One punishes more severely a group member who contributes below the average and whose contribution has generated a more negative feeling.

We now investigate the emotional impact of sanctioning decisions. Our findings are stated in Result 4.

Result 4. *The more punishment points an individual assigns, the more he is physiologically aroused and the more positive is his reported valence.*

Support for Result 4. The average valence of emotions increases from 4.96 to 6.09 when subjects assign punishment points, compared with the moment when they learn others' contributions. This indicates that punishing others increases the satisfaction of the subjects who have been hurt by learning the contributions of others. In contrast, the subjects who do not sanction experience a slight deterioration of their emotional state. They report an average valence of 9.33 at the time of being informed of others' contributions and 8.09 at the time of the sanction decision. The average SCR magnitude is 0.26 μ S when no sanction is assigned, 0.31 μ S when one is assigning one or two points, and 0.44 μ S when one assigns more than two points in total to the other group members.

We estimated two models that are reported in Table 5. Model (1) is an ordered Probit model with robust standard errors and clustering at the group level in which the dependent variable is the valence of feelings reported when making a sanction decision for each other group member. The independent variables include the self-reported valence when the sanctioner was

informed of the targeted player's contribution, the number of punishment points assigned to this player and a dummy variable indicating anti-social punishment, which occurs when the target contributes more than the punisher. Model (2) is a random-effects Tobit model in which the dependent variable is the SCR magnitude when assigning punishment points. The independent variables include the SCR of the subject when informed of the contributions of other members, the total number of punishment points assigned and a dummy variable for anti-social punishment. A time trend and the sanctioner's gender are included in both models.

Table 5. Determinants of the physiological arousal and the valence of emotions when assigning sanctions

	Valence (1)	SCR (2)
SCR when receiving feedback about others' contributions	-	0.068 (0.055)
Valence of emotions when receiving feedback about others' contributions	0.335*** (0.056)	-
Punishment points assigned to group members	-	0.072*** (0.023)
Punishment points assigned to cooperators	-	-0.111 (0.115)
Punishment points assigned to j	0.339** (0.134)	-
Punishment points assigned to j if j is a cooperator	-0.806* (0.469)	-
Time trend (t)	0.006 (0.009)	-0.003 (0.005)
Male	-0.928** (0.395)	0.421*** (0.156)
Constant	-	-0.065 (0.107)
N	1440	480
Left-censored observations	-	174
Log-likelihood (pseudo)	-2181.851	-348.209
R ² (pseudo)	0.108	-

Notes: Model (1) is an ordered Probit models with robust standard errors (in parentheses) and clustering at the group level. Model (2) is a random-effects Tobit model. *, **, and *** indicate statistical significance at the 0.10, 0.05, and 0.01 level, respectively.

Controlling for the valence of emotions when being informed of another individual's contribution, model (1) shows that a more severe punishment is associated with more positive emotional valence. This suggests that the valence of emotions experienced when sanctioning is influenced both by the negative feelings associated with the free-riding of a group member and by the satisfaction of punishing the free-rider. Punishing reduces the disutility experienced when learning about others' free-riding. This finding supports our hypothesis H3, according to which individuals trade off the monetary cost of sanctions and the emotional benefit of punishing free riders. It is also consistent with previous neuroeconomic findings (de Quervain *et al.*, 2004). In addition, model (2) indicates that the greater the number of points assigned,

the greater the physiological arousal, with no additional effect of punishing a cooperator. Additional regressions (available upon request) show that this relationship is significant only for subjects who report neutral or positive feelings (≥ 5).

4.4. Emotional and behavioral impact of the receipt of sanctions

Receiving punishment triggers negative emotions. It also motivates people to revise upward their future contribution levels. This is stated precisely in Result 5.

Result 5. a) *The more strongly an individual is punished, the greater is his SCR and the less positive his reported valence.* b) *The individuals are incited to increase their contribution to avoid both the monetary and emotional cost associated with sanctions.*

Support for Result 5. The valence of emotions, when informed about the sanctions one is receiving, is 8.99 when the individual is not punished (S.D.=1.54) and 3.89 when he is punished (S.D.=2.41). The average SCR is 0.14 (S.D.=0.365) when a participant is not punished and 0.21 (S.D.=0.389) when he is punished.

