Journal of Economic Psychology 39 (2013) 268–277

Contents lists available at ScienceDirect

Journal of Economic Psychology

journal homepage: www.elsevier.com/locate/joep

# Third-party sanctioning and compensation behavior: Findings from the ultimatum game

Alex K. Chavez [a,*], Cristina Bicchieri [b]

[a] Department of Psychology, University of Michigan, Ann Arbor, MI, USA
[b] Departments of Philosophy and Psychology, University of Pennsylvania, Philadelphia, PA, USA

ARTICLE INFO

*Article history:*
Received 21 September 2012
Received in revised form 12 August 2013
Accepted 12 September 2013
Available online 19 September 2013

*JEL classification:*
C38
C72
C91
D63

*PsycINFO classification:*
3020
2300

*Keywords:*
Social norms
compensation
rewarding
third-party punishment
Fairness
Equality

ABSTRACT

We measured the beliefs and behavior of third parties who were given the opportunity to add to or deduct from the payoffs of individuals who engaged in an economic bargaining game under different social contexts. Third parties rewarded bargaining outcomes that were equal and compensated victims of unfair bargaining outcomes rather than punishing perpetrators, but were willing to punish when compensation was not an available option. Beliefs of whether unequal bargaining outcomes were fair differed based on the normative context, but actual punishment, compensation, and rewarding behavior did not. This paper makes a contribution to the literature on informal mechanisms of social norm enforcement by comparing negative sanctions, positive sanctions, and compensation behavior by third parties.



## 1. Introduction

Without legal enforcement, social norms rely on informal sanctions to produce norm-conforming behavior. Empirical studies have reported evidence of negative sanctions (e.g., punishment) both by parties directly involved in the norm-governed interaction (de Quervain et al., 2004), and by uninvolved third parties (Fehr & Fischbacher, 2004; Kahneman, Knetsch, & Thaler, 1986; Kurzban, DeScioli, & O'Brien, 2007). In these studies, the punished transgression refers to a violation of social norms of fairness or reciprocity, but usually no independent evidence is presented about the players' consensus about the norm's existence and relevance to the experimental situation. In the definition that we adopt, a social norm (Bicchieri, 2006, p. 11) is a behavioral rule for which it must hold for sufficiently many people that: (1) they know that such an approved behavioral rule exists; (2) they prefer to comply with the rule if they believe (a) that others will also comply

* Corresponding author. Tel.: +1 734 619 0407; fax: +1 781 623 0407.
*E-mail addresses:* achavez@sas.upenn.edu (A.K. Chavez), cb36@sas.upenn.edu (C. Bicchieri).

0167-4870/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.joep.2013.09.004

(empirical expectations) and (b) that others believe they ought to comply and might sanction non-compliance (normative expectations). We explore the possibility that individuals may employ compensation and rewards – in addition to sanctions – as mechanisms for upholding social norms.

Notice this definition allows a social norm to exist while not always being followed. Only if sufficiently many people have the appropriate empirical and normative expectations about others' behavior and beliefs will a social norm reliably be followed. Consensus and compliance may thus differ. While observed behavior remains a crucial measure of compliance with elicited norms, recent experimental work has introduced the explicit use of questionnaires to assess normative consensus (Bicchieri & Chavez, 2010; Bicchieri, Xiao, & Muldoon, 2011; Krupka & Weber, 2013; Rauhut & Winter, 2010; Reuben & Riedl, 2013). In this paper, we therefore measure both the beliefs and behavior of third parties who tradeoff sanctioning, compensating, and rewarding in response to the violation of or compliance with a social norm.

An example of informal sanctioning is the punishment of individuals who divide a good unfairly, when the understanding of what constitutes fair division depends on context. Under an *equality* context, a fair division is one in which goods are allocated equally amongst all parties. Under an *equity* context, a fair division is one that divides goods according to each party's share due to merit or acquired right. Assessing normative consensus should show different judgments of what counts as fair in equity versus equality conditions and this differential assessment should be reflected in behavior. So we would expect third-party punishment levels to differ under contexts that invoke rules either of equity or equality. In particular, we would expect uneven divisions to be punished less under equity contexts in which one party is perceived as being entitled to a greater share of the good.

