Fairness Perceptions and Prosocial Emotions in the Power to Take

Ernesto Reuben [*,a] *and Frans van Winden* [b]

[a] *Columbia Business School, Columbia University, 3022 Broadway, New York, NY 10027, USA.*

[b] *Department of Economics University of Amsterdam, Roetersstraat 11, 1018 WB Amsterdam, Netherlands.*

ABSTRACT

This experimental study investigates how behavior changes after receiving punishment. The focus is on how proposers in a power-to-take game adjust their behavior depending on their fairness perceptions, their experienced emotions, and their interaction with responders. We find that fairness plays an important role: proposers who take what they consider to be an unfair amount experience higher intensities of prosocial emotions (shame and guilt), particularly if they are punished. This emotional experience induces proposers to lower their claims. We also find that fairness perceptions vary considerably between individuals. Therefore, it is not necessarily the case that proposers who considered themselves fair are taking less from responders than other proposers. Lastly, we provide evidence that suggests that eliciting emotions through self-reports does not affect subsequent behavior.

Keywords: shame, guilt, punishment, prosocial emotions, proposers, power-to-take game

JEL classification codes: Z13, C92, C72, K42

Psychological classification codes: 2360; 3040

---

[*] Corresponding author: *E-mail address:* ereuben@columbia.edu.

Note: This is the authors' version of a work that was accepted for publication in the *Journal of Economic Psychology*. Changes resulting from the publishing process, such as peer review, editing, corrections, structural formatting, and other quality control mechanisms may not be reflected in this document. Changes may have been made to this work since it was submitted for publication. A definitive version was subsequently published in http://dx.doi.org/10.1016/j.joep.2010.08.001

1. Introduction

By now, it is a well-documented fact that individuals who participate in economic experiments are willing to spend money in order to punish people who treat them unkindly (Fehr et al., 2002). Moreover, several studies have found emotions to be an important motivating factor behind this type of behavior (e.g., Bosman and van Winden, 2002; Sanfey et al., 2003; de Quervain et al., 2004). However, thus far, research has concentrated on the motivations and behavior of the individuals who do the punishing. The goal of this paper is to investigate through an experiment how individuals who experience punishment behave and the role of emotions in motivating their actions. Shedding light on this subject can help us understand why punishing someone sometimes promotes future prosocial behavior but at other times it is ineffective in doing so (e.g., Houser et al., 2008; Reuben and Riedl, 2009) and can even provoke retaliation (Denant-Boemont et al., 2007; Nikiforakis, 2008).

Whereas emotions seem to play an important role in motivating individuals to punish others, it is not clear in which way (if at all) emotions affect the decisions of those that may be punished. Prosocial emotions such as shame and guilt might be important in these situations as most individuals experience this type of emotions when they violate social norms or their own personal values (Haidt, 2003).[1] In this paper, we investigate whether punishment in an economic experiment does indeed trigger prosocial emotions, and if so, how they relate to fairness perceptions and how they affect an individual's reaction to being punished.

For our study, we use a repeated version of the power-to-take game. In this game, the proposer can make a claim on the resources of a responder. Then, the responder can destroy any part (including nothing and everything) of her own resources. Hence, the power-to-take game approximates social environments in which one party can appropriate the resources of another.[2] In the experiment, this game was played for two consecutive periods.

---

[1] Bowles and Gintis (2005) argue that prosocial emotions and notions of fairness evolved specifically to solve problems such as social dilemmas and resource allocation.

[2] A clear example is taxation, but one can also think of monopolistic selling—where the monopolist can appropriate the consumers' surplus and they in turn can destroy by buying less—and common

An important part of the experimental design is that we measure the emotions of proposers after they observe the amount destroyed by the responders in the first period. This allows us to study how the proposers' emotional reaction to being punished by the responder in the first period (through destruction) affects their decision in the second period. We find that emotions play an important role in determining how proposers change their decision from one period to the next. Proposers that experienced punishment and experienced high intensities of prosocial emotions lower their claim. Proposers who did not experience punishment and felt high intensities of regret increase their claim. Furthermore, we find that the experience of prosocial emotions does not simply depend on the size of the proposer's claim but is associated with the proposer's *perception* of what is fair.

The experimental design also allows us to observe how fairness perceptions vary among proposers. Current theories that incorporate a notion of fairness typically assume that people agree on what is fair or unfair (see Fehr and Schmidt, 2006). Although we find support for the idea that fairness matters, we do not find much support for the presence of a clearly shared fairness perception.

Finally, given that we use self-reported measures of emotions, we also run a control treatment to test whether measuring emotions affects the subjects' subsequent behavior. We do not find evidence that this is the case.

The paper is organized as follows. In Section 2 we review related research, while the experimental design is described in Section 3. Results are presented in Section 4, and Section 5 concludes.

2. Related Research

Our work is related to three different areas of research. First, our work is related to research focusing on proposer behavior in ultimatum games. Second, we contribute to the growing literature on the economic significance of emotions and their role in

---

agency situations—where the principal decides on the incentive scheme for an agent who may destroy by choosing a low effort level, which is costly to the agent because it conflicts with the selected incentive scheme (see, Bosman and van Winden, 2002; Bosman et al., 2005).

the enforcement and compliance of norm and values. Third, this paper is related to studies concerned with how fairness perceptions affect behavior.

Even though there are no studies exploring proposer behavior in the power-to-take game, there is considerable research on proposer behavior in the ultimatum-bargaining game (Güth et al., 1982).[3] Space constraints allow only a quick overview of the main findings (for an extensive summary see Camerer, 2003). Broadly speaking, proposers seem to be motivated by a combination of "profit-maximizing" and "nonprofit-maximizing" behavior. Profit-maximizing behavior—in the sense of going for the highest offer that will not be rejected—is clearly observed since proposers adjust their offers depending on the likelihood of rejection. For example, offers go down in cases where responders are less likely to reject, such as when the total size of the pie is unknown (Camerer and Loewenstein, 1993; Straub and Murnighan, 1995; Rapoport et al., 1996), when there is competition among responders (Grosskopf, 2003; Fischbacher et al., 2009), or in the extreme case of a dictator game in which responders cannot reject at all (Forsythe et al., 1994). However, the fact that, even in completely anonymous dictator games there are positive offers seems to indicate that there is a degree of nonprofit-maximizing (perhaps fairness-guided) behavior.[4] By analyzing whether emotions play a role in proposer behavior, we can shed light on the relative importance of profit-maximizing and nonprofit-maximizing factors in the decisions of proposers.

There are various papers that explore the potential role of emotions—in particular, the avoidance of guilt and shame—in the observed nonprofit-

---

[3] The power-to-take game differs in 3 important ways from the well-known ultimatum game. First, in the power-to-take game each participant has an endowment. Second, in this game only the endowment of the responder is at stake. And third, the responder can destroy any amount of her endowment. See also Andreoni et al. (2003) for a similar game regarding the third difference.

[4] Further but less conclusive evidence of nonprofit-maximizing behavior is provided by Lin and Sunder (2002) and Brenner and Vriend (2006) who find that, given the responders' reactions, the proposers' offers are higher than the profit-maximizing offer. Henrich et al. (2001) also find a few hunter-gatherer societies where such offers are evident. However, we should note that the most comprehensive datasets of behavior in ultimatum games indicate that the offer that maximized expected profits is in fact the equal split, in which case one cannot convincingly argue that proposers are motivated by reasons other than profit-maximization (see, Güth et al., 2003; 2007).

maximizing behavior of individuals. For instance, feelings of shame are conjectured to explain why proposers in dictator games take more as the level of anonymity increases (see, Hoffman et al., 1996; Dana et al., 2006). Similarly, shame is also attributed to the positive effect of nonmonetary sanctions on norm compliance (e.g., Masclet et al., 2003; Noussair and Tucker, 2005; Xiao and Houser, 2005). With respect to guilt, there is a growing experimental literature testing theoretical models of guilt aversion (Charness and Dufwenberg, 2006). In this case, guilt is modeled as disutility caused when individuals do not fulfill the expectations of others. Experimental evidence in favor or this interpretation of guilt is mixed: Charness and Dufwenberg (2006) and Reuben et al. (2009) find that changes in the expectations of others do produce changes in the amount returned by responders in a trust game, whereas Ellingsen et al. (2010) fail to find an effect. In this paper, we complement this literature by actually measuring emotions. This way we can confirm, as conjectured, that the observed behavioral changes are indeed driven by shame and guilt.

There are a small but growing number of studies that measure emotions and relate them to costly punishment. However, they concentrate on the emotions and behavior of responders. By now, there is strong evidence suggesting that the destruction of income by responders is induced by emotions such as anger, irritation, and contempt. For example, Bosman and van Winden (2002) introduced the power-to-take game with the specific purpose of investigating the importance of emotions for reciprocity. Their results show that the destruction of income is related to the intensity of experienced anger, which in turn is positively related to the actual take rate and negatively related to the expected take rate (see also, Bosman et al., 2005; Reuben and van Winden, 2008). Recently, evidence has been reported that responders who punish exhibit heightened arousal levels (Sanfey et al., 2003; Ben-Shakhar et al., 2007). More revealingly, effective punishment has been found to be correlated with neural activity in areas of the brain that are involved in the processing of rewards (e.g., de Quervain et al., 2004). With this paper, we take an important next step in this line of research by studying how emotions affect the decisions of proposers. To the best of our knowledge, the only paper that investigates the neural activity of proposers is Spitzer et al. (2007). They find that the threat of punishment is correlated with brain activity in areas that are involved in social cognition and interestingly also the insula, which is associated with emotions such as anger and fear. Our focus differs from that of Spitzer et al. (2007) in that we concentrate on the proposers' emotional reaction *after* receiving punishment and how it affects their subsequent behavior.

