Games **2012**, *3*, 78-96; doi:10.3390/g3020078

OPEN ACCESS

games

ISSN 2073-4336

www.mdpi.com/journal/games

Article

What Behaviors are Disapproved? Experimental Evidence from Five Dictator Games

Raúl López-Pérez [1,*] and Marc Vorsatz [2,3]

[1] Universidad Autónoma de Madrid, Facultad de Ciencias Económicas, Departamento de Análisis Económico, Cantoblanco, 28049 Madrid, Spain

[2] Departamento de Análisis Económico II, Universidad Nacional de Educación a Distancia (UNED), Paseo Senda del Rey 11, 28040 Madrid, Spain; E-Mail: mvorsatz@fedea.es

[3] Fundación de Estudios de Economía Aplicada (FEDEA), Calle Jorge Juan 46, 28001 Madrid, Spain

* Author to whom correspondence should be addressed; E-Mail: raul.lopez@uam.es; Tel.: +34-91-497-6801; Fax: +34-91-497-6930.

Received: 28 February 2012 / Accepted: 30 March 2012 / Published: 23 April 2012

Abstract: The literature on social norms has often stressed that social disapproval is crucial to foster compliance with norms and promote fair and cooperative behavior. With this in mind, we explore the disapproval of allocation decisions using experimental data from five dictator games with a feedback stage. Our data suggests that subjects are heterogeneous in their disapproval patterns, distinguishing two main groups: (1) Subjects who only disapprove choices that harm them, and (2) subjects who disapprove socially inefficient choices.

Keywords: approval; disapproval; dictator game; experiment; social norms

1. Introduction

Social scientists have often stressed that social disapproval can foster normative behavior and cooperation within groups—see [1] for numerous illustrations. Indeed, a growing experimental literature shows that the availability of approval/disapproval increases contributions in public goods games (see, for instance, [2–4]), cooperation in social dilemmas [5], or generosity in dictator games [6,7]. Given this evidence, it seems important to understand which kind of behavior tends to be disapproved, as this could clarify further how (and when) social disapproval affects cooperation and generous behavior.

Another, perhaps more speculative reason to study disapproval patterns has to do with social norms. In effect, much of the existing literature on social norms emphasizes that norms are partly sustained by the disapproval of others. [8], (p. 104), for instance, contends that "when people obey norms, they often want to avoid the disapproval—ranging from raised eyebrows to social ostracism— of other people." Similarly, [9] (p. 366) notes that "norms are also enforced by expressions of disapproval, by ridicule, and in extreme cases, by ostracism. [...] If I dress bizarrely or spit in public, I may find myself shunned" (see also [10] and references therein, [11,12]. All this suggests that social norms might be elicited by studying what people disapprove [13], an important issue if we believe that behavior is often shaped by social norms.

This paper uses experimentation to study the disapproval of allocation decisions. For this, participants in our study play five dictator games with an additional stage in which the recipients can evaluate the dictator's choice. We then apply the analysis of [14] in order to classify recipients according to their individual disapproval pattern. The results show that recipients are heterogeneous in this respect. In particular, we find that the best distributional model in terms of parsimony and empirical relevance distinguishes two types of recipients: those who disapprove of choices if they are harmed and those who disapprove of socially inefficient choices. Perhaps paradoxically, in our games equity considerations or payoff comparisons only play a relatively minor role. The data also provides insights into the content of social norms and more precisely on *distributive norms*, that is, norms regulating how to distribute income or other goods.

We proceed as follows. Section 2 describes and discusses our experimental design. In Section 3, we report some aggregate statistics and perform the classification analysis. We conclude in Section 4. The experimental instructions are relegated to the Appendix.

2. Experimental Design and Procedures

We consider five binary dictator games in which the recipient has the possibility to express her/his opinion about the dictator's choice. The games are presented one at a time and each consists of two stages. In the first stage, the dictator chooses between two payoff allocations, which differ across games. In order to discriminate between different disapproval patterns (see next section), we employed the following ECU allocations (20 ECU = 1 Euro), where the first number in each parenthesis indicates the dictator's payoff:

1. $(180, 140)$ and $(100, 220)$
2. $(180, 120)$ and $(100, 100)$
3. $(240, 120)$ and $(100, 120)$
4. $(180, 120)$ and $(120, 180)$
5. $(200, 120)$ and $(160, 160)$

In the second stage, the recipient can send a closed–form, costly message (for a 10 ECU fee) to the dictator. The message depends on the dictator's choice and must be chosen from the following: "Your choice was (1) good, (2) neither good nor bad, (3) bad." In what follows, we interpret the first message as

approval, the second as neutral, and the third as disapproval. One subtlety of our design is that subjects do not receive any feedback during the experiment about the choices they make of their match. For this reason, we elicited the recipient's messages by means of the strategy method. In each allocation of each dictator game, more precisely, the recipient is first asked whether she/he wants to pay the message fee. Afterwards, and *independently* of her/his prior answer, the recipient is asked which of the three available messages she/he would like to send had she/he paid the fee. In this manner, we elicited 10 messages from each recipient (one for each allocation). Of course, the message is delivered only if the recipient decided to pay the fee at the allocation chosen by the dictator.

Two of the above mentioned features of our study merit some discussion. *First*, we used the strategy method because it maximizes the amount of statistical evidence (a crucial point for our classification analysis) and because we did not want subjects to receive any feedback. This was important to avoid variations in the players' mood across games (*e.g.*, a recipient might get angry at the dictator's choice and that could affect her subsequent choices), and thus order effects which would complicate our classification analysis. While we believe that the use of multiple games and the strategy method is the most efficient manner to explore subjects' disapproval patterns, some caution is warranted when interpreting our results, as the hypothetical character of this method might have an effect on its own. For instance, emotions could be attenuated and that could in turn affect the disapproval (see [15] for a survey of experiments using the strategy and the direct response method). *Second*, our design allowed us to know whether a recipient disapproved a choice even if the fee had not been paid. In other words, we elicited messages both when they are *costly* but also in the hypothetical case in which they were *costless* (as we asked the recipients even if they had not paid the fee). Although this is not typical in the experimental literature on non–monetary sanctions, which usually focus on costless messages, we find it an important design feature because a comparison between costly and costless messages allows us to see if the latter have been chosen carelessly or at random.