At the individual level, we estimated two models (see Table A in Appendix C). Model (1) is an ordered Probit model in which the dependent variable is the valence of feelings when informed on receiving sanctions (N=480, log pseudolikelihood=-711.863). Model (2) is a random-effects Tobit model in which the dependent variable is the SCR magnitude during this feedback (N=480, log-likelihood =-331.579). In both models, the independent variables include the number of punishment points received, the number of points received when the subject contributed more than the average, a time trend and gender. Model (1) indicates that the number of punishment points a person receives exerts a strong negative impact on the valence of his emotions (coeff.=-0.520, $p < 0.001$). This effect is reinforced when punishment is anti-social (coeff.=-1.100, $p < 0.001$). Moreover, model (2) shows that the physiological arousal is positively related to the number of points received (coeff.=0.065, $p = 0.051$), with no additional effect of anti-social punishment ($p=0.260$).¹⁵

As observed in previous studies (Fehr and Gächter, 2000; Masclet *et al.*, 2003), the change in contributions from one period to the next depends on the number of punishment points the individual received in the previous period. Indeed, a participant who was not sanctioned in the

¹⁵ Separate regressions of SCR for subjects who report a valence below 5 show a significant positive effect of the number of points received ($p=0.032$), but not the regression for those whose valence is equal or higher than 5 ($p=0.254$). A higher SCR when receiving a feedback on sanctions is associated with negative feelings.

previous period, decreases his contribution in the current period by on average -0.12 ECU (N=388; S.D.=1.47); receiving one or two points increases contribution by 1.11 ECU (N=47; S.D.=3.74); receiving three points or more increases it by 5.76 (N=21; S.D.=8.35). To measure how the emotions experienced during various stages of the game influence subsequent cooperation levels, we estimated four random-effects GLS models with robust standard errors and clustering at the group level in which the dependent variable is the change in individuals' contributions between periods $t-1$ and t .¹⁶ In Table 6, models (1) is for the Baseline and models (2) for the Sanction treatment. The regressions are conducted separately for the subjects who contributed an amount greater than or equal to the group average ("high contributors") and those who contributed less than the average ("low contributors"). The explanatory variables include the SCR magnitude and the valence of emotion when contributing in period $t-1$, in addition to the difference between the subject's contribution and the mean contribution of group members in $t-1$. In the Sanction treatment, the regressions also take into account the number of points received, as well as the SCR magnitude and the associated valence when informed about sanctions in $t-1$.

Table 6. Emotions, punishment, and the change in contribution between periods $t-1$ and t

Dependent variable: Change in contribution between $t-1$ and t	Baseline treatment (1)		Sanction treatment (2)	
	High contributors	Low contributors	High contributors	Low contributors
Difference between own contribution and the average contribution in $t-1$, $(c_i - \bar{c}_{-i})^{t-1}$	-0.750*** (0.088)	-0.523*** (0.149)	-0.065 (0.050)	-0.800*** (0.056)
Physiological arousal when contributing in $t-1$	-0.535 (0.611)	0.738 (1.450)	-0.216 (0.228)	0.204 (1.348)
Valence of emotions when contributing in $t-1$	0.248*** (0.077)	-0.204 (0.238)	0.013 (0.033)	0.199 (0.227)
Punishment points received in $t-1$	-	-	-1.258*** (0.397)	0.377*** (0.107)
Physiological arousal at the time of receipt of sanctions in $t-1$	-	-	0.266 (0.270)	3.391** (1.677)
Valence of emotions at the time of receipt of sanctions in $t-1$	-	-	-0.170** (0.078)	0.086 (0.288)
Time trend (t)	0.007 (0.056)	-0.041 (0.068)	-0.029 (0.021)	-0.013 (0.027)
Male	-1.028 (1.025)	-0.767 (1.632)	-0.458** (0.212)	0.058 (0.859)

¹⁶ We have also estimated our models with fixed effects and conducted Hausman tests. The results of the tests are not significant. These findings indicate that it is therefore safe to use random effects models.