Finally, this study is related to research on the perception of fairness norms. A few researchers explicitly asked for the fairness perceptions of individuals (e.g., Binmore et al., 1991), but most have not analyzed how fairness perceptions interact with behavior. Pillutla and Murnighan (1996) do measure the fairness perceptions of responders in an ultimatum game and find that perceived unfairness is related to the rejection of offers. However, more recent research by Reuben and van Winden (2008) indicates that, once the effect of the offered amount and the responder's expected offer are taken into account, the perceived unfairness of an offer has no longer a significant effect on responder behavior. Nonetheless, fairness perceptions might play a significant role when it comes to proposer behavior.

3. The Experiment

We use a repeated version of the power-to-take game. Subjects play the game for two periods. In each period, one subject, the proposer (with endowment E_{prop}), is matched with another subject, the responder (with endowment E_{resp}). The game consists of two stages. In the first stage, the randomly chosen proposer decides on the "take rate" $t \in [0, 1]$, which is the fraction of the responder's endowment after the second stage that is transferred to the proposer. In the second stage, the responder decides to destroy a part $d \in [0, 1]$ of her own endowment E_{resp}. The payoff of the proposer equals her endowment plus the transfer from the responder: $E_{prop} + t(1 - d)E_{resp}$. The responder's payoff equals the part of her endowment that she does not destroy minus the amount transferred to the proposer: $(1 - t)(1 - d)E_{resp}$. In the experiment all endowments were equal ($E_{resp} = E_{prop}$).

In each of the two periods, subjects are randomly assigned to either the proposer's role or the responder's role. Each proposer is randomly paired with a responder using a perfect-stranger matching protocol. Note that this eliminates any incentive to build a reputation. In addition, this procedure produces a group of subjects that had the same role in both periods and another group that switches roles between periods. Although we are primarily interested in the behavior of subjects who are proposers in both periods, we designed the experiment to be a sequence of identical one-shot games, and therefore, we randomized roles in both periods (i.e., we worried that behavior might be different from a one-shot setting if proposers know they will keep their role for two consecutive periods and that they will face a subject who will be responder for two periods).

We briefly discuss the theoretical predictions of this game. Traditional economic theory, assuming own-payoff maximization, predicts that a proposer will choose to take essentially all of the responder's endowment and that the responder will not destroy. However, previous work has shown that this is not the case. Responders consistently destroy some or all of their endowment when faced with high take rates, and proposers hardly ever choose to take all of the responder's endowment (van Winden, 2001). In order to explain behavior in this and similar games, researchers have constructed models that incorporate different kinds of fairness norms (for a review see, Fehr and Schmidt, 2006).[5] However, it is not the aim of this paper to test the performance of individual models of social preferences. Instead, we wish to investigate what motivates proposers without making a specific assumption on the type of fairness norm that they might be using. This way, we hope to improve our knowledge of proposer behavior in order to produce more accurate theoretical models in the future.

3.1 Treatments

We ran two different treatments. In the first treatment, we use self-reports during the experiment to measure the proposers' emotions. We refer to this treatment as

[5] Some of these models predict remarkably well the behavior of responders in the power-to-take game. Although, to the best of our knowledge, none of these models simultaneously explain proposer behavior in both the ultimatum and power-to-take game. In particular, if a model is calibrated to explain proposer behavior in the ultimatum game, it predicts that in the power-to-take game proposers take considerably less than they actually do. For a more thorough discussion see Reuben (2006).

Treatment E. The second treatment is a control treatment used to ensure that measuring emotions does not affect subsequent behavior. In other words, we use the same game and procedures as in Treatment E except that we do not measure the proposers' emotions. We refer to the second treatment as Treatment No-E. In both treatments, we measure expectations and fairness perceptions.

In each period in Treatment E, the proposers' emotions towards the responder are measured after proposers observe the destruction rate. Emotions are measured simultaneously by providing subjects with a list of fourteen emotions and asking them to report on a seven-point scale with what intensity they experienced each emotion.[6] The scale ranges from "no intensity at all" (1) to "very intensely" (7). The list includes the following emotions: pride, envy, anger, guilt, joy, shame, irritation, gratitude, surprise, contempt, disappointment, admiration, regret, and sadness. A variety of emotions is used to avoid pushing subjects in a particular direction. Subjects were not told in the initial instructions that they will be asked to report their emotions. However, they were told that they will be asked to answer a few questionnaires and that their responses would not be revealed to others. The screen used to elicit emotions is available in the online appendix (Figure A3).

Self-reported measures of emotions are commonly used in social psychology (Robinson and Clore, 2002). Furthermore, in this paper, the use of self reports has an important advantage. Namely, it allows us to study prosocial emotions such as shame and guilt, for which a clear physiological reaction pattern has not been identified (Adolphs, 2002).[7] Since we conjecture that prosocial emotions are important for the study of norm enforcement, we require a technique that is capable of clearly measuring them. This is not to say that self reports do not have limitations. In particular, one could worry that subjects do not report their true

[6] The precise wording (translated from Dutch) was: "When you saw the percentage chosen by B [the responder], with what intensity did you feel the following emotions? Note that there are no right or wrong answers. Do not take too much time and give your first impression of how you felt."

[7] The closest neurological correlate for a prosocial emotion is reported by Takahashi et al. (2004), who find a correlation between self-reported guilt and the medial prefrontal cortex and posterior superior temporal sulcus (both brain areas are associated with the theory of mind).

emotions and instead report a different emotional reaction—for example, what they think the experimenter expects. However, we think that in this study this type of conduct is unlikely to arise as the list of fourteen emotions makes it hard for a subject to identify what the emotions of interest are. Reassuringly, considerable research has demonstrated that self-reported emotional experiences are correlated with various physiological measures like heart rates, facial movements, and brain activation (Bradley and Lang, 2000; Bjork et al., 2004). For example, self-reported anger has been shown to be correlated with measures of physiological arousal in the power-to-take game (Ben-Shakhar et al., 2007). This evidence lends support to our use of self-reports as a trustworthy measure of emotional intensity.

In both treatments, the proposers' expected destruction rate is measured by asking proposers, after they choose a take rate, to indicate the most likely value for d.[8] Although expectations can be incentivized by paying subjects depending on their accuracy, we opted for self-reported measures of the expected destruction rate for two reasons: first, incentivized procedures require complicated elicitation techniques (Offerman et al., 2009) that might distract subjects from the game they are playing,[9] and second, to the best of our knowledge studies comparing incentivized versus non-incentivized beliefs have not found that the latter are biased (e.g., Friedman and Massaro, 1998; Sonnemans and Offerman, 2001; and for a more detailed discussion see Blanco et al., 2008).

Lastly, we measure the subjects' fairness perceptions by asking them to indicate the fair amount a proposer should take (as a percentage of the responder's endowment).[10] In Treatment E, we asked all subjects to report their "fair take rate" at the end of the experiment in a debriefing questionnaire. The same procedure was used for subjects that kept the same role in both periods in Treatment No-E (the

---

[8] We measure expectations in this way since subjects might have difficulty reporting a probability distribution over the interval [0, 1]. The question used (translated from Dutch) was: "What percentage do you think B [the responder] is going to choose?"

[9] We wanted to keep subjects focused on the game as we might be already distracting them with self-reports of emotions and beliefs.

[10] The precise question (translated from Dutch) was: "What would be in your opinion, the *fair* percentage that A [the proposer] should choose in this experiment?"

subjects whose behavior we are interested in). Subjects that switched roles in Treatment No-E we asked for their fair take rate at the beginning of the game, specifically, after they know their role in the first period but before they submit their take rate. By asking some subjects for their fair take rate before they play the game, we can test whether playing the game affects the subjects' fairness perceptions.

We decided to use self-reported fair take rates because our main interest is to evaluate the effect on behavior of each subject's subjective opinion of what is fair. Needless to say, this type of belief cannot be incentivized and therefore must be self-reported. As shown by Krupka and Weber (2008) one can elicit fairness norms in an incentive compatible way by having subjects guess what other subjects think is the fair action to take. The drawback of this approach is that it fails to capture disagreement between what the subject thinks is fair and what the subject thinks others think is fair.[11] We opted for the individual (non-incentivized) fairness measure as we think it is more relevant when evaluating the subjects' emotional response.

3.2 Experimental Procedures

The computerized experiment was run in October 2003, May 2006 (Treatment E), and March 2009 (Treatment No-E) at the CREED laboratory of the University of Amsterdam. In total 308 undergraduate students participated in the experiment. About one third of the subjects were women (36% in Treatment E and 38% in Treatment No-E). Moreover, 55% were students of economics (49% in Treatment E and 67% in Treatment No-E) and the rest were students from other fields such as biology, political science, and law. Subjects received a show-up fee of 2.5 euros, independent of their earnings in the experiment, and 10 euros as endowment in each of the two periods. On average, subjects were paid 21.18 euros. The standard deviation of earnings was 8.67 euros and the minimum and maximum earnings were 2.50 euros and 39.50 euros respectively. The experiment took one hour.

---

[11] This two are not necessarily the same. For example, an individual in the United Sates might think that low-income families pay a fair share of taxes, but being in the minority, if she is asked what others think is fair, she should answer that they pay too much (see Bowman, 2009).