The experiment was computerized and programmed using z–Tree [16]. In total, 180 subjects from the Faculty of Economics and Business Administration at Maastricht University participated in the experiment. Since this experiment is rather short, we implemented it together with another experiment on a prisoner's dilemma (PD) game, which consisted of three different treatments and whose results are detailed in [5].[1] More precisely, after the corresponding PD treatment was finished, subjects received instruction sheets explaining the general structure of the dictator games (information about the payoffs was only provided for the first of the five games), were assigned a fixed role (dictator or recipient), and were randomly re-matched with another participant for the five games. All this was publicly announced. We decided not to re-match subjects after each dictator game because no repeated game effects were possible due to the absence of feedback; in fact, from a theoretical point of view, each game can be analyzed as one-shot. To ensure that the PD did not affect behavior in the dictator games, subjects were not informed about the results of the PD.[2] In order to prevent income effects, only one game (among the PD and the five dictator games) was randomly selected for payment when all subjects had finished their

[1] The complete set of original instructions corresponding to the so-called *feedback* treatment can be found in the Appendix. Note that in the Appendix the PD game is called Scenario 1. As a result, the first dictator game is referred to as Scenario 2.

[2] As we noted before, the absence of feedback should attenuate order effects due to mood variations (subjects cannot condition their choices on the earlier behavior of others). However, we cannot totally exclude that some subjects condition their behavior in a given scenario on their *own* earlier behavior.

choices (subjects knew this from the beginning). Finally, subjects were informed about their co–player's actions in the payoff–relevant game. The average payoff of the 45 minutes session was about 9 Euros.

3. Results

This section is organized as follows. First, we briefly report on some aggregate data regarding *costless* messages (all messages, independently whether the fee was paid or not) and the dictators' behavior. Then, we focus on our main objective, which is the elicitation and classification of the disapproval patterns of the recipients using again costless messages. Finally, we comment on the costly messages (the subset of messages the recipients decided to pay for).

3.1. Aggregate Statistics

Table 1 presents the costless messages of the recipients.[3] For the left–hand allocation of game 1 (180/140 *vs.* 100/220), for instance, we see that 27.78 % of the recipients approve that choice, whereas 40 % disapprove it (the remaining recipients send a neutral message). We observe that in each game, one allocation receives mostly disapprovals and the alternative allocation mostly approvals. This suggests that recipients tend to follow some common rule.

Table 1. Approval [Disapproval] rates of the recipients in the five dictator games using costless messages.

	Game			Costless Messages	
	Left–Hand		Right–Hand	Left–Hand Allocation	Right–Hand Allocation
1	(180,140)	*vs.*	(100,220)	0.2778 [0.4000]	0.7333 [0.1667]
2	(180,120)	*vs.*	(100,100)	0.6889 [0.1333]	0.1556 [0.6444]
3	(240,120)	*vs.*	(100,120)	0.6222 [0.1667]	0.2000 [0.5444]
4	(180,120)	*vs.*	(120,180)	0.1778 [0.4556]	0.6222 [0.1667]
5	(200,120)	*vs.*	(160,160)	0.2000 [0.5444]	0.7333 [0.1222]

Although our goal in this paper is to explore the recipients' behavior, we can also make a brief reference to the actions of the dictators. In most games, they choose the allocation that maximizes their own payoff. More precisely, the frequency of that choice is 98.88 % in the first, 94.44 % in the second, 100 % in the third, 98.88 % in the fourth, and 81.11 % in the fifth game. One might be tempted to infer from this that dictators are selfish, but we note that in most of our games, many theories of other–regarding preferences make the same predictions as the standard model of selfish players.[4] In any

[3] The behavior of the recipients does not vary significantly depending on the PD treatment played before. For this reason, we decided to pool the data, irrespective of the PD treatment. See the Appendix for the corresponding statistical analysis.

[4] Models of inequity–aversion [17,18], for instance, share predictions with the standard model in all games except game 5. In the latter, models of inequity-aversion predict that some players should choose the egalitarian allocation (as almost 19 %

case, our games were not chosen to study the dictators' behavior, but that of the recipients'. We do this in detail in what follows.

3.2. Classification Analysis

We apply the classification procedure from [14] to analyze the motives behind the recipients' disapproval. More precisely, we posit that the recipients follow deterministic decision rules when disapproving (the rules may differ from subject to subject), but also that they deviate with probability $\varepsilon > 0$. This classification procedure has several favorable attributes. By selecting the decision rule that best fits each subject's behavior, we can classify subjects by types. It also helps us finding the best single decision rule or the combination of two, three, *etc.* decision rules that best account for the behavior in our dictator games. Given this, we can then apply the *Akaike Information Criterion* (AIC) to infer the number of decision rules necessary to provide an accurate but parsimonious explanation of the observed disapproval pattern. Importantly, this procedure circumvents the multicollinearity problems that would appear in a classical regression analysis if the decision rules were entered as independent variables and allows for appropriate inferences even when testing all possible decision rules —no matter how similar their predictions are—at the same time.

