Article

RATIONALITY AND SOCIETY

Not 'Just the two of us': Third party externalities of social dilemmas

Rationality and Society
23(3) 347–370

DOI: 10.1177/1043463111412716
rss.sagepub.com
SAGE

Hanne van der Iest
Independent Researcher

Jacob Dijkstra and Frans N. Stokman
University of Groningen, The Netherlands

Abstract
Many real-life social dilemmas contain third parties who cannot make decisions in the dilemma, but are affected by its outcome (receive externalities) nonetheless. Dilemmas with identical payoffs for decision-making *actors* may greatly vary in their externalities for *third parties*. If actors value the welfare of thirds, externalities will affect actors' decisions. We test behavioral predictions from three leading ideas on social preferences (altruism, inequality aversion, competition) in two studies that employ four one-shot, 2-person prisoner's dilemmas (PDs) that differ only in their externalities. The PDs respectively include a third party that (i) is indifferent, (ii) prefers defection, (iii) prefers cooperation. Our results show that while aggregate behavior is not affected by externalities, individual behavior is. Compared to a PD without externalities, prosocial individuals cooperate more when a third benefits from cooperation, but do not defect more when a third benefits from defection. The opposite pattern is found for competitive individuals.

Keywords
externality, prisoner's dilemma, social preferences, social value orientation, third party

Corresponding author:
Hanne van der Iest, c/o Jacob Dijkstra, University of Groningen, Grote Rozenstraat 31, 9712 TG Groningen, The Netherlands.
Email: hannevanderiest@hotmail.com

Introduction

Social dilemmas are situations in which the concurrence of individually rational behavior of a set of actors leads to an undesirable outcome for these same actors (Dawes, 1980). But apart from affecting the actors, dilemma outcomes frequently affect the well-being of *third parties* as well. By definition, third parties cannot make decisions in the dilemma, and are thus unable to influence its outcome directly. The effects of outcomes of social dilemmas for third parties are therefore referred to as *externalities*. In this paper we will theoretically argue and empirically show that externalities for third parties are important for understanding and predicting behavior in social dilemmas.

In fact, many real-life social dilemmas produce externalities for third parties. As a first example consider the case of a price cartel. Companies in a price cartel are supposed to cooperate by keeping prices high, but could individually decide to exploit other cartel members (i.e., defect) by lowering their prices. Not only cartel members are affected by price decisions, however. Above all, consumers experience favorable consequences of *defection* (i.e., price cutting). Thus, price decisions of cartel members yield externalities for consumers.

As a second example, consider the case of a set of fishermen exploiting the same fishing grounds. Each fisherman is individually better off catching more than quotas allow (i.e., defection), but if quotas were transgressed regularly, the fish population would deteriorate. But not only the fishermen bear the consequences of overfishing. Consumers are benefited by a stable fish supply and biodiversity is generally valued. In this case, and as opposed to the cartel example, third parties experience favorable consequences of the *cooperation* of fishermen.

Note that from the perspective of a rationally egoistic actor in the dilemma, the payoff structures in both dilemmas are the same. Defection is a dominant strategy in both cases as it leads to a higher payoff than cooperation regardless of what the other actors do. However, mutual cooperation yields higher payoffs for all than mutual defection.

The two dilemmas are nevertheless essentially different with respect to their externalities for third parties. While in a price cartel defection usually benefits third parties, in a resource dilemma third parties are mainly benefited by cooperation. Cooperation in social dilemmas can thus no longer be unconditionally regarded as the optimal (Pareto superior) collective outcome, as it may damage third parties.

Contrary to the rational egoist assumption, we argue that externalities for third parties *do* affect the choices actors make in social dilemmas. The central

claim of this paper is that behavior in many real-life social dilemmas cannot be correctly interpreted when externalities for third parties are neglected.

The literature on social preferences offers ample evidence that decision makers take others' payoffs into consideration (e.g. Gintis *et al*., 2003; Kollock, 1998a; Simpson and Willer, 2008). Actors in social dilemmas who value (either positively or negatively, and either in absolute or relative terms) the well-being of third parties experiencing externalities, consequently account for this well-being when making their decisions. Thus, although from a rational egoist's point of view the payoff structure for the actors in the cartel dilemma and the resource dilemma is similar, behavior in these dilemmas will differ because of the externalities for third parties.

Kollock (1998b) distinguishes strategic and motivational solutions to social dilemmas with a fixed structure (see also Yamagishi, 1995). Externalities can affect actors' decisions in social dilemmas either because they may have strategic consequences or because they inherently have motivational consequences. Specifically, if third parties themselves or other agencies acting on their behalf (e.g., the government) can punish or reward an actor for externalities caused to them, these anticipated sanctions are likely to affect the actor's choice in the current social dilemma (e.g. Axelrod, 1984; Ostrom *et al*., 1992; Yamagishi, 1986) for strategic reasons. In addition, actors who directly value the well-being of third parties in one way or another may account for externalities in their decisions, even in the absence of strategic considerations. Thus, externalities have motivational consequences in as far as actors hold social preferences.