After arrival in the lab's reception room, each subject drew a card to be randomly assigned to a seat in the laboratory. Once everyone was seated, the instructions for the experiment were read out loud. Subjects were told that the experiment consisted of two independent parts (each part being one of the two periods of the game). We emphasized the fact that their choices in the first part of the experiment would not affect their earnings in the second part. Furthermore, it was explained that the instructions for each part would be given at the beginning of the respective part. At this point, the power-to-take game was described and the subjects completed a few exercises to confirm their understanding of the game. Subjects were informed of their role in the first period of the game (proposer or responder) by opening an envelope on their desk. Subsequently, they played the first period of the power-to-take game via the computer.[12] The instructions for the second part consisted of simply informing subjects that they would play the same game once again. We stressed that they would play against a different person and that their role would be randomly determined once again. Thereafter, subjects played the second period of the two-period power-to-take game. A translation of the instructions can be found in the online appendix.

4. Results

In this section, we present the experimental results. First we provide descriptive statistics and compare behavior across treatments. We pay particular attention to how the responders' behavior affects the proposers' take rate choice. Having established that emotion elicitation does not affect proposer behavior, we study how the proposers' expectations and emotions in period one influence their choice in period two. Lastly, we investigate how fairness perceptions influence the proposers' decision. Unless it is otherwise noted, we use the data from subjects who are proposers in *both* periods. Descriptive statistics for subjects who play as proposer for only one period are available in the online appendix (Table A1).

[12] The experiment was programmed and conducted with the software z-Tree (Fischbacher, 2007).

TABLE 1 – DESCRIPTIVE STATISTICS FOR PROPOSERS

	Treatment E		Treatment No-E	
Variable	Period 1	Period 2	Period 1	Period 2
Take rate	53.8%	59.0%	61.3%	68.0%
	(22.1)	(21.4)	(20.7)	(16.4)
Destruction rate	12.0%	9.0%	15.6%	18.5%
	(28.9)	(26.2)	(36.1)	(39.6)
Expected destruction rate	14.0%	9.9%	23.9%	13.3%
	(25.7)	(19.7)	(33.7)	(24.3)
Frequency of positive destruction	21.8%	14.5%	18.5%	18.5%
Expected frequency of positive destruction	36.4%	32.7%	40.7%	37.0%
Observations	55	55	27	27

Note: Means or frequency for each variable. Standard deviations are in parenthesis.

4.1 Proposer behavior

Proposers in this experiment behave similarly to proposers in other power-to-take experiments. The average take rate (over both periods) in Treatment E is 56.4% and the median take rate is 60.0%. Similarly, in Treatment No-E, the average take rate is 64.7% and the median take rate is 66.5%.[13] In other words, in spite of the fact that both proposers and responders start with an endowment of 10 euros, proposers claim slightly more than half of the responders' endowment. The distribution of take rates for both treatments can be seen in the online appendix (Figure A1). These take rates are very similar to take rates in one-shot power-to-take games. For example, Bosman et al. (2005) find a mean take rate of 60.0%.

Table 1 shows additional descriptive statistics divided by period for both treatments. In each period, between 14.5% and 21.8% of the responders destroy a positive percentage of their endowment (i.e., $d > 0$). Average destruction rates per period range between 9.0% and 18.5% of the responders' endowment. Conditional

[13] As in previous studies, we do not find that gender or age have an effect on the chosen take rate. Consistent with some experiments (e.g., Carter and Irons, 1991), we find that economics students take significantly more than non-economists (Wilcoxon-Mann-Whitney test using individual averages across periods as independent observations, $p = 0.014$, $n = 82$)

TABLE 2 – CHANGE IN TAKE RATES FROM PERIOD 1 TO PERIOD 2

	Mean (std. dev.)	% decreasing	% constant	% increasing	Obs.
		Treatment measuring emotions			
Overall	5.2% (12.0)	12.7%	38.2%	49.1%	55
No destruction	7.0% (12.3)	7.0%	37.2%	55.8%	43
Positive destruction	–1.3% (8.3)	33.3%	41.7%	25.0%	12
		Treatment not measuring emotions			
Overall	6.7% (15.7)	18.5%	29.6%	51.9%	27
No destruction	9.8% (15.3)	9.1%	31.8%	59.1%	22
Positive destruction	–7.0% (9.7)	60.0%	20.0%	20.0%	5

on positive destruction, the average destruction rate is 57.8% in Treatment E and 92.0% in Treatment No-E. As in other experiments, there is a positive and significant correlation between the take rate and the destruction rate: the Spearman's rank correlation coefficient between the two is $\rho = 0.26$ ($p = 0.007$, $n = 110$) in Treatment E and $\rho = 0.52$ ($p < 0.001$, $n = 54$) in Treatment No-E.

Next, we take a look at how proposers adjust their take rate. The mean change in take rates between period 1 and period 2, as well as the fraction of proposers that increase, decrease, or keep constant their take rate is seen in Table 2. The same table displays these statistics depending on the action of the responder in period 1: no destruction or positive destruction. The distribution of the change in take rates for both treatments can be seen in the online appendix (Figure A2).

On average, proposers increase their take rate by 5.16 percentage points in Treatment E and by 6.70 percentage points in Treatment No-E. The increase in each treatment is significant with a Wilcoxon matched-pairs signed-rank test (WSR, p = 0.001, $n = 55$, and $p = 0.036$, $n = 27$, respectively).

As one would expect, the proposers' decision to change the take rate is affected by the behavior of the responder in the first period. The average take rate of proposers who do not experience destruction increases significantly from 50.2% to 57.2% in Treatment E (WSR test, $p < 0.001$, $n = 43$) and from 58.2% to 68.0% in Treatment No-E (WSR test, $p = 0.003$, $n = 22$). In contrast, the average take rate of proposers who experience positive destruction does not statistically change in either treatment: it goes from 66.6% to 65.3% in Treatment E (WSR test, p = 0.623, n = 12) and from 75.0% to 68.0% in Treatment No-E (WSR test, p = 0.172, n = 5).[14]

We summarize these findings as our first result.

RESULT 1: On average, proposers claim slightly more than half of the responders' endowment, and responders react with higher destruction rates the higher the proposers' claim is. Moreover, if proposers experience no destruction in period 1, they increase their take rate in period 2, whereas if they experience positive destruction in period 1, proposers keep or lower their take rate in period 2.

4.2 Emotion elicitation and behavior

In this subsection, we investigate whether eliciting emotions through self-reports affects the subsequent behavior of proposers. In order to do this, we look for differences in behavior between Treatment E and Treatment No-E. A quick look at Table 1 might suggest that take rates and expected destruction rates are higher in Treatment No-E. To test this impression, as well as other potential differences between treatments, we compare the distributions for each period of each statistic in Table 1 with both a Kolmogorov-Smirnov test (KS) and a Wilcoxon-Mann-Whitney test (WMW) or in the case of binary variables with a χ^2 test (all tests are two-sided).

We do not find differences at conventional levels of statistical significance. Comparing take rates results in the following p-values: p = 0.188 for period 1 and p = 0.334 for period 2 using KS tests, and p = 0.068 for period 1 and p = 0.080 for period 2 using WMW tests. For the remaining variables, p-values are even higher: for destruction rates $p \geq 0.866$ (KS tests) and $p \geq 0.514$ (WMW tests); for expected destruction rates $p \geq 0.370$ (KS tests) and $p \geq 0.365$ (WMW tests); for the frequency of destruction $p \geq 0.643$ (χ^2 tests); and for the expected frequency of destruction $p \geq 0.699$ (χ^2 tests).[15]

[14] This lack of statistical significance might be due to the small number of observations, in particular in Treatment No-E.

[15] For all these tests the sample size n equals 82.

One might worry that in the case of take rates, the *p*-values of the WMW tests are close to the standard 0.05 threshold. However, the fact that the *p*-value is already low in the first period indicates that this is not due to the measurement of emotions, which is done *after* proposers choose the take rate in the first period. On closer inspection, the potential difference in take rates appears to be driven by a difference in the subject pool between the two treatments (after all they were run three years apart). Namely, whereas 40% of proposers are students of economics in Treatment E, 74% are students of economics in Treatment No-E. To evaluate the effect of economists, we run two OLS regressions with the take rate as the dependent variable. In the first regression, we use a treatment dummy variable as the only independent variable (1 equals Treatment No-E). The coefficient and *p*-value for the treatment variable are: $b_1 = 8.3$ and $p = 0.056$. In the second regression, we add a dummy variable indicating whether a subject studies economics as a second independent variable. Consistent with the seeming treatment difference being driven by the difference in the subject pool, the coefficient for the economist variable equals $b_2 = 9.6$ and is statistically significant ($p = 0.041$) and the coefficient of the treatment variable decreases to $b_1 = 5.0$ and its *p*-value increases to $p = 0.275$.[16]

Next, we look for treatment differences in the way proposers adjust their take rate. After all, if there is an effect of measuring emotions on the proposers' decisions, it ought to show up between period 1 and period 2. As we saw, on average, proposers tend to increase their take rate in both treatments. However, if we test whether the change in the take rates differs between the two treatments, we do not find a statistically significant difference (KS test, $p = 1.000$; WMW test, $p = 0.940$; $n = 82$).[17] Moreover, we do not find a significant difference in the change in take takes when we test separately proposers who experienced no destruction (KS

[16] The regressions are done with robust standard errors and clustering on subjects. These results are robust to running the regressions for each period separately, including demographic characteristics (gender and age), and including an interaction term between the variables for treatment and area of studies.