3.2.1. Individual Decision Rules

For simplicity, we restrict our analysis to binary decision rules; that is, rules that indicate for each allocation whether or not the subject expresses disapproval (without specifying the message sent if the subject does not disapprove). More precisely, any rule can be formally described as a vector of ten numbers, one for each allocation of each game. The rule takes value one if it predicts disapproval at the corresponding allocation, and value zero otherwise. We consider seven rules that seem especially appealing. Letting (x_1^l, x_2^l) refer to the left–hand and (x_1^r, x_2^r) to the right–hand allocation at any of our dictator games (1 denotes the dictator and 2 the recipient), the rules are defined as follows:

1. The *No–Disapproval rule* never predicts disapproval;
2. the *Envy rule* predicts disapproval at allocation $a \in \{l, r\}$ if, and only if, $x_1^a > x_2^a$ –that is, if the dictator gets a larger payoff than the recipient;
3. the *Harm rule* predicts disapproval at allocation $a \in \{l, r\}$ if, and only if, the alternative allocation $al \neq a$ is such that $x_2^a < x_2^{al}$ –that is, if the recipient was harmed by the choice of the dictator;
4. the *Spite rule* predicts disapproval at all allocations;
5. the *Efficiency rule* predicts disapproval at allocation $a \in \{l, r\}$ if, and only if, $x_1^a + x_2^a < x_1^{al} + x_2^{al}$, where $al \neq a$ –that is, subjects disapprove socially inefficient choices;
6. the *Equity rule* predicts disapproval at allocation $a \in \{l, r\}$ if, and only if, $|x_1^a - x_2^a| > |x_1^{al} - x_2^{al}|$, where $al \neq a$ –that is, subjects disapprove inequitable choices; and,

of our dictators did). Further, theories based on ideas like altruism [19]; also the "maximin" model of [20], reciprocity [21], or spite [22] predict the same choices as the standard model in all games.

7. the *Maximin rule* predicts disapproval at $a \in \{l, r\}$ if, and only if, $\min\{x_1^a, x_2^a\} < \min\{x_1^{al}, x_2^{al}\}$, where $al \neq a$ –that is, subjects disapprove deviations from a maximin or Rawlsian norm.

Table 2 presents the predictions of the Non–Trivial rules in the five dictator games.

Table 2. Predictions of disapproval in the five dictator games for the Non–Trivial Decision rules. The following notation is used: EN=Envy, HA=Harm, EQ=Equity, EF=Efficiency, MA=Maximin.

	Game			Predictions of Disapproval	
	Left–Hand		Right–Hand	Left–Hand Allocation	Right–Hand Allocation
1	(180,140)	*vs.*	(100,220)	EN, HA	EQ, MA
2	(180,120)	*vs.*	(100,100)	EN, EQ	HA, EF, MA
3	(240,120)	*vs.*	(100,120)	EN, EQ	EF, MA
4	(180,120)	*vs.*	(120,180)	EN, HA	
5	(200,120)	*vs.*	(160,160)	EN, HA, EQ, MA	

For instance, the Efficiency rule only predicts disapproval at the right–hand allocation of games 2 ($180/120 vs.\ 100/100$) and 3 ($240/120 vs.\ 100/120$). In these allocations, the social surplus is lower than in the corresponding alternative allocation (in contrast, games 1, 4, and 5 are surplus–preserving). As another example, the Maximin rule does not predict any disapproval in game 4 ($180/120 vs.\ 120/180$) because the minimum payoff is the same at both allocations. Note also that Table 2 motivates our selection of games: As the reader can confirm, the games allow us to discriminate between the above mentioned rules; that is, no two decision rules make the same predictions for all five games. For instance, game 5 ($200/120 vs.\ 160/160$) is useful to distinguish between Harm and Efficiency; the first rule predicts disapproval at the left–hand allocation, whereas Efficiency predicts no disapproval. As another example, game 2 ($180/120 vs.\ 100/100$) can be used to discriminate between rules like Efficiency and No-Disapproval.

To finish, we note that our games do not only permit disapproval behavior to be studied and whether subjects are heterogeneous in this respect. In effect, we can also infer something about social norms *if one assumes* that deviations from social norms are disapproved. More precisely, we are thinking here about *distributive norms*; that is, norms regulating how to distribute income or other goods. We can think of many examples: (i) Efficiency, or "choose the allocation that maximizes the social surplus", (ii) Equity, or "choose the most equitable allocation", (iii) Maximin, or "choose the allocation that maximizes the minimum payoff", (iv) No–Harm, or "choose an allocation so that the co–player's payoff is not reduced below a reference point" (this is inspired by [23] and the literature on reciprocity), and (v) No–Envy, or "choose any allocation that does not leave your co–player disadvantaged."[5] Most of our rules described before have a clear relation with one of these norms. For instance, a person who disapproves according

[5]The No-Envy norm is inspired by models of inequity aversion [17,18]. If *monetary* punishment was available (*i.e.*, punishment which reduces the material payoff of the punished person), these models would predict that a sufficiently

to the Envy rule might do so because she finds the No–Envy norm binding, while following the Harm rule might be a signal that the corresponding player places importance on the No–Harm norm. In turn, a person who does not care about norms should act according to the No–Disapproval rule.[6]

3.2.2. Methodology

Let D be the set of all recipients (of cardinality d) and let r_k be any set of k rules ($k = 1, ..., 7$) from the seven decision rules introduced above. Recipients are indexed by i, decision rules by j. Further, $x_i(j)$ denotes the number of allocations in the five dictator games where the rule j correctly predicts the choice (disapproving/not disapproving) of recipient i. Obviously, $x_i(j)$ cannot be larger than 10. Similarly, $x_i(r_k)$ is the number of allocations for which recipient i follows the rule(s) in r_k that best replicates her behavior; that is, $x_i(r_k) = \max_{j \in r_k} x_i(j)$. For example, if $r_2 = \{Harm, Envy\}$ and recipient i only disapproves the dictator at the left–hand allocation of game 1, then it follows from Table 2 that $x_i(\{Harm, Envy\}) = 7$ because the Harm rule correctly predicts her behavior for 7 allocations whereas the Envy rule predicts only for 6 allocations. Given any set r_k, let $f_{r_k} : D \to r_k$ be a mapping that assigns each recipient i to the decision rule j in r_k that replicates her/his behavior best—*i.e.*, $x_i(j) = x_i(r_k)$.[7] A *model* (r_k, f_{r_k}) consists of a set r_k and a mapping f_{r_k}.