Constant	-0.899 (1.085)	2.524 (1.652)	1.882* (0.979)	-3.040*** (1.158)
N	260	196	411	45
R ²	0.237	0.104	0.103	0.729

Notes: All the models are random-effects GLS models with robust standard errors (in parentheses) and clustering at the group level. *, **, and *** indicate statistical significance at the 0.10, 0.05, and 0.01 level, respectively.

In the Baseline but not in the Sanction treatment, the more positive the emotions when contributing in $t-1$, the more the high contributors increase (or rather the less they decrease, since on average the change is negative) their contribution in the current period. The sanctions received, the SCR magnitude, and the valence of emotions when being informed of the quantity of sanctions received are strong predictors of changes in behavior. Receiving sanctions results in low contributors increasing their contribution, whereas the high contributors adjust it downward in response to sanctions. Controlling for this direct effect of sanctions, the more physiologically aroused the low contributors were when informed about how much they have been punished, the greater the upward adjustment of contribution. The valence of feelings exerts no significant influence. While SCR magnitude does not predict the high contributors' decisions, the less negative their self-reported feelings when being informed about the sanctions they receive, the lower the upward adjustment of their contribution. In addition to the monetary consequences of sanctions, individuals dislike being punished and they adjust their behavior accordingly.

5. DISCUSSION AND CONCLUSION

We consider the emotional patterns associated with behavior in the Voluntary Contributions Mechanism. We record two measures of emotion, skin conductance response and self-reported valence of emotional state, as our subjects choose how much to cooperate, observe others' levels of cooperation, sanction others, and are themselves sanctioned.

We observe a number of patterns. First, we find that higher contributions are associated with more positive self-reported emotions and less arousal on the part of the contributor. Second, individuals are more aroused and report less positive emotions if they learn that other players have not cooperated. After controlling for the contribution of others, learning that one contributes more than another group member affects the valence of feelings negatively, while learning that one contributes less increases it. Third, a higher arousal and more negative feelings when learning the difference between one's own and others' contributions predict the

punishment behavior in the Sanction treatment. For a given difference in contributions levels, the emotions experienced by an individual contribute to his decision to sanction another subject. Sanctioning non-cooperators mitigates and partially relieves the negative emotions associated with others' free-riding. Lastly, we find that receiving punishment is associated with an emotional arousal and an adverse emotional state. Those who are more emotionally aroused when receiving sanctions have a stronger tendency to modify their subsequent behavior in the direction of greater cooperation.

Emotions, therefore, appear to play a role in the emergence and maintenance of cooperation. The emotional responses of group members are, however, not sufficient on their own to allow the group to attain a cooperative outcome. Appropriate institutions must be in place to allow emotions to express themselves in behavior that can be communicated to other parties (Ostrom,1990). Emotions cannot make cooperation emerge when no sanctions or other forms of communication are available, as the Baseline treatment shows. In the Sanction treatment, the mechanism whereby emotions induce cooperation operates in conjunction with the ability to punish free-riders. The application of punishment can then set in motion a "virtuous emotional circle" that promotes cooperation. Cooperators experience positive feelings, which become even more positive if they learn that others are also cooperating. When they find themselves in groups with free-riders, they relieve negative emotions when they punish free riders. The free-riders experience negative emotions in response to punishment, and increase their subsequent level of cooperation. Thus, both positive and negative emotional reinforcement, in conjunction with the availability of sanctions, encourages cooperation.

While our view is that the measures of emotional state that we have studied here help in understanding economic decision-making, we acknowledge that they have some limitations. The SCR and the self-reports on hedonic valence are two separate measures of emotions. At some specific moments in our experiment, these two measures are correlated, especially when assigning sanctions and when receiving sanctions. But at some other moments, they are independent: for example, when subjects decide on their contribution level, those who are more physiologically aroused are not those who report the most pleasant or the most unpleasant feelings. In such cases the interpretation of the emotional state is more difficult. However, despite limitations, these measures of emotions provide useful information for explaining the dynamics of cooperation and the influence of altruistic sanctions in social

dilemmas. This suggests that further research should be developed on the role of emotions on decision-making in groups