[17] We also do not find a significant difference in the frequency of proposers that increase, decrease, or keep the same take rate across treatments (χ^2 test, $p = 0.665$, $n = 82$).

test, p = 1.000; WMW test, p = 0.765; n = 65) and proposers who experienced positive destruction (KS test, p = 0.328; WMW test, p = 0.304; n = 17). Lastly, we check whether these results are sustained once we control for the area of studies. We do so by running an OLS regression with the change in the take rate as the dependent variable. As independent variables we add: the responder's destruction rate in period 1, a treatment dummy variable, a dummy variable indicating whether a subject studies economics, and an interaction terms for each dummy variable and the destruction rate. The only statistically significant coefficients are the constant (b_0 = 7.7, p = 0.001) and the coefficient of the destruction rate (b_1 = –10.3, p = 0.005). In other words, the change in the take rate is positive unless there is enough destruction (to be precise, at least 75%) and neither economists nor proposers in Treatment No-E react significantly different to the responder's action ($p \geq 0.243$).[18]

Our second result sums up these findings.

RESULT 2: By and large, the measurement of emotions through self reports does not seem to affect the behavior of proposers or their reaction to the responders' destruction rate.

4.2 Emotions and proposer behavior

As we have seen, proposers who face a responder who does not destroy increase their take rates whereas proposers who face a responder who does destroy keep their take rate more or less constant. However, the experience of destruction explains only a part of the proposers' decision to change the take rate. Not all proposers who experience zero destruction increase their take rate: in Treatment E (No-E) only 55.8% (59.1%) do. Similarly, in spite of observing positive destruction, 25% (20%) of proposers nevertheless increase their take rate.

The next step in our analysis is to try to explain why, when faced with a similar situation, some proposers decide to change their take rate and some do not. In order to do so, we take a look at the subjects' emotional response. Descriptive

[18] Once again, the regression is done with robust standard errors and clustering on subjects.

TABLE 3 – EMOTIONAL REACTION OF PROPOSERS DEPENDING ON DESTRUCTION

	Responder did not destroy		Responder destroyed a positive amount	
Emotion	Period 1	Period 2	Period 1	Period 2
admiration	3.07 (1.92)	3.36 (2.08)	2.42 (1.44)	1.50 (1.07)
anger	1.09 (0.48)	1.26 (0.74)	3.67 (2.46)	3.25 (1.58)
contempt	1.35 (0.90)	1.36 (0.92)	2.50 (1.73)	2.75 (1.75)
disappointment	1.33 (0.89)	1.32 (0.84)	4.33 (2.10)	4.25 (2.38)
envy	1.23 (0.68)	1.34 (0.92)	3.00 (1.86)	2.00 (1.60)
gratitude	4.67 (1.86)	4.94 (1.69)	2.50 (1.45)	2.38 (1.92)
guilt	2.07 (1.35)	2.45 (1.56)	2.00 (1.35)	1.75 (1.49)
irritation	1.40 (1.28)	1.28 (0.83)	4.17 (1.85)	3.88 (2.03)
joy	5.02 (1.42)	4.96 (1.82)	2.83 (1.70)	3.25 (2.55)
pride	3.91 (1.84)	4.15 (2.13)	2.75 (1.14)	1.50 (1.07)
regret	2.05 (1.36)	2.04 (1.53)	3.25 (1.48)	2.88 (2.17)
sadness	1.19 (0.79)	1.17 (0.56)	2.58 (1.78)	2.00 (1.77)
shame	1.86 (1.30)	2.26 (1.45)	2.42 (1.56)	1.50 (1.41)
surprise	3.00 (1.89)	3.26 (1.70)	4.42 (2.19)	4.25 (2.31)

Note: Mean emotional intensities. Standard deviations are in parenthesis.

statistics of the proposers' emotional response depending on whether or not they experienced destruction in period 1 are available in Table 3. Furthermore, the correlation coefficients between the different emotions are available in the online appendix (Table A2).

It is well-known that in spite of the fact that different emotions have distinct eliciting conditions and action tendencies, in many cases similar emotions will be experienced simultaneously (Haidt, 2003). For this reason, emotion researchers often divide emotions into different categories or families. Unfortunately, since emotions that are normally correlated in some contexts can be uncorrelated in others (Zeelenberg and Pieters, 2006), it is hard to define *ex ante* the appropriate categorization.

To find the families of emotions that best describe our data, we analyze the proposers' emotional response using principal component analysis. This allows us to group emotions according to the correlations observed in the experiment. Given that we are interested in how emotions affect the adjustment of take rates, we

TABLE 4 – FACTORS SUMMARIZING THE PROPOSERS' EMOTIONAL RESPONSE

Factor	1st	2nd	3rd	4th	5th
Proportion of variance explained	0.36	0.13	0.11	0.11	0.07
			Factor loadings		
anger	0.42	0.02	–0.01	0.03	–0.05
disappointment	0.41	–0.01	0.00	0.04	0.01
envy	0.40	0.01	0.09	0.01	0.02
sadness	0.33	0.27	–0.02	–0.01	–0.21
irritation	0.37	–0.07	0.00	–0.03	0.08
contempt	0.39	–0.06	0.20	0.03	0.07
regret	0.00	0.03	0.00	0.00	0.96
shame	–0.03	0.66	–0.03	–0.01	0.08
guilt	0.02	0.66	–0.01	–0.01	0.01
pride	0.06	–0.03	0.78	–0.06	0.01
joy	–0.16	0.03	0.47	0.10	0.02
gratitude	–0.17	0.03	0.19	0.46	0.01
admiration	–0.07	0.13	0.10	0.60	–0.11
surprise	0.15	–0.14	–0.26	0.63	0.09
			KMO test = 0.82		
Family of emotions	Negative	Prosocial	Positive	Praising	Regret
Mean	0.04	0.12	–0.01	0.04	0.04
Standard deviation	2.30	1.33	1.26	1.24	0.98

Note: Principal component analysis of the proposers' emotional response in period 1. Reported factors are rotated with orthogonal varimax rotation and account for 80% of the variance.

concentrate on the emotional response in period 1 and in order to have as many observations as possible for the estimation we utilize the emotional response of all proposers. However, if we do the analysis with the emotional response in period 2 or if we restrict ourselves to proposers who kept their role we obtain very similar factors (available Tables A3 and A4 in the online appendix). Moreover, the results reported in the paper remain qualitatively similar if we use these alternative factors.

We find that five factors do a good job at summarizing the proposers' emotional response. These factors are presented in Table 4. The first factor consists of the emotions of anger, disappointment, envy, sadness, irritation, and contempt. We refer to this factor as being the family of *negative* emotions (since it picks up almost all of them). The second factor clearly represents the emotions of shame and guilt. Thus, following Bowles and Gintis (2005), we call this the family of *prosocial* emotions. The third factor seems to be a combination of pride and joy, which for lack of a better name we call the family of *positive* emotions. The fourth factor picks up gratitude, admiration, and surprise. We refer to this factor as the family of *praising* emotions. Lastly, the fifth factor is basically the emotion of *regret*. At the bottom of Table 4 we provide the descriptive statistics of each factor.

It is worth noting that the emotion families that emerge from the principal component analysis are quite similar to the families of moral emotions proposed by Rozin et al. (1999) and Haidt (2003). Most of the emotions in the first factor fall in the family they call "other-condemning" (the exceptions are sadness and envy). The second factor corresponds to the "self-conscious" family and the fourth factor to the "other-praising" family. They do not discuss the emotions in the other two factors.

Next, we test whether emotions explain how proposers change their take rate from period 1 to period 2 controlling for the effect of the responders' destruction. [19] For our dependent variable we use a ternary variable that indicates whether proposers increased (1), kept constant (0), or decreased (–1) their take rate. As independent variables we use a binary variable indicating whether the responder destroys (1) or not (0), and two interaction variables per family of emotions: the first interaction variable equals the value of the respective factor if there is no destruction and zero otherwise, and the second interaction variable equals the value of the

[19] Expected destruction rates are also a potential explanation of why some of these proposers increase their take rates while others do not. It is reasonable to suppose that proposers who expect positive destruction and observe zero destruction are more likely to increase their take rate than proposers who correctly anticipate no destruction. However, WMW and KS tests do not find a significant difference between these two groups ($p = 0.371$ and $p = 0.696$; $n = 43$). Similarly, one might expect that the reaction of proposers who expect no destruction and face positive destruction is different from those who expect and observe positive destruction. However, once again, this is not the case (WMW test, $p = 0.290$; KS test, $p = 0.513$; $n = 12$). Moreover, the results from the regressions reported in Table 5 are not affected by the inclusion of the expected destruction rate, whose coefficient does not have a statistically significant coefficient. This might suggest that expectations do not have a strong impact on the proposers' decisions. However, since we only know the expected destruction rate conditional on the take rate, it is premature to conclude that the proposers' expectations do not play a role.

factor if there is positive destruction and zero otherwise. Lastly, given its initial effect on take rates, we also control for the area of studies. We estimate ordered probit coefficients with robust standard errors. The results are presented in Table 5. Columns I to V correspond to a separate regression for each family of emotions and column VI is a regression that includes all the families of emotions.

As we can see, when included separately, the only families of emotions that have a significant effect on the change in take rates are prosocial emotions and regret. Specifically, among proposers who experienced no destruction, proposers who report higher intensities of regret are significantly more likely to increase their take rate (p = 0.044), and among proposers who experienced positive destruction, those who report higher intensities of prosocial emotions are significantly more likely to decrease their take rate (p = 0.013). Moreover, these two findings are not affected when we run the regression using all the families of emotions (see column VI). We summarize these findings as our next two results.

RESULT 3: Proposers who increase their take rate after experiencing zero destruction are proposers that report higher intensities of regret.