No model (r_k, f_{r_k}) is likely to perfectly replicate the actual disapproval pattern of all recipients, so we must allow for some error. More precisely, we posit that every recipient deviates from her corresponding rule with probability $\varepsilon > 0$ at each allocation. Therefore, the probability that recipient i follows her assigned rule exactly $x_i(r_k)$ times out of the 10 choices is $(1 - \varepsilon)^{x_i(r_k)} \cdot \varepsilon^{10 - x_i(r_k)}$.[8] Since the decision of the recipients are independent, the joint probability that the data is generated by the model (r_k, f_{r_k}) is then equal to

$$\prod_{i \in D} (1 - \varepsilon)^{x_i(r_k)} \cdot \varepsilon^{10 - x_i(r_k)}$$

Given the joint probability distribution, one can show by standard optimization techniques that the maximum likelihood estimator of the error rate for the model (r_k, f_{r_k}) coincides with the proportion of overall deviations:

$$\hat{\varepsilon}(r_k, f_{r_k}) = \frac{10 \cdot d - \sum_{i=1}^{d} x_i(r_k)}{10 \cdot d}$$

To finish, observe that for any two models (r_k, f_{r_k}) and (r_s, f_{r_s}) such that $s < k$, it follows $\hat{\varepsilon}(r_s, f_{r_s}) \leq \hat{\varepsilon}(r_k, f_{r_k})$, as $x_i(r_s)$ can never be larger than $x_i(r_k)$ for any i. However, the model (r_k, f_{r_k}) contains more rules and it is hence more complex than (r_s, f_{r_s}). For this reason, it makes sense to introduce a penalty that depends on the number of decisions rules in a model. For this, we use the *Akaike Information Criterion* (AIC), according to which the best model (r_k, f_{r_k}) has to maximize the value of the log–likelihood function minus the number of parameters $(10 + d) \cdot k$.

inequity–averse recipient should punish an advantaged dictator. The No–Envy norm is somehow akin to this idea, as disapproval (*i.e.*, non–monetary punishment) is predicted towards dictators who deviate from that norm and hence become advantaged.

[6] We also consider the reverse spite rule for completeness and to check that people do not act incoherently.

[7] Observe that f_{r_k} may not be unique, as there can be several rules j in r_k that fit i's behavior best.

[8] We posit that choices across allocations and games are independent—*i.e.*, the probability of following a rule at any allocation does not depend on what the subject did before. This seems a reasonable assumption because subjects are given no feedback and, hence, there appears to be no reason for mood changes and no room for learning.

3.2.3. Results

Table 3. Error rates, AIC, and the optimal assignment of subjects in the best models of 1, 2, and 3 decision rules (models with more decision rules have a very negative AIC). The following abbreviations are used: ND = No–Disapproval, EN = Envy, HA = Harm, EF = Efficiency, EQ = Equity, and MA = Maximin.

Model			Error Rate	AIC	Assignment of Subjects		
Efficiency			0.2967	–647			
Harm			0.3256	–668			
No–Disapproval			0.3344	–674			
Maximin			0.3544	–685			
Envy			0.4944	–724			
Equity			0.5322	–722			
Spite			0.6656	-764			
Efficiency	Harm		0.1900	–638	49 EF	41 HA	
No–Disapproval	Harm		0.2122	–665	46 ND	44 HA	
Efficiency	Envy		0.2177	–672	63 EF	27 EN	
Harm	Maximin		0.2388	–695	49 HA	41 MA	
Efficiency	No–Disapproval		0.2455	–702	53 EF	37 ND	
Harm	Efficiency	No–Disapproval	0.1566	–690	35 HA	33 EF	22 ND
Efficiency	Harm	Envy	0.1566	–690	46 EF	30 HA	14 EN
Efficiency	Harm	Equity	0.1655	–703	46 EF	34 HA	10 EQ
No–Disapproval	Harm	Maximin	0.1744	–717	40 ND	40 HA	10 MA
Harm	Maximin	Equity	0.1811	–725	41 HA	36 MA	13 EQ

Table 3 presents the results of the classification analysis. We first consider all models where r_k consists of just one rule—*i.e.*, all models assuming homogeneous players—and report for each one the corresponding error rate and the value of the AIC. Note that the models are ordered according to the value of the AIC. We also report the error rate and the AIC for the five best models with two and three rules, together with the assignment of the 90 recipients to their respective rules.[9] For instance, the model with three rules and the maximum AIC includes the Efficiency, the Harm, and the No–Disapproval rule: 33 subjects predominantly followed the first rule, 35 the second, and 22 the third rule. Note also that this model has an error rate of 0.1566, which means that the model can explain almost 85 % of the choices of the average recipient.

[9]When several of the rules in the corresponding set r_k fit the behavior of a recipient best, this recipient is "divided" equally between the tied rules.

If we focus on models with just one rule, the reader can observe that the Efficiency rule describes behavior best, followed by the Harm and the No–Disapproval rule. If we consider models with two rules, the best model has an estimated error rate of 0.1900 and includes the Efficiency and the Harm rule (49 subjects are assigned to the Efficiency and 41 subjects to the Harm rule). Overall, this is the best model in terms of parsimony and accuracy, as it maximizes the Akaike information criterion (–637.60). Models with three rules, in contrast, have a lower AIC. Taking into account both empirical relevance and parsimony, therefore, our analysis suggests that there are two main types of recipients in our dictator games: those who disapprove of inefficient choices and those who disapprove harming ones.

3.3. Costly Messages

The previous analysis has focused on all messages, irrespective of whether the message fee was effectively paid—*i.e.*, we studied what would happen in case messages were costless. Now, we concentrate on the costly messages. There were two main reasons to include them into the design. *First*, some people might be more likely to pay the fee than others, and this can be studied with our design. *Second*, an exclusive focus on costless messages as it is usual in the existing literature is potentially problematic because selfish types might choose messages randomly, which could contaminate our analysis of which behaviors are disapproved. Our elicitation of costly messages can clarify this issue. In effect, extremely different patterns for the costless and the costly messages could be interpreted as evidence that a lot of noise has been introduced by the selfish types choosing randomly.[10]

For several reasons, we highly doubt that (many) recipients chose their hypothetical, costless messages in a random or careless manner. Note first that a model in which all recipients choose randomly in an uniform manner should have an error rate of 0.5000. However, even models with only one rule have much lower error rates. For example, the Efficiency rule alone can explain more than 70 % of the data. This would not be the case if many players acted randomly. To provide another example: the best model in terms of accuracy and parsimony, the one consisting of the Efficiency and the Harm rules, can explain more than 80 % of the recipients' choices. Again, this appears to be totally at odds with the idea that *many* recipients chose carelessly. To make a final point along those lines, note that the Spite rule has the least predictive power among all of our rules (in fact, it performs even worse than a totally random rule). Since this rule predicts an arguably incoherent behavior (disapproval always), its failure is another apparent signal that people took some care when making their choices.