References

- Andreoni J. (1988). Why free ride? Strategies and Learning in Public Goods Experiments. *Journal of Public Economics* 37, 291-304.
- Andreoni, J. (1990). Impure Altruism and Donations to Public Goods: A Theory of Warm Glow Giving. *Economic Journal C*, 464-477.
- Barrett, L. F., L. Robin, P. R. Pietromonaco, and K.M. Eysell (1998). Are women the 'more emotional' sex? Evidence from emotional experiences in social context. *Cognition & Emotion* 12, 555-578.
- Barrett, L. F., and J.A. Russell (1999). Structure of current affect. *Current Directions in Psychological Science*, 8, 10-14.
- Ben-Shakhar, G., G. Bornstein, A. Hopfensitz, and F. van Winden (2007). Reciprocity and emotions in bargaining using physiological and self-report measures. *Journal of Economic Psychology* 28(3), 314-323.
- Blau, P.M., (1964). *Exchange and power in social life*, John Wiley and Sons, Inc., New York.
- Bosman, R. and F. van Winden (2002) Emotional Hazard in a Power-to-take Experiment. *Economic Journal* 112, 147-169.
- Boucsein, W. (1992). *Electrodermal Activity*, Plenum Series in Behavioral Psychology and Medicine, Plenum.
- Bowles, S. and H. Gintis (2001). *The Economics of Shame and Punishment*. Mimeo, University of Massachusetts.
- Brekke K., Kverndokk S., Nyborg K. (2003). An economic model of moral motivation, *Journal of Public Economics*, 87, 1967-1983.
- Bradley, M.M., M. Codispoti, B.N. Cuthbert, and P.J. Lang, (2001a). Emotion and motivation I: Defensive and appetitive reactions in picture processing. *Emotion* 1, 276-298.
- Bradley, M.M., M. Codispoti, D. Sabatinelli, and P.J. Lang (2001b). Emotion and motivation. II: Sex differences in picture processing. *Emotion* 1, 300-319.
- Carpenter, J.P. (2007). The demand for punishment. *Journal of Economic Behavior & Organization* 62(4), 522-554.
- Coleman, J.S. (1990). *Foundations of social theory*, Belknap, Cambridge, MA.
- Coricelli, G., M. Joffily, C. Montmarquette and M.C. Villeval, 2010. "Cheating, Emotions and Rationality: An Experiment on Tax Evasion", *Experimental Economics* 13 (2), 226-247.
- Critchley H.D., Elliott R., Mathias C.J. and Dolan R.J. (2000). Neural activity relating to generation and representation of galvanic skin conductance responses: a functional magnetic resonance imaging study. *Journal of Neuroscience* 20, 3033-3040.
- Croson, R. (2007). Theories of commitment, altruism and reciprocity: evidence from linear public goods games. *Economic Inquiry* 45, 199-216.
- Damasio, A.R. (1994). *Descartes' Error: Emotion, Reason, and the Human Brain*, Grosset Putnam Publishers, New York.