Result 3 is quite intuitive. As one would expect, if a proposer reports feeling regret after observing zero destruction, it is because she realizes that she could have chosen a higher take rate.[20] What is more interesting is that feeling regret does not seem to be related to the proposer's expectations. One could think that proposers that report high intensities of regret are proposers that expected responders will destroy and then experience no destruction. However, we do not find evidence to support this hypothesis. For example, Spearman's rank correlation coefficient between regret and the expected destruction rate for these proposers is not statistically significant (ρ = 0.18, p = 0.248, n = 43). A possible explanation for this result is that some proposers do not want to take more than the amount they are

[20] This role of regret is, in some sense, at odds with reinforcement learning, which implies that proposers who do not experience destruction keep choosing the same take rate and those who experience destruction decrease their take rate (Roth and Erev, 1995). It is more in line with learning direction theory (Selten and Stoecker, 1986) and theories that integrate regret into best-reply functions (Engelbrecht-Wiggans and Katok, 2008).

TABLE 5 – CHANGE IN THE TAKE RATE DEPENDING ON DESTRUCTION AND EMOTIONS

Variable	I	II	III	IV	V	VI
Positive destruction	–1.01	–0.84*	–1.12*	–0.87*	–1.24**	–0.63
	(0.61)	(0.40)	(0.49)	(0.40)	(0.40)	(0.63)
Negative emotions	0.01					–0.05
and no destruction	(0.14)					(0.18)
Negative emotions	0.03					–0.18
and positive destruction	(0.11)					(0.18)
Prosocial emotions		0.09				0.03
and no destruction		(0.12)				(0.15)
Prosocial emotions		–0.47*				–0.73**
and positive destruction		(0.19)				(0.28)
Positive emotions			–0.08			–0.08
and no destruction			(0.14)			(0.17)
Positive emotions			–0.17			–0.36
and positive destruction			(0.26)			(0.42)
Praising emotions				0.15		0.25
and no destruction				(0.12)		(0.14)
Praising emotions				0.47		1.27
and positive destruction				(0.56)		(0.71)
Regret					0.48*	0.57*
and no destruction					(0.24)	(0.26)
Regret					0.30	–0.08
and positive destruction					(0.22)	(0.34)
Economist	0.02	0.09	0.07	0.01	0.11	0.21
	(0.35)	(0.34)	(0.35)	(0.35)	(0.33)	(0.39)
Log pseudolikelihood	–50.79	–48.69	–50.56	–49.85	–48.06	–43.55
Wald χ^2	5.92	13.67**	6.00	7.63	13.50**	32.93**

Note: The dependent variable is a ternary variable indicating whether a proposer increases (1), keeps (0), or decreases (–1) the take rate. Ordered probit estimates and robust standard errors in parenthesis. In all cases n = 55. * and ** indicate significance that the 5% and 1%.

already taking, and thus, they do not feel regret even though they realize they could have chosen a higher take rate. This is consistent with the fact that proposers who keep their take rates constant do lower their expected destruction rate (from 14.0% to 1.88%, WSR test, p = 0.026, n = 16), which shows that they understand they initially overestimated the risk of destruction.

RESULT 4: Proposers who decrease their take rate after experiencing positive destruction are proposers that report higher intensities of prosocial emotions (shame and guilt).

Result 4 gives us an important hint of what motivates proposers to decrease their take rate. The emotions of shame and guilt are generally triggered when an individual violates a norm, even in anonymous settings such as in this experiment (see Tangney et al., 1996). If destruction by responders makes proposers feel bad by triggering these emotions, one would expect proposers to lower their take rates in order to feel better. Naturally, this opens up the question: why do some proposers feel shame and guilt while others do not? As these emotions are triggered by the violation of a norm, this seems related to different beliefs regarding the fair amount to take. We explore the effect of fairness perceptions and prosocial emotions in the following subsection.

4.3 Emotions and fairness perceptions

A casual look at the data reveals that prosocial emotions are not triggered by high take rates. The correlation between the take rate and the factor of prosocial emotions is close to zero (Spearman's $\rho = 0.105$, $p = 0.444$, $n = 55$). Thus, if prosocial emotions are indeed triggered by deviations from a fairness norm or value, it appears that not all proposers consider that choosing a high take rate is a violation of such a norm/value. On closer inspection, this seems to be the case since we get a clear result if we take into account the proposers' perceived fair take rate t^f.

If we divide the proposers into those who choose a take rate that they consider unfair ($t > t^f$) and those who choose a take rate that they consider (more than) fair ($t \leq t^f$), we find that proposers who choose an unfair take rate (61.8% of the sample) report higher intensities of prosocial emotions than proposers who choose a fair take rate (KS test, $p = 0.001$; WMW test, $p = 0.001$; $n = 55$).

Moreover, this result is unaffected when we control for other variables. Specifically, we run an OLS regression with the factor of prosocial emotions as the dependent variable. As independent variables we use: the dummy variable indicating whether proposers think their take rate is fair (1) or unfair (0), an interaction variable that equals the destruction rate if the proposer thinks his take rate is fair and zero otherwise, a second interaction variable that equals the destruction rate if the proposer thinks his take rate is unfair and zero otherwise, and a dummy variable indicating the area of studies.[21] Figure 1 illustrates the regression's output. It shows the predicted value (and 95% confidence intervals) of the factor of prosocial emotions depending on the destruction rate and separating proposers who think their take rate was fair from those who think their take rate was unfair. As can be seen, proposers who believe they made a fair choice experience low intensities of prosocial emotions (even when they face destruction), and proposers who acknowledge they made an unfair choice experience higher intensities of prosocial emotions. This is confirmed by the coefficient and p-value for the fairness dummy variable, which are $b_1 = -0.91$ and $p = 0.017$, and by the fact that the coefficient of the destruction rate is smaller for proposers who think they are being fair (0.10 vs. 1.40). This brings us to our last observation, which is that the destruction rate has a positive effect on the proposers' experienced prosocial emotions. Thus, destruction not only imposes a monetary cost on proposers but it also imposes an emotional cost. What is interesting is that the coefficient of the destruction rate is significantly positive only among proposers who think their take rate is unfair ($p < 0.001$) and it is not significantly different from zero among proposers who think their take rate is fair ($p = 0.906$). We summarize these findings in our fifth result.

RESULT 5: Proposers who chose take rates that they consider unfair experience higher intensities of prosocial emotions—in particular, after observing a high destruction rate.

Note that the overall influence of prosocial emotions is likely to be bigger than the one we observe. After all, proposers who correctly anticipate experiencing high intensities of prosocial emotions will avoid feeling them by not taking more than

[21] The regression equation including robust standard errors (between parenthesis) is: 0.46 (0.31) – 0.91 (0.37) × Take rate is fair + 1.40 (0.38) × Destruction rate when take rate is unfair + 0.10 (0.86) × Destruction rate when take rate is fair – 0.17 (0.30) × Economist.

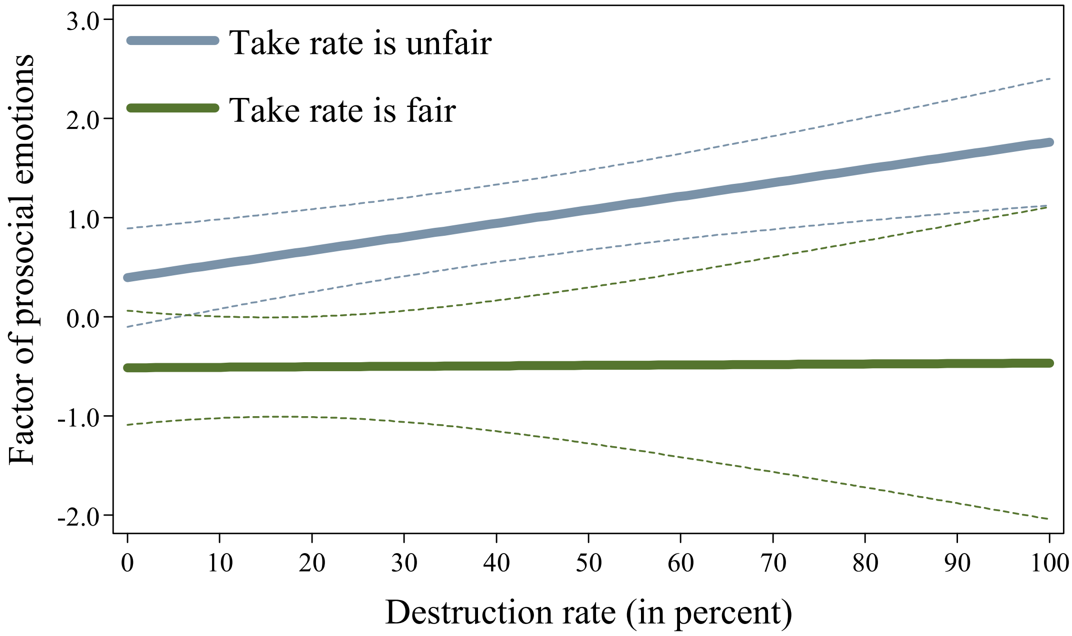

FIGURE 1 –PREDICTED INTENSITY OF PROSOCIAL EMOTIONS DEPENDING ON DESTRUCTION

Note: Predicted value of the factor of prosocial emotions depending on the destruction rate and the proposer's belief that his take rate was unfair or fair. An unfair (fair) take rate is a take rate that is above (below or equal to) the proposer's fair take rate. OLS estimates and 95% confidence intervals (calculated with robust standard errors). There are 55 observations, 34 with unfair rate rates and 21 with fair take rates.

their fair take rate. Thus, in addition to the observed *ex post* adjustment in take rates, prosocial emotions can be expected to have an effect by deterring the choice of high take rates in the first place. Put differently, proposers who choose an unfair take rate must expect only moderate intensities of prosocial emotions as they are willing to risk experiencing negative prosocial emotions in order to obtain a higher expected monetary payoff.