In real-life social dilemmas, these two different motives will often occur simultaneously. But to gain insight in the effects of externalities, they need to be examined separately. The influence of externalities on behavior that is motivated by social preferences is independent of specific decision-settings (such as the likelihood of repeated interaction, or the presence or absence of a punishing agency) whereas the effects of future sanctions and rewards are dependent on the decision setting. This paper examines the first. Based on the social preference literature (e.g., Fehr and Fischbacher, 2002) we expect that people are sensitive to the presence of third parties and that they will take their welfare into account when making their decisions. Our main hypothesis is therefore:

> H1. In a social dilemma, externalities for a third party affect the behavior of actors in this dilemma.

We will investigate the effect of externalities by means of two experiments building upon the standard one-shot, 2-person prisoner's dilemma

game (which we will refer to as the 'standard PD'). We examine how dilemma decisions are affected by the presence of (i) an indifferent third party, (ii) a third party that is benefited by defection, and (iii) a third party that is benefited by cooperation. The third party cannot affect the game's payoffs directly.

We will derive predictions from four well known *motivational principles,* of which the latter three concern principles based on social preferences. These principles are: *egoism*, *altruism*, *inequality aversion*, and *competition*.

In Study 1 we evaluate our main hypothesis and the predictions of the motivational principles in a between-subjects experiment, in which we do not attempt to measure the motives of the participants. The results of Study 1 suggest that there are no differences in cooperation between a standard PD and PDs with externalities. There are nevertheless differences in cooperation between the various PDs that include externalities. However, no single one of the motivational principles accurately predicts participants' behavior.

In Study 2 we therefore conduct a within-subjects experiment in which we measure participant motivation with the Triple-Dominance Scale of Social Values (e.g., Van Lange *et al*., 1997). The results of Study 2 show that behavioral changes due to externalities are moderated by Social Value Orientation (SVO).

In the next section we will discuss the four motivational principles, derive predictions from them and introduce the game we use in our two studies. In the two subsequent sections we describe Study 1 and Study 2, and report the results. We conclude with a discussion and pose questions for future research.

Theory and predictions

Social preferences are usually modeled by including the payoffs of others in an individual's utility function (e.g. Bolton and Ockenfels, 2000; Fehr and Schmidt, 1999; Levine, 1998; Rabin, 1993). One problem with utility functions is that they can take many different forms. One and the same motivational principle (such as altruism) can be modeled in an indefinite number of ways. So when a model's predictions are refuted it is impossible to conclude whether the motivational principle itself or its formalization is flawed. For this paper we therefore refrain from specifying the precise functional form of actors' utility functions. Instead we design an experimental game where different hypotheses result from different *general assumptions* about actors' social preferences.

Assumptions A1 and A2 below, are common to all our models of the four motivational principles. The principles are distinguished by a third assumption:

A1. Actors are expected utility maximizers
A2. Actors' utilities increase when their payoff increases, *ceteris paribus*
A3. 0 *Egoism* – payoffs of others are irrelevant
A3. 1 *Altruism* – actors' utilities increase when the payoff of others increases, *ceteris paribus*[1]
A3. 2 *Inequality aversion* – actors' utilities decrease when the difference between their own payoff and payoffs of others increases, *ceteris paribus*
A3. 3 *Competition* – actors' utilities increase when their payoff is larger than the payoff of others, and their utilities decrease when their payoff is smaller than the payoff of others, *ceteris paribus*

The first motivational principle is *egoism*. In egoism other actors' payoffs are irrelevant to the decision maker. According to egoism, the classical game theory postulate, an individual's own payoff is the *only* relevant incentive. It therefore predicts that actors will always defect in a standard PD, and thus cannot explain any cooperation in this situation.

The second motivational principle we consider is *altruism* (e.g. Gintis *et al*., 2003). The higher the payoffs of others, the better for the altruist. In fact, Gintis *et al*. (2003) consider an act altruistic when it benefits another at a cost to the individual. This implies A3.1, but the reverse is not true. The third principle we consider is *inequality aversion*. Inequality aversion is people's preference for equal outcomes over unequal outcomes (e.g. Dawes *et al.,* 2007; Eek and Gärling, 2006; Fehr and Schmidt, 1999). The fourth principle we consider is *competition* (e.g. Frank, 1985; Van Lange*,* 1999). Competitive individuals like having a higher payoff than others.