- Dawson, M.E., A.M. Schell, and D.L. Filion (2000). The electrodermal system, in *Handbook of Psychophysiology*, J. T. Cacioppo, L.G. Tassinary, G.G. Berntson, Eds., 2nd ed. Cambridge University Press, Cambridge.
- De Quervain, D. J. F., U. Fischbacher, V. Treyer, M. Schellhammer, U. Schnyder, A. Buck, and E. Fehr (2004). The neural basis of altruistic punishment. *Science* 305, 1254-1258.
- Edelberg, R. (1967). Electrical properties of the skin, in *Methods in Psychophysiology*, C.C. Brown, Ed., Williams & Wilkens, Baltimore, 1-53.
- Elster, J. (1998). Emotions and Economic Theory. *Journal of Economic Literature* 36(1), 47-74.
- Falk, A., and U. Fischbacher (2006). A theory of reciprocity. *Games and Economic Behavior* 54(2), 293-315.
- Fehr, E. and S. Gächter (2000). Cooperation and punishment in public goods experiments. *American Economic Review* 90, 980-994.
- Fehr, E. and S. Gächter (2002). Altruistic punishment in humans, *Nature* 415, 10 January, 137-140.
- Fehr, E. and Schmidt, K. M., (1999). "A Theory of Fairness, Competition and Cooperation." *Quarterly Journal of Economics* 114, 817-868.
- Fehr, E., and K.M. Schmidt (1999). A Theory of Fairness, Competition and Cooperation. *Quarterly Journal of Economics* 114, 817-868.
- Figuières C., D. Masclet and M. Willinger. Weak Moral Motivation Leads to the Decline of Voluntary Contribution. Forthcoming in *Journal of Public Economic Theory*.
- Frijda, N. H. (1986). The emotions. Cambridge: Cambridge University Press.
- Fujita, F., E. Diener, and E. Sandvik (1991). Gender differences in negative affect and well-being: The case for emotional intensity. *Journal of Personality and Social Psychology* 61, 427-434.
- Gächter, S., E. Renner, and M. Sefton (2008). The Long-Run Benefits of Punishment. *Science* 322, 5 December, 1510.
- Glimcher P.W., C.F. Camerer, E. Fehr, and R.A. Poldrack (Eds.). 2009. *Neuroeconomics. Decision-making and the brain*. Amsterdam: Elsevier.
- Greiner, B. (2004). An Online Recruitment System for Economic Experiments. in *Forschung und wissenschaftliches Rechnen*, K. Kremer, V. Macho, Eds., GWDG Bericht 63, Göttingen, Ges. für Wiss. Datenverarbeitung.
- Gross, J. M. and Barrett, L. F. (2011) Emotion Generation and Emotion Regulation: One or Two Depends on Your Point of View. *Emotion Review* 3(1), 8-16.
- Hermann, B., S. Gächter, and C. Thöni (2008). Antisocial Punishment Across Societies. *Science* 319, 7 March, 1362-1367. *Science* 319, 1362 (2008).
- Homans, G.C. (1961). *Social behavior: Its elementary forms*, Harcourt Brace & World, New York.
- Houser, D., E. Xiao, K. McCabe, and V. Smith (2008). When punishment fails: Research on sanctions, intentions and non-cooperation. *Games and Economic Behavior* 62(2), 509-532.

- Hopfensitz, A. and E. Reuben (2009). The importance of emotions for the effectiveness of social punishment. *Economic Journal* 119, 1534-1559.
- Kreps D., Milgrom P., Roberts J., Wilson R., (1982), "Rational cooperation in the finitely repeated Prisoners' Dilemma." *Journal of Economic Theory*, 27, 1982,245-52.
- Ledyard, J. (1995), Public goods: A survey of experimental research, in J.H. Kagel and A.E. Roth (eds.), *The handbook of experimental economics*, Princeton University Press.
- Loewenstein, G. (1996). Out of Control: Visceral Influences on Behavior. *Organizational Behavior and Human Decision Processes*, 65(3), 272-292.
- Loewenstein, G. F., C.K. Hsee, E.U. Weber, and N. Welch (2001). Risk as feelings. *Psychological Bulletin*, 127, 267–286.
- Marwell, G. and R.E. Ames (1979). Experiments on the provision of public goods. I: Resources, interest, group size, and the free-rider problem. *American Journal of Sociology* 84(6), 1335–1360.
- Masclet, D., C. Noussair, S. Tucker, and M.C. Villeval (2003). Monetary and Non-Monetary Punishment in the Voluntary Contributions Mechanism. *American Economic Review* 93, 366-380.
- Masclet, D., and M.C. Villeval (2008). Punishment, Inequality, and Welfare: A Public Good Experiment. *Social Choice and Welfare* 31(3), 475-502.
- McRae, K., K.N. Ochsner I.B. Mauss, J.J.D. Gabrieli, J.J. Gross. (2008). Gender Differences in Emotion Regulation: An fMRI Study of Cognitive Reappraisal. *Group Processes & Intergroup Relations* 11(2), 143-162.
- Meshulam, M., E. Winter, G. Ben Shahaar, and Y. Aharaon. Rational Emotions in the Lab. Forthcoming in *Social Neuroscience*.
- Milinski, M., and B. Rockenbach (2006). The efficient interaction of indirect reciprocity and costly punishment. *Nature* 444, 7 December, 718-723.
- Milinski, M., and B. Rockenbach (2008). Human behaviour: Punisher pays. *Nature* 452, 20 March, 297-298.
- Noussair, C., and S. Tucker (2005). Combining Monetary and Social Sanctions to Promote Cooperation. *Economic Inquiry* 43(3), 649-660.
- Nyborg, K. (2000). Homo economicus and homo politicus: Interpretation and aggregation of environmental values. *Journal of Economic Behavior & Organization* 42, 305--322.
- Ostrom, E. (1990). *Governing the Commons: The Evolution of Institutions for Collective Action*, Cambridge Univ. Press, Cambridge.
- Rabin, Matthew, (1993). Incorporating Fairness into Game Theory and Economics. *American Economic Review* 83, 1281-1302.
- Rand, D.G., A. Dreber, T. Ellingsen, D. Fudenberg, and M. A. Nowak (2009). Positive Interactions Promote Public Cooperation. *Science* 325, 1272-1275.
- Rege M.S and K. Telle, (2004). The impact of social approval and framing on cooperation in public good situations, *Journal of Public Economics* 88, 1625-1644.