We think that the combination of Results 4 and 5 give us an important insight of one of the mechanisms through which fairness perceptions affect proposer behavior. We reiterate it as the following corollary.

COROLLARY: Fairness perceptions, by triggering prosocial emotions, have an impact on proposer behavior.

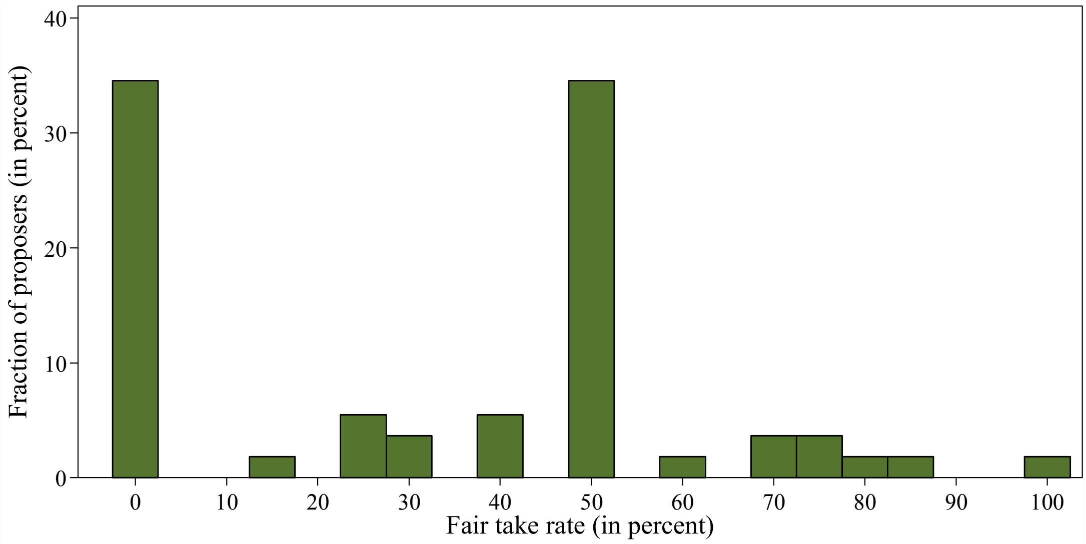

FIGURE 2 – HISTOGRAM OF THE PROPOSERS' FAIR TAKE RATE IN TREATMENT E

4.4 Heterogeneity in fairness perceptions

In this subsection we analyze the reported fair take rates and whether their values depend on the moment we chose to elicit them. Given that we find a link (through prosocial emotions) between fairness perceptions and behavior, it is worth noting that the take rate that is regarded as fair varies considerably. This variation is clearly seen in Figure 2. Roughly one third of proposers think that the fair action is not to take at all, another third thinks taking 50% of the responder's endowment is fair, and the rest is more or less evenly distributed among the remaining take rates. Note that this result is not limited to proposers. We find a very similar distribution of fair take rates for subjects playing twice as a responder. The heterogeneity in the reported fair take rates is also consistent with the findings from video experiments that show that individuals often disagree on what is the fair action to take in numerous types of games (Hennig-Schmidt, 2002). This leads us to out sixth result.

RESULT 6: There is little agreement on what the fair take rate is.

Interestingly, it is not necessarily true that proposers who act fairly (in their opinion) are being nicer to responders. The main difference between proposers who choose a take rate they think is fair and proposers who choose a take rate they think is unfair is their fairness perception and not their chosen take rate. For example, in the first period, proposers who think they are unfair choose a take rate that is 6.6

percentage points higher than proposers who think they are fair, but at the same time, they report a fair take rate that is 32.9 percentage points lower than proposers who thought they are fair.

Result 6 is important because it illustrates a missing characteristic in most models of fairness. Namely, that there can be disagreement as to what is fair. In other words, in addition to there being differences in the degree to which individuals care about fairness (e.g., the value of β and α in the model of Fehr and Schmidt, 1999), there can also be disagreement with respect to what the fairness reference point is (e.g., should income that is not at stake be used for interpersonal comparisons).[22]

Given that we elicit fair take rates at the end of the experiment, it is reasonable to worry about the causal assertions of the previous subsection. In particular, instead of fair take rates affecting behavior through prosocial emotions, it is possible that behavior in the game is the one affecting the reported fair take rates. As the last part of our analysis, we perform a series of tests to ensure that the proposers' fairness perceptions are not affected by the events in the game.[23]

First, we make use of the observations of subjects that were randomly assigned to switch roles in Treatment No-E. Recall that these subjects were asked to report their fair take rate *before* they played the game (but after they knew their role for the first period), and therefore, they serve as a control to see whether fair take rates elicited after playing the power-to-take game are different from those elicited before. In Figure 3, we plot the histograms of fair take rates reported by proposers in Treatment No-E. The left panel corresponds to the 27 proposers who kept their role and reported their fair take rate after the game (like in Treatment E). The right panel

[22] In this view, fairness reference points are similar to expectations in models using psychological game theory. For example, in principle, in models of guilt aversion behavioral differences between individuals can be due to differences in the guilt parameter and/or to differences in the expectation they use as a reference point.

[23] We do acknowledge that we did not elicit fair take rates twice, once before and once after the game, and therefore we cannot do a within-subjects test of the effect of the game on fair take rates. The reason that we did not do this is that we wanted to avoid priming fairness concerns and/or anchoring by eliciting fair take rates before proposers chose their take rate.

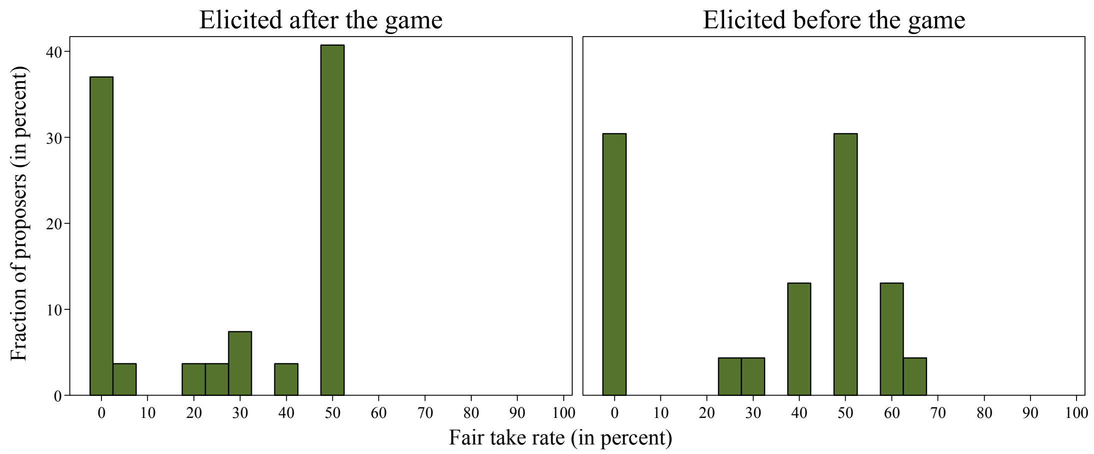

FIGURE 3 – HISTOGRAMS OF THE PROPOSERS' FAIR TAKE RATE IN TREATMENT NO-E

corresponds to the 23 proposers who later switched roles and reported their fair take rate before playing the game. As can be seen, the two distributions are very similar to each other and also to the distribution of fair take rates seen in Figure 2. We confirm this visual impression with WMW and KS tests. We do not find that the distribution of fair take rates elicited before the game significantly differs from the distribution of fair take rates elicited after the game in Treatment No-E (WMW test, $p = 0.254$, KS test, $p = 0.764$; $n = 50$) or in Treatment E (WMW test, $p = 0.829$, KS test, $p = 0.915$; $n = 78$). Moreover, this result does not change if we include the fair take rates reported by responders, which are very similar to those of proposers in both treatments.[24] In this case, the p-values are $p = 0.346$ (WMW test, $n = 100$) and $p = 0.923$ (KS test) when comparing within Treatment No-E and $p = 0.432$ (WMW test, $n = 156$) and $p = 0.832$ (KS test) when comparing to Treatment E. Lastly, if we control for the area of studies, we still do not find a significant difference.[25]

Second, we take a look at the reported fair take rates and we evaluate whether, *ceteris paribus*, proposers that faced destruction report different fair take rates than

[24] Unlike Schmitt (2004), we do not find evidence that role matters. Subjects who played twice as proposers do not report significantly different fair take rates than subjects who played twice as responders in either Treatment E (on average 33.3% vs. 29.5%; WMW test, $p = 0.580$, KS test, $p = 0.965$; $n = 110$) or in Treatment No-E (on average 25.9% vs. 20.0%; WMW test, $p = 0.237$, KS test, $p = 0.526$; $n = 54$). The same is true for proposers and responders who reported their fair take rate before they played the game (on average 33.5% vs. 20.2%; WMW test, $p = 0.070$, KS test, $p = 0.124$; $n = 46$).