Although costly punishment is more commonly studied, it is not the sole mechanism for upholding a social norm. Costly rewarding also plays a role in supporting pro-social behavior. When both punishments and rewards were available to the players, second-party rewarding was frequent and promoted cooperation (Andreoni, Harbaugh, & Vesterlund, 2003; Rand, Dreber, Ellingsen, Fudenberg, & Nowak, 2009). Third-party rewarding was also common when there was the possibility of reputation formation (Milinski, Semmann, & Krambeck, 2002; Seinen & Schram, 2006). Few experimental studies, however, have compared third party punishment and rewarding in a non-repeated, reputation free game (cf. Almenberg, Dreber, Apicella, & Rand, 2011). Whereas positive and negative sanctions promote norms in that they increase their long-term expected realization, another mechanism for upholding social norms is compensation of the victim of a violation.

Compensation is not directly norm promoting. It rather *honors* norms, as it points to the fact that the victim *ought* to have received a fair share or had a right to be treated fairly. If indeed compensation has such an important signaling function, we would expect all three regulatory mechanisms – compensation, punishment, and rewarding – to be employed in upholding social norms.

Our hypothesis is thus that third parties will employ both compensation and positive sanctions in addition to negative sanctions when all three regulatory mechanisms are available. Furthermore, it is worthwhile to investigate whether the presence of all three mechanisms leads individuals to discount one in favor of another. For example, the opportunity to compensate victims might focus third parties on feelings of compassion, which has been found to decrease punitive sentiment (Condon & DeSteno, 2011).

To test our hypothesis, we focused on a version of the Ultimatum Game (Guth, Schmittberger, & Schwarze, 1982) in which *Proposers* proposed a division of a sum of $10 to *Responders*, who accepted or rejected the offer. In the case of a rejection, both parties received nothing. Previous studies found that both Proposer behavior and second-party punishment were sensitive to different fairness contexts. When the roles of Proposer and Responder were assigned randomly, Proposers offered an average of 45% of the sum (for a review, see Camerer, 2003, chap. 2). But when the role of Proposer was earned by higher scorers on a general knowledge quiz, Proposers offered an average of only 35%. Moreover, overall rejection rates were the same whether roles were assigned randomly or based on quiz performance (Hoffman, McCabe, Keith, & Smith, 1994), although Proposers offered moderately less in the latter condition, apparently feeling entitled to a larger share under an equity context. Similarly, Falk, Fehr, and Fischbacher (2003) created different offer contexts by allowing the Proposer to (1) choose between offering 20% and 50%, and (2) choose between offering 20% and 0%. They found that in the first context, Proposers offered 20% only 31% of the time, but in the second context they always offered 20%. Moreover, in the first context, Responders who were offered 20% rejected the offer 44% of the time, whereas in the second context, Responders who were offered 20% rejected only 9% of the time. In the first context, the presence of a 50/50 option prompts a norm of fairness, whereas in the second case an offer of 20% may just be perceived as generous (see Bicchieri, 2006, chap. 3). Thus, second-party punishment is clearly sensitive to context, and whether a given context elicits a specific norm.

Is third-party punishment, compensation, or rewarding behavior sensitive to fairness context? To answer this question, we conducted a two-stage study in which participants engaged in an Ultimatum Game under an equity or equality context in Stage 1, and in Stage 2, third parties decided to add to or deduct from the payoffs of participants from Stage 1 based on their bargaining outcome. To determine whether third parties were willing to trade off compensation and punishment, we created a separate experimental condition in which third parties only had the option of deducting from the payoffs of Stage 1 participants. This design allowed us to assess whether third parties preferred to compensate victims or punish perpetrators of unfair bargaining outcomes, and to assess whether third parties would reward fair bargaining outcomes. We discuss data only from Stage 2 below.