For further evidence, consider now costly messages. Table 4 shows the approval and disapproval rates for costly messages, together with the total number of costly messages sent at each allocation. We observe that the fee was paid in about 8.67 percent of all possible cases, and that 44 (29) messages approved (disapproved) the corresponding dictator's choice. Although the relative scarcity of observations warrants some caution in analyzing these results, some preliminary evidence suggests that motives of efficiency and harm remain important factors. For instance, costly disapprovals are totally absent for some allocations in which there is no loss of social efficiency and/or the recipient is not harmed, as at the right–hand allocation of the first, fourth, and fifth games. In contrast, costly disapprovals are relatively frequent for some allocations in which the recipient is harmed, like the

[10]Note that an alternative interpretation could be that people who pay the fee disapprove in a very different manner from those who are not willing to.

right–hand allocation of game 2 and the left–hand allocation of game 5. Yet, there is some evidence that other factors also play a role. For example, the choice of the socially efficient (even Pareto efficient) allocation 240/120 is disapproved by 66 % of the costly messages, and a similar situation is observed in game 2. This suggests that payoff comparisons play a non–negligible role for the subjects who are most willing to pay in order to express disapproval (note however that the strictly egalitarian allocation 100/100 in game 2 is also disapproved by 70 % of the choices), a point that future research should address.

Table 4. Approval [Disapproval] rates of the recipient in the five dictator games using costly messages. The total number of messages being paid for is presented to the right of the disapproval rates.

	Game			Costly Messages				
	Left–Hand		Right–Hand	Left–Hand Allocation			Right–Hand Allocation		
1	(180,140)	*vs.*	(100,220)	0.3636	[0.5455]	11	1.0000	[0.0000]	13
2	(180,120)	*vs.*	(100,100)	0.4444	[0.5556]	9	0.2000	[0.7000]	10
3	(240,120)	*vs.*	(100,120)	0.3333	[0.6666]	6	0.5000	[0.5000]	4
4	(180,120)	*vs.*	(120,180)	0.2857	[0.2857]	7	1.0000	[0.0000]	9
5	(200,120)	*vs.*	(160,160)	0.2500	[0.7500]	4	1.0000	[0.0000]	5

In order to gain more confidence that the costless messages were not chosen randomly, we have also performed the classification analysis (using again costless messages) on (i) the 65 recipients who never used costly messages and (ii) the remaining recipients who used messages at least once (20 out of these 25 subjects used a costly message more than once). Table 5 presents the error rates, the AIC, and the optimal assignment of subjects within the two subgroups in question.

Most importantly, the results of the subgroup analysis are arguably very similar to the overall findings and across subgroups. In particular, we find that the ordering of the seven decision rules in terms of their explanatory power is the same within both subgroups (Efficiency $>$ Harm $>$ No–Disapproval $>$ Maximin $>$ Envy $>$ Equity $>$ Spite). It also turns out that the corresponding error rates are not significantly different across subgroups at the 5 percent level: the two–sided p–values of the Mann Whitney U tests are equal to 0.9160 for the Efficiency rule, 0.4432 for the Harm rule, 0.0512 for the No–Disapproval rule, 0.4069 for the Maximin rule, 0.3237 for the Envy rule, 0.3338 for the Equity rule, and 0.0512 for the Spite rule. Overall, the evidence seems to be at odds with the idea that the costless messages were chosen in a careless manner, so that we conclude that all our prior analysis based on this type of messages is informative.

Table 5. Error rates, AIC, and the optimal assignment of subjects in the best models of 1 and 2 decision rules (models with more decision rules have a very negative AIC) for the two subgroups using hypothetical messages. The following abbreviations are used: ND = No–Disapproval, EN = Envy, HA = Harm, EF = Efficiency, EQ = Equity, and MA = Maximin.

Model		Error Rate	AIC	Assignment of Subjects	
Subjects who sent at least one costly message					
Efficiency		0.2920	–186		
Harm		0.3000	–188		
No–Disapproval		0.3000	–188		
Maximin		0.3640	–198		
Envy		0.4480	–207		
Equity		0.5000	–208		
Spite		0.7000	–188		
Efficiency	Harm	0.2040	–196	14 EF	11 HA
Efficiency	Envy	0.2040	–196	17 EF	8 EN
No–Disapproval	Harm	0.2120	–199	13.5 ND	11.5 HA
Efficiency	No–Disapproval	0.2280	–204	13 EF	12 ND
No–Disapproval	Envy	0.2440	–209	17 ND	8 EN
Subjects who never sent a costly message					
Efficiency		0.3090	–471		
Harm		0.3454	–486		
No–Disapproval		0.3575	–495		
Maximin		0.3606	–496		
Envy		0.5196	–525		
Equity		0.5515	–527		
Spite		0.6575	–494		
Efficiency	Harm	0.1969	–461	35 EF	30 HA
No–Disapproval	Harm	0.2242	–486	33 ND	32 HA
Efficiency	Envy	0.2348	–495	46 EF	19 EN
Harm	Maximin	0.2454	–503	34 HA	31 MA
Efficiency	No–Disapproval	0.2636	–517	40.5 EF	24.5 ND

4. Conclusion

Our analysis indicates that subjects are heterogenous in regards to which behavior they disapprove and we identify two main behavioral types in this respect: one disapproves when she has been harmed, while the other disapproves when a socially inefficient outcome has been chosen. Further, we observe that other factors, such as envy and equity only play only minor role *in our games* (at least when messages

are costless). This is somehow surprising. Since abundant evidence suggests that inequity–aversion or some kind of egalitarian motives are important to explain monetary punishment (see [17,18,24,25]), we expected a similar result with non–monetary punishment (*i.e.*, disapproval). Perhaps the causes of monetary and non–monetary punishments are different.[11] Another possibility has simply to do with our set of games. In this respect, we hope that our study motivates additional research on disapproval patterns so as to see whether our results generalize to other games or other socioeconomic environments.