- Reisenzein, R. (1994). Pleasure-arousal theory and the intensity of emotions. *Journal of Personality and Social Psychology* 67, 525–539.
- Rilling, J.K., D.A. Gutman, T.R. Zeh, G. Pagnoni, G.S. Berns, and C.D. Kilts (2002). A neural basis for social cooperation. *Neuron* 35, 395-405.
- Rilling, J.K., A.G. Sanfey, J.A. Aronson, L.E. Nystrom, and J.D. Cohen (2004). Opposing BOLD responses to reciprocated and unreciprocated altruism in putative reward pathways. *NeuroReport* 15(15), 2539-2543.
- Russell, J.A. (1980). A Circumplex Model of Affect. *Journal of Personality and Social Psychology* 39(6), 1161-1178.
- Sanfey, A.G., J.K. Rilling, J.A. Aronson, L.E. Nystrom, and J.D. Cohen (2003). The neural basis of economic decision-making in the Ultimatum Game. *Science* 300, 5626, June 13, 1755-1758.
- Yamagishi, T. (1986). The Provision of a Sanctioning System as a Public Good. *Journal of Personality and Social Psychology*, 51(1), 110-116.
- Zeiliger, R. A Presentation of Regate, Internet Based Software for Experimental Economics, <http://www.gate.cnrs.fr/~zeiliger/regate/RegateIntro.ppt>, (GATE, Lyon, 2000).

Appendix A. Instructions to participants

The following are the instructions, translated from the original French, for the Sanction treatment. The instructions for the Baseline treatment are a subset of those printed here. The instructions describe the timing of decisions, the decision rules, and the process of payoff determination. Subjects had to answer several control questions at the end of the instructions to ensure that they understood the rules of the game.

--- Text of Instructions

Thank you for taking part in this experiment during which you can earn money. Your earnings depend on your decisions and on the decisions of the other participants with whom you will interact.

All of the transactions during the experiment and your entire earnings will be calculated in ECU (Experimental Currency Units). At the end of the experiment the total amount of ECU you have earned during the session will be converted to Euros and paid to you in cash in a separate room by somebody who is not aware of the content of the experiment, according to the following rules:

- Your final payoff in ECU consists of the total of your payoffs in each of the 20 periods comprising this session.
- This final payoff in ECU will be converted into Euros at the rate: 100 ECU = 2 Euros.
- In addition, you will be given a show-up fee of 5 Euros.

At the beginning of the session, the participants are divided into groups of four. You will therefore interact with three other participants. **During the 20 periods, you will interact with the same persons.** You will never be informed of the identity of these persons.

The four participants belonging to a group can participate in a project, by contributing to a group account that will be shared among them. The amount of this group account is determined by the sum of the individual contributions of the four members of the group.

Description of each period

Each period consists of two stages.

First stage. At the beginning of each period each participant receives an endowment of 20 ECU.

You, as well as the three other members of your group, simultaneously decide how much of your endowment you will contribute to the project, by indicating a number between 0 and 20, inclusive.