[25] Ordered probit estimates with robust standard errors: $p = 0.172$ when comparing within Treatment No-E and $p = 0.640$ when comparing to Treatment E.

those that do not. In particular, one could think that proposers that experience destruction are "taught" by the responder that it is unfair to take a high amount. To start, given that two values (0% and 50%) account for a majority of the fair take rates, we construct a dependent variable with the following values: 0 if t^f = 0%, 1 if 0% < t^f < 50%, 2 if t^f = 50%, and 3 if t^f > 50%, which we use to run ordered probit regressions with robust standard errors. In our first regression we use the following independent variables: the take rate in the first period, the destruction rate in the first period, and the area of studies.[26] As one would expect, the coefficient for the take rate is positive and statistically significant (p = 0.009). The coefficient for the destruction rate is not statistically significant (p = 0.166) and paradoxically it is positive, with goes against the "teaching" hypothesis described above. Lastly, the economist dummy variable is negative but is not statistically significant (p = 0.984). We also tried running the regression with an interaction term between the take rate and the destruction rate. Neither the destruction rate nor the interaction term are statistically significant (p = 0.392 and p = 0.714). Moreover, we cannot reject the hypothesis that both coefficients are equal to zero (F-test, p = 0.311).[27] As a final robustness check, we run a probit regression with the same independent variables and a binary dependent variable indicating whether the take rate in the first period was considered fair ($t \leq t^f$) or unfair ($t > t^f$) by the proposer (i.e., the variable used to arrive to Result 5). Again, neither the destruction rate nor the interaction term are statistically significant (p = 0.142 and p = 0.368), and an F-test for joint significance

[26] Result 5 suggests that if we include the take and destruction rates from the second period in the regression then we would an endogeneity problem. For this reason we opted for the regression with first-period variables. However, our results remain unchanged if we use the average take and destruction rates over both periods.

[27] The reported results are for regressions using the observations from Treatment E. If we include subjects from Treatment No-E then the p-values for the destruction rate and the interaction term are p = 0.610 and p = 0.720 and the p-value for the test of joint significance is p = 0.830. Alternatively, if we run the same regression using the actual fair take rate as the independent variable then the p-values for the destruction rate and the interaction term are p = 0.534 and p = 0.962 and the p-value for the test of joint significance is p = 0.107.

also fails to yield a significant result ($p = 0.101$). Given this evidence, we conclude that the elicited fair take rates have not been influenced by behavior in the game.

5. Conclusions

In this paper we investigate how proposers in the power-to-take game adjust their behavior depending on their interaction with responders, their fairness perceptions, and experienced emotions. In addition, we test whether eliciting emotions through self-reports affects the subjects' subsequent behavioral response. Our main results can be summarized by the following statements:

- Measurement of emotions through self-reports does not affect the behavior of proposers.
- The emotional experience of proposers helps explain their reaction to receiving or not receiving punishment. Proposers who are not punished and feel regret tend to increase their take rates, and proposers who are punished and experience prosocial emotions (shame and guilt) tend to decrease them.
- Fairness considerations appear to play an important role by influencing the emotional reaction of proposers. In particular, only proposers who consider their take rate was unfair feel higher intensities of prosocial emotions.
- Fairness perceptions vary considerably between proposers.

Our results show that prosocial emotions play a significant role in the proposers' decision-making process. Of particular importance is the observation that if responders destroy, they provoke higher intensities of prosocial emotions among some proposers. Therefore, destruction not only reduces the proposers' income but it also makes them feel bad, which makes each unit of punishment of opportunistic behavior more effective. This highlights the importance for responders to signal their displeasure as the mere existence of a notion of fairness might not be enough to restrain the proposers' behavior.

The fact that some proposers do not adjust their behavior after being punished combined with the finding that these proposers consider their take rate to be fair and do not feel high intensities of prosocial emotions, highlights the importance of punishment to be legitimate for it to be effective. This observation helps explain why punishment or the threat of punishment promotes cooperation in certain situations whereas in others it does not (see e.g., Houser et al., 2008; Reuben and Riedl, 2009).[28] In other words, punishment on its own is not enough to promote prosocial behavior. There must also be an understanding of what is fair and what is unfair.

Given the reported importance of fairness perceptions, it is interesting that they vary substantially among proposers. Considering that even philosophers find it hard to define fairness, it is not that surprising that proposers disagree on what the fair take rate is. The difficulty of defining fairness means that, even if proposers want to be fair, they first have to figure out what fairness means in a specific context. In the power-to-take game, about a third of the subjects settle on a definition of fairness that implies an equal split of the subjects' total income. In contrast, another third settles on a definition of fairness that splits equally only the money that is at stake (i.e., the endowment of the responder). Although the latter definition might seem quite unfair to some, it is consistent with Rabin's (1993) formalization of fairness and with some of the fairness notions described by Konow (2003).

The results of this paper are encouraging for the modeling of emotions in economics. The fact that experiencing prosocial emotions affects the proposers' subsequent behavior is consistent with models that incorporate the emotions of shame and guilt. Knowing that the prosocial emotions play a role in this type of setting can inform us on how to best use tools such as psychological game theory (Geanakoplos et al., 1989) to model fairness-influenced behavior.[29] For example, research on emotions tells us that people feel more shame in situations in which others can clearly observe their actions and show disapproval (Tangney and

---

[28] This might also be an explanation of why, when it is allowed, we see some individuals retaliate against those that are doing the punishing (see, Denant-Boemont et al., 2007; Nikiforakis, 2008). For a study specifically designed to investigate the role of emotions in "counter-punishment" see Hopfensitz and Reuben (2009).

[29] There are already a few attempts of formally modeling specific prosocial emotions. These include Bowles and Gintis (2005), Charness and Dufwenberg (2006), and Miettinen (2008).

Dearing, 2002). This would be consistent with proposers asking more in ultimatum games in which the amount to be divided is unknown. If uncertainty over the size of the pie prevents responders from clearly judging the actions of proposers, proposers might feel less shameful when making a low offer.

As mentioned in the previous section, prosocial emotions can have additional effects on behavior if people anticipate their own emotional response and act in order to avoid feeling bad. In this experiment, we cannot test whether this is true. However, recent work by Dana et al. (2006) and Lazear et al. (2006) provides support to this idea. If proposers make positive offers in dictator games for the sole reason of avoiding prosocial emotions then, if given a choice, they would prefer to avoid being in the dictator game in the first place. Lazear et al. (2006) study precisely this situation and find that proposers who give the most when forced to play a dictator game are also the most willing to avoid playing. This is exactly what a model that incorporates individuals who anticipate their own prosocial emotions would predict.

Another implication of this paper is that the motivations of proposers are actually quite different from those of responders. Whereas proposers seem to be influenced by fairness and emotions such as regret, shame, and guilt, responders are affected more by their expectations and emotions such as anger and irritation (van Winden, 2001). This suggests that although very low take rates and positive destruction rates are both deviations from profit maximization, they are the result of two very different decision-making processes.[30] In fact, in the experiment we find no relationship between the two. If we look at subjects who switched roles, subjects who destroy in the first period do not choose lower take rates in the second period (WMW test, p = 0.745). Similarly, subjects who choose above-average take rates in the first period do not destroy less in the second period (WMW test, p = 0.758). In other words, individuals that are willing to punish others for treating them badly

[30] In this context, see Loewenstein et al. (1989) for a discussion on the qualitative difference between the reactions of individuals depending on whether they have a positive of negative relationship with others.

are not necessarily willing to treat others nicely. More generally, these findings suggest that proposer and responder behavior should be modeled separately.

Finally, we wish to advocate that measuring the emotional reaction of subjects can help us understand what is motivating them to make certain decisions. In this case, we identify prosocial emotions as an important motivation for proposers to lower their take rates. We do not argue that it is always necessary to know the precise emotional and cognitive processes by which subjects arrive at a decision. However, whenever we have a situation in which theory does not provide us with good predictions, a better understanding of the motivations of individuals can give us insights to improve the predictive power of our models.[31]

References

Adolphs R. (2002). Recognizing emotion from facial expressions: Psychological and neurological mechanisms. *Behavioral and Cognitive Neuroscience Reviews* 1: 21-61.

Andreoni, J., M. Castillo, and R. Petrie (2003). What do bargainers' preferences look like? Experiments with a convex ultimatum game. *American Economic Review* 93: 672-685.

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: 314-323.

Binmore, K., P. Morgan, A. Shaked, and J. Sutton (1991). Do people exploit their bargaining power? An experimental study. *Games and Economic Behavior* 3: 295-322.

[31] Our experiment is a good example of the added predictive power of emotions. If one is interested in predicting how proposers adjust their take rate, then using the data on emotions one obtains a model with a substantially better fit: the pseudo R^2 of the regression in column VI of Table 5 equals 0.191, and is significantly higher than 0.057, which is the pseudo R^2 of a regression that does not include the emotions data (χ^2 test, p = 0.008; if we run the two regressions using OLS the corresponding R^2s are 0.323 and 0.115).

Bjork, J.M., B. Knutson, G.W. Fong, D.M. Caggiano, S.M. Bennett, and D.W. Hommer (2004). Incentive-Elicited Brain Activation in Adolescents: Similarities and Differences from Young Adults. *The Journal of Neuroscience* 24: 1793-1802.

Blanco, M., D. Engelmann, A.K. Koch, and H.-T. Normann (2008). Belief Elicitation in Experiments: Is there a Hedging Problem? Discussion Paper No. 3517, IZA.

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

Bosman, R., M. Sutter, and F. van Winden (2005). On the impact of real effort and emotions in power-to-take experiments. *Journal of Economic Psychology* 26: 407-429.

Bowles, S. and H. Gintis (2005). Prosocial emotions. In Blume, L.E. and S. Durlauf (Eds.) *The Economy as a Complex Evolving System III: Essays in Honor of Kenneth Arrow*. Oxford: Oxford University Press.

Bowman, K. (2009). Public Opinion on Taxes. AEI Studies in Public Opinion, American Enterprise Institute for Public Policy Research.

Bradley, M.M. and P.J. Lang (2000). Measuring emotion: behavior, feeling and physiology. In Lang, R.D. and L. Nadel (Eds.) *Cognitive Neuroscience of Emotions*. Oxford: Oxford University Press.