While our study attempts to elicit norms by looking at disapproval patterns, related literature has studied other–regarding preferences by looking at how agents allocate payoffs. For instance, [19,27], and [20] report data from several simple games like our dictator games, and classify dictators according to their choices in the games. Since allocation decisions might be driven by fairness norms, the relation to our study is clear. More speculatively, one could guess that allocation decisions are somehow the mirror image of disapproval decisions: a person choosing allocation a as a dictator should disapprove a choice different from a as a recipient. While this idea has obviously some flaws (e.g., self–serving agents are not likely to care about disapproval when being recipients; see also our comment later on self-serving biases), some preliminary evidence seems to be in line with it. For instance, [20] emphasize the importance of social welfare maximization, which seems to coincide with our result that many people disapprove socially inefficient choices.

Even though our focus was not on the effect of the disapproval on the dictators' behavior, we also note that our findings qualify previous results in this respect in the experimental literature. For instance, [28] reports that about one–third of the dictators are willing to pay one dollar to "exit" the game and hence ensure that the recipient never knows that a dictator game was played. Exiting might reflect a desire to not violate the other's expectations, or possibly a desire to not feel disapproved of. In [6,7] ex post messages significantly foster generous behavior in the dictator game. Also, [29] observe that cursing a cheater is a widely–used and effective method to deter cheating in collusive agreements. In contrast, generosity in our game 5 remains at an arguably low level *even* when approval/disapproval is available. One possible reason for these contrasting results is the particular kind of messages that were available, as the latter three cited studies used open–form messages. Maybe it is not only disapproval per se that people care about, but more particularly the way in which it is framed.

There are other issues that merit future attention. A particularly appealing one concerns self–serving biases [30]: do people have a biased view of norms (maybe inadvertently) depending on their position in the game? In our experiment, for instance, we conjecture that dictators would probably exhibit rather different disapproval patterns than the recipients if they were asked what choices merit approval or disapproval (we doubt, for instance, that harm would play such a significant role). However, since dictators might have an interest to manipulate the answers to such questions in order to provide a justification of their choices, the challenge of future research is to obtain sincere answers in this respect.

[11] [26] review the experimental literature on punishment and explored the patterns of both types of punishment in a different set of dictator games with a punishment stage. In line with our results here, they show that harm is important for disapproval, while payoff disadvantages (inequity–aversion) are not. In contrast, both factors are important for monetary punishment.

Acknowledgements

The authors are very grateful to the Economics Department (AE1) of Maastricht University for being able to use their computer lab for running the experiments. Vorsatz gratefully acknowledges financial support from the Spanish Ministry of Education and Science, through the project ECO200907530.

References

1. Gächter, S.; Fehr, E. Collective action as a social exchange. *J. Econ. Behav. Organ.* **1999**, *39*, 341–369.
2. Masclet, D.; Noussair, C.; Tucker, C.; Villeval, M.-C. Monetary and non–monetary punishment in the voluntary contributions mechanism, *Amer. Econ. Rev.*, **2003**, *93*, 366–380.
3. Noussair, C.; Tucker, S. Combining monetary and social sanctions to promote cooperation. *Econ. Inquiry* **2005**, *43*, 649–660.
4. Rege, M.; Telle, K. The impact of social approval and framing on cooperation in public good situations, *J. Public Econ.* **2004**, *88*, 1625–1644.
5. López-Pérez, R.; Vorsatz, M. On Approval and Disapproval: Theory and Experiments. *J. Econ. Psych.* **2010**, *31*, 527–541.
6. Ellingsen, T.; Johannesson, M. Anticipated verbal feedback induces pro-social behavior. *Evol. Hum. Behav.* **2008**, *29*, 100–105.
7. Xiao, E.; Houser, D. Avoiding the sharp tongue: Anticipated written messages promote fair economic exchange. *J. Econ. Psychol.* **2009**, *30*, 393–404.
8. Elster, J. Social norms and economic theory. *J. Econ. Perspect.* **1989**, *3*, 99–117.
9. Posner, R. Social norms and the law: An economic approach. *Amer. Econ. Rev.* **1997**, *87*, 365–369.
10. Horne, C. Sociological Perspectives on the Emergence of Social Norms. In *Social Norms*; Hechter, M., Opp, K.–D., Eds.; Russel Sage Foundation: New York, NY, USA, 2001.
11. Cowen, T. The esteem theory of norms. *Public Choice* **2002**, *113*, 211–224.
12. Fehr, E.; Falk, A. Psychological foundations of incentives. *Europ. Econ. Rev.* **2002**, *46*, 687–724.
13. Krupka, E.; Weber, R. Identifying social norms using coordination games: Why does dictator game sharing vary? Mimeo, 2008.
14. El–Gamal, M.; Grether, D. Are people Bayesian? Uncovering behavioral strategies. *J. Amer. Statistical Assoc.* **1995**, *90*, 1137–1145.
15. Brandts, J.; Charness, G. The strategy *versus* the direct-response method: A first survey of experimental comparisons. *Exper. Econ.* **2011**, *14*, 375–398.
16. Fischbacher, U. z-Tree: Zurich toolbox for ready-made economic experiments. *Exper. Econ.* **2007**, *10*, 171-178.
17. Bolton, G.; Ockenfels, A. ERC: A theory of equity, reciprocity, and competition. *Am. Econ. Rev.* **2000**, *90*, 166–93.
18. Fehr, E.; Schmidt, K. A theory of fairness, competition and cooperation. *Quart. J. Econ.* **1999**, *114*, 817–868.
19. Andreoni, J.; Miller, J. Giving according to GARP: An experimental test of the consistency of preferences for altruism. *Econometrica* **2002**, *70*, 737–53.