After all group members have made their decision, your screen will show you the total amount of ECU contributed to the project by the members of your group (including your own contribution). You are also informed of the amount contributed by each of the three other members of your group to the project. Note: the order in which each contribution is displayed is changed randomly in each period (in other words, the number that appears first on your screen does not always correspond to the decision of the same player). You are also informed about your first-stage payoff .

Your income consists of two parts:

- the amount of your endowment which you have kept for yourself (that is, 20 – your contribution to the project),
- the income from the project: this income represents 40% of the total contribution of all four group members to the project

Your income in ECU in the first stage of each period is therefore:

$$(20 - \text{your contribution to the project}) + 40\% * (\text{total contributions of the group to the project})$$

The payoff of each group member is calculated in the same way, which means that each group member receives the same income from the project.

Suppose the total of the contributions of all group members is 60 ECU. In this case each member of the group receives an income for the first stage from the project of 40% (of 60 ECU) = 24 ECU. If the total contribution to the project is 9 ECU, then each member of the group receives an income of 40% (of 9 ECU) = 3.6 ECU from the project.

For each ECU that you keep for yourself you earn an income of 1 ECU. For every ECU you contribute to the project instead, the total contribution to the project increases by one ECU. Your income from the project will increase by 40% (of 1 ECU) = 0.4 ECU. The income of the other group members will also rise by 0.4 ECU each, so that the total income of the group from the project rises by 1.6 ECU. This means that your contribution to the project also increases the income of the other group members.

On the other hand you will earn money from each ECU contributed by other members to the project. For each ECU contributed by any member you earn 40% (1) = 0.4 ECU.

Second stage. You can, if you like, indicate your disapproval of members of your group by assigning points that reduce their first-stage payoff. You can assign a particular number of points to a member of your group to express a level of disapproval (10 points for the highest disapproval, 0 points for no disapproval). Each point assigned reduces her first-stage income by 10%. Similarly, your income can be modified if the other members of your group wish to do so.

You decide how many points to give to each of the other three group members to reduce their income or leave it unchanged. You must enter a value for each member, between 0 and 10 points. If you do not wish to change the income of a specific member, then you must enter 0.

If you distribute one point to a member, you reduce his first-stage payoff by 10%; if you assign two points, you reduce his payoff by 20%, etc. The number of points you assign determines by how much you are willing to reduce his first-stage payoff.

If you assign points, you pay a cost in ECU that depends on the number of points you distribute to each subject. The more points you give to any subject, the higher your costs. Your total costs are equal to the sum of the costs of distributing points to each of the other three group members. The following table illustrates the relationship between points distributed to a subject and the cost of doing so in ECU:

<i>Number of punishment points</i>	0	1	2	3	4	5	6	7	8	9	10
Cost to the punisher	0	1	2	4	6	9	12	16	20	25	30

Percentage of reduction of the target's payoff	0	10	20	30	40	50	60	70	80	90	100
--	---	----	----	----	----	----	----	----	----	----	-----

If you assign two points to one group member, this will cost you 2 ECU; if you assign 9 points to another member, this will cost you 25 ECU more; if you give the last subject no points, this does not cost you anything. In this example, the total cost of the assigned points is 27 ECU (2+25+0). These costs will be displayed on your screen.

Your final income in ECU in each period is calculated as follows:

If you received less than 10 points:

$$\text{final income} = (\text{income from the 1st stage}) * [(10 - \text{number of received points}) / 10] - \text{cost of points you assigned}$$

If you received 10 or more points:

$$\text{final income} = - \text{cost of points you assigned}$$

The maximum number of points received that can count against you in a period is 10. For example, if you received 3 points from other group members your first-stage payoff will be reduced by 30%. If you received 4 points from other group members your first-stage payoff will be reduced by 40%. If you received 10 points or more, you will lose 100% of your first-stage payoff. In this case, you make a loss if you have assigned points to other members of your group. The amount of the loss equals the cost of the assigned points.

Therefore, your income in ECU at the end of the second stage can be negative, if the costs of the points you distribute exceed your income from the first stage. You can, however, avoid such losses with certainty through your own decisions.