Existing versions of these models (that *do* include a functional form) make the simplifying assumption that actors have the same social preferences towards different individuals. In our setting, this implies that actors have the same preferences regarding actors and third parties. We will maintain this simplifying assumption, while we realize that social preferences may be more nuanced in reality.

Note that the assumptions formulated above are too general to specify actors' utility functions. They are specific enough, however, to allow the derivation of hypotheses for our experimental game.

In the present paper we extend the standard PD with positive payoffs by including positive externalities for a third party. We use the standard PD since it is the basic model used in much social dilemma research, facilitating comparisons of our results with those of previous studies. Our experiments are designed to test whether and how externalities influence behavior, not to

Table 1. Payoffs for every choice combination of participants A and B.

		Participant B	
		Cooperation ($)	Defection ($)
Participant A	Cooperation	3, 3, X[a]	0, 5, 2[a]
	Defection	5, 0, 2[a]	1, 1, Y[a]

[a] The first, second, and third character in each cell represent the payoffs of Participant A, Participant B and the third party, respectively.

Table 2. Values for payoffs of the third party in Condition I–IV; C = Cooperation, D = Defection.

	CC ($X)	CD($)	DC($)	DD ($Y)
Condition I	n/a	n/a	n/a	n/a
Condition II	2	2	2	2
Condition III	0	2	2	5
Condition IV	5	2	2	0

mimic any particular real-life social dilemma. The payoff matrix of our experimental PD is displayed in Table 1.

Note that in Table 1 the externalities for the third party are indicated by letters in two cells (e.g. $X and $Y). The conditions of our experiments are distinguished by the values these letters take (see Table 2). Table 1 shows that the third party receives a payoff of $X when both actors cooperate, receives a payoff of $2 when either actor A or actor B cooperates while the other defects, and earns a payoff of $Y when A and B both defect. The experiments of Study 1 and Study 2 contain the same four conditions.

Condition I is the standard PD without a third party. Condition II includes a third party that is indifferent with respect to the behavior of A and B: $X = $Y = $2. In Condition III, the third party profits from defection of A and B: $X = $0, $Y = $5. In Condition IV, the third party profits from cooperation of A and B: $X = $5, $Y = $0. Condition II, III, and IV all include a third party, unlike Condition I. Condition II deviates from III and IV only in its *externalities* for the third party.

In our studies, we use positive payoffs only. We are aware that people respond differently toward gains and losses (Kahneman and Tversky, 1979) and may thus behave differently in dilemmas with positive versus negative externalities or payoffs. But since the aim of our experiment is to detect

whether actors respond to externalities, we leave the refinement of our basic design to future research. Using only positive payoffs has the additional benefit of assuming that the scaling of utilities is irrelevant, which simplifies our models.

We can now derive predictions from the four motivational principles for the comparisons between conditions II, III, and IV. Predictions are based on an actor's expected utility of either defecting or cooperating in a particular condition. To avoid sticking to just one functional form of a motivational principle, we do not make any assumptions about the shape of actors' utility function. Instead, we choose to test the motivational principles more generally (see Appendix for the mathematical derivation).

Since a rational egoist is not affected by other actors' payoffs, *egoism* predicts no behavioral differences between conditions. I.e., the proportion of cooperation in Condition II (π_{II}) = the proportion of cooperation in Condition III (π_{III}) = the proportion of cooperation in Condition IV (π_{IV}). The *altruism assumption,* however, predicts that people will cooperate more (or at least not less) when the third party benefits from cooperation than when the third is indifferent. Additionally, this assumption predicts that people will cooperate more (or at least not less) when the third party is indifferent than when it benefits from defection (i.e. $\pi_{IV} \geq \pi_{II} \geq \pi_{III}$).

The *inequality aversion assumption* predicts that people will cooperate more (or at least not less) when the third party is indifferent than when it prefers cooperation (i.e. $\pi_{II} \geq \pi_{IV}$).

According to the *competition assumption,* people cooperate more (or at least not less) when the third party prefers defection than when it is indifferent. Furthermore, people cooperate more (or equally) when the third party is indifferent than when it benefits from cooperation (i.e. $\pi_{III} \geq \pi_{II} \geq \pi_{IV}$). Note how the competition and inequality aversion assumptions agree in their ordering of Condition II and Condition IV. If competition is the driving motivational force, this in addition implies a specific ordering of Condition III and Condition II. Thus, although corroboration of the competition prediction does not logically rule out inequality aversion as a motivational principle, observing $\pi_{III} \geq \pi_{II}$ would suggest competition. The following section describes Study 1 that is designed to test the predictions.