20. Charness, G.; Rabin, M. Understanding social preferences with simple tests. *Quart. J. Econ.* **2002**, *117*, 817–869.
21. Rabin, M. Incorporating fairness into game theory and economics. *American Economic Review* **1993**, *83*, 1281–1302.
22. Kirchsteiger, G. The role of envy in ultimatum games. *J. Econ. Behav. Organ.* **1994**, *25*, 373–389.
23. Holländer, H. A social exchange approach to voluntary cooperation. *Amer. Econ. Rev.* **1990**, *80*, 1157–1167.
24. Dawes, C.; Fowler, J.; Johnson, T.; McElreath, R.; Smirnov, O. Egalitarian motives in humans. *Nature* **2007**, *446*, 794–796.
25. Zizzo, D. Money burning and rank egalitarianism with random dictators. *Econ. Letters* **2003**, *81*, 263–66.
26. Leibbrandt, A.; López-Pérez, R. Different carrots and different sticks: Do we reward and punish differently than we approve and disapprove? Mimeo, 2011.
27. Charness, G.; Grosskopf, B. Relative payoffs and happiness: An experimental study. *J. Econ. Behav. Organ.* **2001**, *45*, 301–328.
28. Dana, J.; Cain, D.; Dawes, R. What you don't know won't hurt me: Costly (but quiet) exit in dictator games. *Organ. Behav. Hum. Dec.* **2006**, *100*, 193–201.
29. Cooper, D.; Kühn, K. Communication, renegotiation, and the scope for collusion. CEPR Discussion Paper No. 7563. Publisher: city, country, 2009.
30. Konow, J. Fair shares: Accountability and cognitive dissonance in allocation decisions. *Amer. Econ. Rev.* **2000**, *90*, 1072–1091.

Appendix: Pooling data over PD treatments

This appendix presents the statistical analysis that supports our claim in footnote 3 that the recipients' behavior in our dictator games does not vary substantially conditional on the prisoner's dilemma (PD) treatment conducted before. Detailed information on the PD games can be found in López-Pérez and Vorsatz (2010), so we try here to be as concise as possible. There were three different PD treatments. In the *control* treatment, a standard PD was played. In the *expectations* treatment, we elicited players' expectations about approval/disapproval from their co-player before playing the control PD. Finally, the *feedback* treatment consisted of a two-stage game. In the first stage, subjects played the PD game. In the second stage (the feedback stage), they were able to approve/disapprove the co-player's prior choice in the PD by means of a message that had no effect on the receiver's material payoff.

Table 6 uses hypothetical messages and presents the disapproval rates in each allocation of the five dictator games, conditional on the PD treatment played before.

We observe at first sight that disapproval rates do not vary much across treatments. For a more formal comparison, Table 7 reports the two-sided p-values of the Z-tests that analyze whether the mean disapproval rate is the same across treatments (pairwise comparisons for all possible outcomes). We find that out of the 30 possible tests, only one difference is significant at the five percent confidence level (the one that appears in boldface). Consequently, it is possible to pool the data.

Table 6. Disapproval rates of the recipient in the five dictator games (costless messages) for all three PD treatments.

Treatment	Game 1		Game 2		Game 3		Game 4		Game 5	
	Left	Right	Left	Right	Left	Right	Left	Right	Left	Right
Control	0.45	0.14	0.03	0.66	0.14	0.55	0.45	0.17	0.52	0.07
Expectations	0.29	0.26	0.23	0.68	0.13	0.55	0.45	0.23	0.65	0.20
Feedback	0.47	0.10	0.13	0.60	0.23	0.53	0.47	0.10	0.47	0.10
Overall	0.40	0.17	0.13	0.64	0.17	0.54	0.46	0.17	0.54	0.12

Table 7. Two-sided p-values of the Z-tests that compare the mean disapproval rates across treatments. C = control treatment, E = expectations treatment, F = feedback treatment.

Treatment	Game 1		Game 2		Game 3		Game 4		Game 5	
	Left	Right	Left	Right	Left	Right	Left	Right	Left	Right
C *vs.* E	0.20	0.25	**0.03**	0.86	0.92	0.98	0.98	0.60	0.31	0.16
C *vs.* F	0.88	0.65	0.17	0.66	0.35	0.88	0.88	0.42	0.70	0.67
E *vs.* F	0.15	0.11	0.35	0.53	0.29	0.90	0.91	0.18	0.16	0.30

Appendix: Instructions

The following are the original instructions used in the experiment. As the careful reader may note, the instructions contain a few typos and grammar mistakes. However, these mistakes are minor and could not drive in any manner our results, among other reasons because participants had to answer some control questions to ensure that they understood the instructions.

Welcome

Dear participant, thank you for taking part in this experiment. It will last about 60 minutes. If you read the following instructions carefully, you can – depending on your decisions – earn some money. The entire of money which you earn with your decisions will be paid to you in cash at the end of the experiment. These instructions are solely for your private information.

We will not speak of Euros during the experiment, but rather of ECU (Experimental Currency Units). Your whole income will first be calculated in ECU. At the end of the experiment, the total amount you have earned will be converted to Euro at the following rate:

20 ECU = 1 Euro.

In order to ensure that the experiment takes place in an optimal setting, we would like to ask you to abide by the following rules. If you do not obey them, we will have to exclude you from this experiment and you will not receive any compensation.

- Do not communicate with your fellow students. If you have any doubts, raise your hand and one of the experimenters will clarify them privately.
- do not forget to switch off your mobile phone!

- you may take notes on this instruction sheet if you wish.
- when the experiment finishes, remain seated till we pay you off.

The Experiment

In the experiment you will participate in six different scenarios, and you will be paid for your decisions in one scenario, randomly chosen at the end of the experiment. More precisely, the participant playing at the computer number 9 will roll a die and we will pay you the equivalent in Euros of your ECU earning in the scenario corresponding to the number that turns up.