Additional questions

Several times during each period, we will ask you to describe your feelings. You report your feelings on a scale scoring from 1 (extremely unpleasant feeling) to 10 (extremely pleasant feeling).

You are not paid based on the answers to these questions. They have no influence on the remainder of the session. We do, however, ask you to answer these questions sincerely and with care.

Summary: You receive an endowment. You decide how much you are willing to contribute to a project. You are then informed of the total amount the group contributed, the contribution of each other group member, and your income. You can assign points to express your disapproval. Several times during the period, you answer questions regarding your feelings at the time.

At the end of each period, the next period begins. You receive a new endowment of 20 ECU and you choose your contribution.

* * *

Thank you for answering the questionnaire on your understanding that has been distributed to you. If you have any questions about these instructions, please raise your hand. We will answer your questions in private.

Communicating with the other participants during the experiment is strictly forbidden at the risk of being excluded from the session and from receiving your payment.

Appendix B. Assignment of punishment points

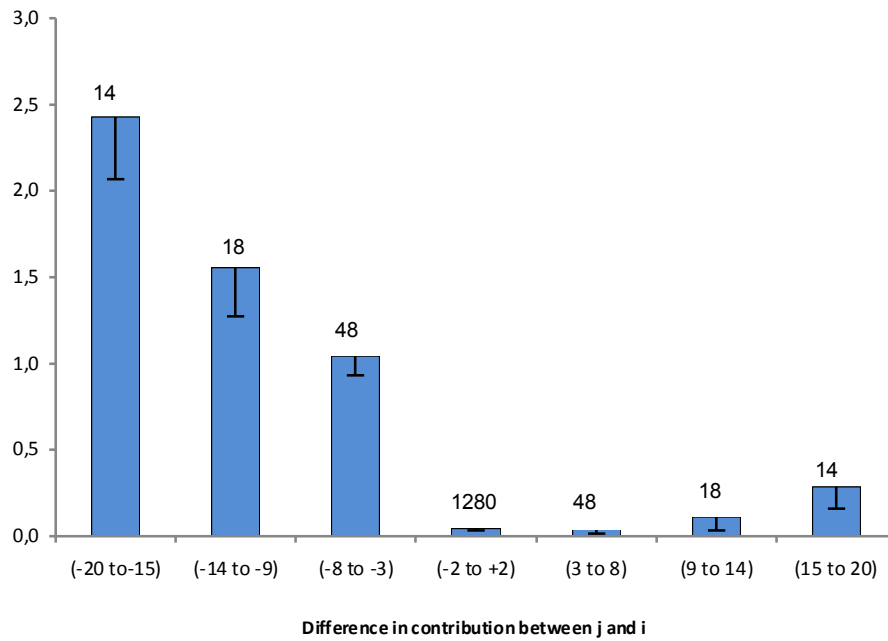


Fig.A. Number of punishment points assigned to j by i as a function of the difference in contribution between j and i

Figure A shows the average number of punishment points assigned by player i to player j . The observations are classified based on the difference between the contributions of players j and i in the current period. The horizontal axis is the difference in contributions, which, because each individual has 20 tokens, ranges from -20 to $+20$. The vertical axis indicates the average assignment of points, which can range from 0 to 10.

Appendix C – The receipt of sanctions

Table A. Determinants of the physiological arousal and the valence of emotions when receiving sanctions

Dependent variable	Valence (1)	SCR (2)
Punishment points received	-0.520*** (0.054)	0.065** (0.033)
Punishment points received by a subject contributing more than the group average	-1.100*** (0.315)	0.309 (0.275)
Time trend	0.015 (0.022)	-0.022*** (0.007)
Male	-0.107 (0.341)	0.292* (0.156)
Constant	-	-0.300** (0.125)
N	480	480
Left-censored observations	-	320
Log-likelihood (pseudo)	-711.863	-331.579
Pseudo R ²	0.100	-

Notes: Model (1) is an ordered Probit model with robust standard errors (in parentheses) and clustering at the group level. Model (2) is a random-effects Tobit model. *, **, and *** indicate statistical significance at the 0.10, 0.05, and 0.01 level, respectively