Brenner, T. and N.J. Vriend (2006). On the behavior of proposers in ultimatum games. *Journal of Economic Behavior & Organization* 61: 617-631.

Camerer, C. (2003). *Behavioral Game Theory*. New Jersey: Princeton University Press.

Camerer, C. and G. Loewenstein (1993). Information, fairness, and efficiency in bargaining. In Mellers B. and J. Baron (Eds.) *Psychological Perspectives on Justice: Theory and Applications*. Cambridge: Cambridge University Press.

Carter, J.R., and M.D. Irons (1991). Are economists different, and if so, why? *Journal of Economic Perspectives* 5: 171-177.

Charness, G. and M. Dufwenberg (2006). Promises and partnership. *Econometrica* 74: 1579-1601.

Dana, J., D.M. Cain, and R.M. Dawes (2006). What you don't know won't hurt me: Costly but quiet exit in dictator games. *Organizational Behavior and Human Decision Processes* 100: 193-201.

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.

Denant-Boemont, L., D. Masclet and C. Noussair (2007). Punishment, counterpunishment and sanction enforcement in a social dilemma experiment. *Economic Theory* 33: 145-167.

Ellingsen, T., M. Johannesson, S. Tjøtta, and G. Torsvik (2010). Testing guilt aversion. *Games and Economic Behavior* 68: 95-107.

Engelbrecht-Wiggans, R. and E. Katok (2008). Regret and Feedback Information in First-Price Sealed-Bid Auctions. *Management Science* 54: 808-819.

Fehr, E., U. Fischbacher, and S. Gächter (2002). Strong reciprocity, human cooperation, and the enforcement of social norms. *Human Nature* 13: 1-25.

Fehr, E. and K. Schmidt (2006). The economics of fairness, reciprocity and altruism: Experimental evidence and new theories. In Kolm, S.C. and J.M. Ythier (Eds.) *Handbook on the Economics of Giving, Reciprocity and Altruism.* Amsterdam: Elsevier.

Fischbacher U. (2007). z-Tree: Zurich Toolbox for ready-made economic experiments. *Experimental Economics* 10: 171-178.

Fischbacher, U., C.M. Fong, and E. Fehr (2009). Fairness, errors and the power of competition. *Journal of Economic Behavior & Organization* 72: 527-545.

Friedman, D. and D.W. Massaro (1998). Understanding Variability in Binary and Continuous Choice. *Psychonomic Bulletin and Review* 5: 370-389.

Forsythe, R., J.L. Horowitz, N.E. Savin, and M. Sefton (1994). Fairness in simple bargaining experiments. *Games and Economic Behavior* 6: 347-369.

Geanakoplos, J., D. Pearce, and E. Stacchetti (1989). Psychological games and sequential rationality. *Games and Economic Behavior* 1: 60-79.

Grosskopf, B. (2003). Reinforcement and Directional Learning in the Ultimatum Game with Responder Competition. *Experimental Economics* 6: 141-158.

Güth, W., C. Schmidt, M. Sutter (2003). Fairness in the mail and opportunism in the internet—A newspaper experiment on ultimatum bargaining. *German Economic Review* 4: 243-265.

Güth, W., C. Schmidt, M. Sutter (2007). Bargaining outside the lab—A newspaper experiment of a three-person ultimatum game. *The Economic Journal* 117: 449-469.

Güth, W., R. Schmittberger, and B. Schwarze (1982). An experimental analysis of ultimatum bargaining. *Journal of Economic Behavior & Organization* 3: 367-388.

Haidt, J. (2003). The moral emotions. In Davidson, R.J., K.R. Scherer, and H.H. Goldsmith (Eds.) *Handbook of Affective Sciences*. Oxford: Oxford University Press.

Hennig-Schmidt, H. (2002). The Impact of Fairness on Decision Making—An Analysis of Different Video Experiments. In Andersson, F. and H. Holm (Eds.) *Experimental Economics: Financial Markets, Auctions, and Decision Making*. Dordrecht: Kluwer.

Henrich, J., R. Boyd, S. Bowles, C. Camerer, E. Fehr, H. Gintis, and R. McElreath (2001). In search of homo economicus: Behavioral experiments in 15 small-scale societies. *American Economic Review* 91: 73-78.

Hoffman, E., K. McCabe, and V.L. Smith (1996). Social distance and other-regarding behavior in dictator games. *American Economic Review* 3: 653-660.

Hopfensitz, A. and R. Reuben (2009). The Importance of Emotions for the Effectiveness of Social Punishment. *Economic Journal* 119: 1534-1559.

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: 509-532.

Konow, J. (2003). Which is the fairest one of all? A positive analysis of justice theories. *Journal of Economic Literature* 41: 1186-1237.

Krupka, E.L. and R.A. Weber (2008). Identifying Social Norms Using Coordination Games: Why Does Dictator Game Sharing Vary? Discussion Paper No. 3860, IZA.

Lazear, E.P., U. Malmendier, and R.A. Weber (2006). Sorting in experiments. NBER Working paper 12041.

Lin, H. and S. Sunder (2002). Using experimental data to model bargaining behavior in ultimatum games. In Zwick R. and A. Rapoport (Eds.) *Experimental Business Research*. Dordrecht: Kluwer.

Loewenstein, G., L. Thompson, and M. Bazerman (1989). Social utility and decision making in interpersonal contexts. *Journal of Personality and Social Psychology* 57: 426-441.

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.

Miettinen, T. (2008). Contracts and Promises—An Approach to Pre-play Agreements. Working Paper No 707. SSH/EFI.

Nikiforakis, N.S. (2008). Punishment and Counter-punishment in Public Goods Games: Can we Still Govern Ourselves? *Journal of Public Economics* 92: 91-112.

Noussair, C. and S. Tucker (2005). Combining monetary and social sanctions to promote cooperation. *Economic Inquiry* 43: 649-660.

Offerman, T., J. Sonnemans, G. van de Kuilen, and P.P. Wakker (2009). A Truth-Serum for Non-Bayesians: Correcting Proper Scoring Rules for Risk Attitudes. *Review of Economic Studies* 76: 1461-1489.

Pillutla, M.M. and J.K. Murnighan (1996). Unfairness, anger, and spite: Emotional rejections of ultimatum offers. *Organizational Behavior and Human Decision Processes* 68: 208-224.

Rabin, M. (1993). Incorporating fairness into game theory and economics. *American Economic Review* 83: 1281-1302.

Rapoport, A., J.A. Sundali, and R.E. Potter (1996). Ultimatums in two-person bargaining with one-sided uncertainty: Offer games. International *Journal of Game Theory* 25: 475-494.

Reuben, E. (2006). *Fairness in the Lab: The effects of norm enforcement in economic decisions*. Amsterdam: Thela Thesis.

Reuben, E., P. Sapienza, and L. Zingales (2009). Is mistrust self-fulfilling? *Economics Letters* 104: 89-91.

Reuben, E. and A. Riedl (2009). Public goods provision and sanctioning in privileged groups. *Journal of Conflict Resolution* 53: 72-93.

Reuben, E. and F. van Winden (2008). Social ties and coordination on negative reciprocity: The role of affect. *Journal of Public Economics* 92: 34-53.

Robinson, M.D. and G.L. Clore (2002). Belief and feeling: Evidence for an accessibility model of emotional self-report. *Psychological Bulletin* 128: 934-960.

Roth, A.E. and I. Erev (1995) Learning in Extensive-Form Games: Experimental Data and Simple Dynamic Models in the Intermediate Term. *Games and Economic Behavior* 8: 164-212.

Rozin. P., L. Lowery, S. Imada, and J. Haidt (1999). The CAD triad hypothesis: A mapping between three moral emotions (contempt, anger. disgust) and three moral codes (community. autonomy. divinity). *Journal of Personality and Social Psychology* 76: 574-586.

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: 1755-1758.

Schmitt, P. (2004). On perceptions of fairness: The role of valuations, outside options, and information in ultimatum bargaining games. *Experimental Economics* 7: 49-73.

Selten, R. and R. Stoecker (1986). End behavior in sequences of finite Prisoner's Dilemma supergames A learning theory approach. *Journal of Economic Behavior & Organization* 7: 47-70.

Sonnemans, J., and T. Offerman (2001). Is the Quadratic Scoring Rule Really Incentive Compatible?" Working Paper, University of Amsterdam.

Spitzer, M., U. Fischbacher, B. Herrnberger, G. Grön, and E. Fehr (2007). The neural signature of social norm compliance. *Neuron* 56: 185-196.

Straub, P. and K. Murnighan (1995). An experimental investigation of ultimatum games: Information, fairness, expectations, and lowest acceptable offers. *Journal of Economic Behavior & Organization* 27: 345-364.

Takahashi, H., N. Yahata, M. Koeda, T. Matsuda, K. Asai, and Y. Okuboc (2004). Brain activation associated with evaluative processes of guilt and embarrassment: an fMRI study. *NeuroImage* 23: 967-74.

Tangney, J. P. and R.L. Dearing (2002). *Shame and Guilt*. New York: Guilford.

Tangney, J. P., R. Miller, L. Flicker, and D.H. Barlow (1996). Are shame, guilt and embarrassment distinct emotions? *Journal of Personality and Social Psychology* 70: 1256-1269.

van Winden, F. (2001). Emotional hazard exemplified by taxation-induced anger. *KYKLOS* 54: 491-506.

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

Zeelenberg, M. and Pieters, R. (2006). Feeling is for doing: a pragmatic approach to the study of emotions in economic behavior. In de Cremer, D., M. Zeelenberg, and K. Murnighan (Eds.). *Social Psychology and Economics*. New Jersey: Erlbaum.