In what follows we will explain to you only the first scenario. Once you made your decision in this first scenario, we will introduce the five remaining scenarios. Note well that each scenario is independent of the others; that is, your payoff in any scenario does not depend on decisions taken in other scenarios.

Scenario 1

In this scenario, you have been randomly and anonymously matched with another participant and both of you have to choose independently between alternative X and alternative Y. Depending on your choices, you will get the following ECU payoff:

- if you both choose X, both of you get 180 ECU.
- if you choose X and the other participant chooses Y, you get 80 ECU and the other participant gets 260 ECU.
- if you choose Y and the other participant chooses X, you get 260 ECU and the other participant gets 80 ECU.
- if you both choose Y, both of you get 100 ECU.

The matrix below summarizes this.

Payoff Table

		The other player	
		X	Y
You	X	(180,180)	(80,260)
	Y	(260,80)	(100,100)

Decisions at this scenario will be private; that is, you will never be informed about the decisions of any other participant in this scenario, and no other participant will know your decision in this scenario. Apart of choosing between X and Y, in this scenario both of you have the possibility to send one message to the other participant with your opinion about her/his choice. Sending a message costs 10 ECU. Since decisions are private, however, you will not know whether the other participant chose X or

Y. For this reason, we will ask you whether you want to send a message for any possible contingency. More precisely, the procedure will consist of three steps:

1. You will be asked the following question: "Suppose both of you chose X. Do you want to send a message paying a cost of 10 ECU?" You can choose either *Yes* or *No*.

2. Independently of your answer to the previous question, you are then asked the following question: "Suppose you decided to send a message to the other participant in case both of you chose X. Which of the following three messages do you send?"

    - Your choice was good.
    - Your choice was neither good nor bad.
    - Your choice was bad.

3. The same two previous steps are then repeated for any of the other three possible combinations of choices: (X, Y), (Y, X), and (Y, Y).

Observe again that decisions are always private; that is, none of you will know the choices of the other participant when going to the next scenario (including messages). The actual decisions in one scenario will be anonymously revealed to both participants only if, at the end of the experiment, it turns out that this scenario is randomly chosen for payment. If the die selects scenario 1, moreover, each participant will receive the message selected by the other participant at that scenario (step 2) only if the other participant previously chose *Yes* in step 1. Finally, observe that once the first scenario has finished, you will be matched to a different participant for the remaining five scenarios.

Control Questions

Please answer the following control questions. Once you have written down all your answers, please raise your hand so that one of the experimenters can check them.

1. How many different scenarios are there?

2. If you choose X and the other participant chooses X, what will be your payoff?

3. If you choose X and the other participant chooses$Y$, what will be your payoff?

4. Are you always matched with the same participant?

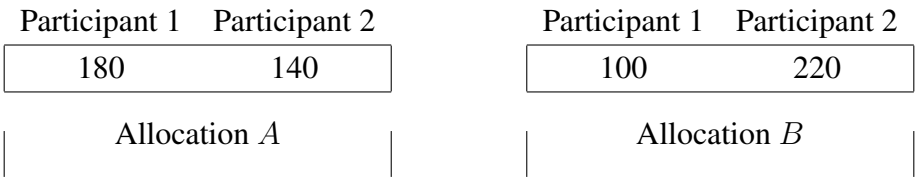
5. How will your final payoff (in Euro) be determined?

Scenarios 2–6

You will be matched throughout the rest of the experiment with a participant different from the one you have been matched with in the first scenario. One of you will be throughout all remaining scenarios a type 1 participant, and the other will be throughout a type 2 participant. Your type will appear soon on the computer screen.

In what follows, we will only explain scenario 2 in detail, since the remaining four scenarios have the same structure. As Figure 1 shows, participant 1 has to decide between two ECU allocations: A and B. Allocation A gives 180 ECU to participant 1 and 140 ECU to participant 2, and allocation B gives 100 ECU to participant 1 and 220 ECU to participant 2.

Figure 1. ECU allocations (to be decided by participant 1).



After participant 1 has made her/his decision, participant 2 gets the possibility to send a message to participant 1 with her/his opinion about the choice being made. Sending a message costs 10 ECU to participant 2. Note well, however, that participant 2 will not be informed about the actual decision of participant 1, instead she/he is asked how she/he would react in case participant 1 chose either of the two allocations. More precisely, the procedure will consist of three steps:

1. Player 2 is first asked the following question: "Suppose participant 1 chose allocation A. Do you want to send a message paying a cost of 10 ECU?" Participant 2 can either choose *Yes* or *No*.

2. Independently of the her/his answer to the previous question, participant 2 is then asked the following question: "Suppose you decided to send a message to the other participant in case she/he chose allocation A. Which of the following three messages do you send?"

    - Your choice was good.
    - Your choice was neither good nor bad.
    - Your choice was bad.

3. The same two previous steps are then repeated with allocation B.

The same procedure applies to the remaining four scenarios, which will differ only in the allocations available to participant 1.

Observe that decisions are always private; that is, neither participant 1 nor participant 2 will know the choices of the other participant when going to the next scenario. The actual decisions in one scenario

will be anonymously revealed to both participants only if, at the end of the experiment, it turns out that this scenario is randomly chosen for payment. Further, and if the die selects one of the scenarios 2 to 6, participant 1 will receive the message selected by participant 2 at that scenario (step 2) only if participant 2 previously chose *Yes* in step 1.

An example will clarify these points. Suppose that scenario 2 is selected by the roll of the die at the end of the experiment, and that participant 1 chose allocation A (180/240) in that scenario. The following two things could then happen:

1. If participant 2 decided to send a message, then participant 1 will receive 180 ECU, participant 2 will receive 230 ECU (240 minus the 10 ECU cost for sending the message), and participant 1 will observe the message sent by participant 2.

2. On the other hand, if participant 2 decided not to send a message, then the final payoff would be 180 ECU to participant 1 and 240 ECU to participant 2.

Similarly, if participant 1 chose allocation B (100/220) instead, the final payoff would be 100 ECU to participant 1 and either 220 ECU or 210 ECU to participant 2. The latter would occur if participant 2 decided to send a message after participant 1 chose allocation B.