Study 1

Methods

Participants and materials. 176 undergraduate students (54.3% women, mean age = 19.6) from the University of South Carolina participated in the

experiment. Students were recruited through an online database. By participating, they were granted a course credit point. In addition, they could earn between $0 and $5, depending on the outcome of the game they played. The experiment was conducted with paper and pen.

Procedure

Each participant was randomly assigned to one of the four conditions: 43 participants to Condition I, 44 to Condition II, 44 to Condition III, and 45 to Condition IV. Upon entering the experiment room, participants were directed to separate cubicles so that they could not communicate with each other in any way. Participants were then given written instructions including the payoff matrix belonging to their experimental condition.

The instructions informed the participants that they were paired with one (Condition I) or two (Conditions II, III, and IV) other participants and that they had to make a single choice between Option 1 or 2 (i.e. cooperation or defection; since these are terms with strong normative connotations, we made no notion of them to participants). It was stressed that the *joined decisions* of themselves and one other participant determined their final payoff, and that these decisions would be made simultaneously.

In Conditions II, III, and IV, participants were informed that the third party could not make a decision, and that their payoffs therefore depended solely on the combined decisions of themselves and one other participant. Since the experiment did not involve deception, an actual third person was paid. Students who had just finished different experiments in the same experiment room took the position of the third party.

All participants knew that their decisions in the experiment were anonymous, that their earnings (if any) would be paid privately, and that they had to leave the experiment room directly after they got paid, to guarantee anonymity.

After reading the instructions, participants answered a quiz to verify their understanding of the rules of the experimental game. Whenever participants made any mistakes in the quiz, extra care was taken that they fully comprehended the game before proceeding. Participants were then given a new sheet of paper with the same payoff matrix they had been presented with in the instructions. They were told to make a single choice between Options 1 and 2 (i.e. cooperation and defection).

To check the robustness of their decisions under small changes in the payoffs, participants from Conditions II, III, and IV then had to make the same choice given a slightly altered payoff matrix. To facilitate the robustness check, these participants were randomly allocated to Group Ψ (psi) or

Group Ω (omega). This time the choice was hypothetical, and it was stressed that the second choice did not affect their monetary payoffs in any way. Subsequently, participants listed their age and gender. They were then paid their earnings privately and left the experiment room. On average one session lasted 20 minutes.

Design

Study 1 had a between-subjects design and in each condition, subjects played the one-shot PD of Table 1. Condition I represented a standard PD without a third party, so no payoffs for the third party were given. In the other three conditions the payoffs for the third party were $X = $Y = $2, $X = $0 and $Y = $5, and $X = $5 and $Y = $0, for Conditions II, III, and IV, respectively (see Table 2).

For the robustness check concerning the choices of participants in Conditions II, III, and IV small changes were made only in the highest payoffs for the third party. The highest payoff of the third party was decreased and increased with two dollars, for participants assigned to groups Ψ and Ω, respectively. Note that although these payoffs were changed, the third party still had the same ordinal preferences over the outcomes (i.e. preferring defection in Condition III, and preferring cooperation in Condition IV).

Results

We conduct logistic regression analyses to investigate whether the participants' choices differed across experimental conditions (and thus whether our main hypothesis is confirmed), and to test the predictions derived from the motivational principles. We take the proportion of cooperative choices as the dependent variable. Dummies of the four conditions are independent variables and we control for gender. Subsequently, we examine the robustness of subjects' behavior in Conditions II, III, and IV by comparing their first decision with their second (hypothetical) decision by means of paired samples t-tests.

Descriptive statistics: Cooperation across conditions. On average across all four conditions, participants made the cooperative choice 29.0% of the time. In Condition I, 30.2% of the participants cooperated. In Condition II this was 38.6%, in Condition III 18.2%, and in Condition IV 28.9%. With respect to the three conditions containing externalities we thus observe $\pi_{II} = 0.386 > \pi_{IV} = 0.289 > \pi_{III} = 0.182$. The differences in cooperation across conditions suggest support for our main hypothesis. Concerning the motivational principles, this ordering partly contradicts and partly confirms

Table 3. Logistic regression coefficient estimates for experimental conditions, with cooperation (= 1; defection = 0) as the dependent variable and Condition I, II and III as reference categories respectively.

N = 176	Reference category		
	Condition I Coefficient (S.E.)	Condition II Coefficient (S.E.)	Condition III Coefficient (S.E.)
Condition I	n/a		
Condition II	0.398 (0.457)	n/a	
Condition III	−0.714 (0.517)	−1.112* (0.505)	n/a
Condition IV	−0.145 (0.475)	−0.544 (0.464)	0.569 (0.513)
Gender	−0.391 (0.347)	−0.391 (0.347)	−0.391 (0.347)

* $p = 0.028$ (two-tailed)

the orders predicted by both the altruistic ($\pi_{IV} \geq \pi_{II} \geq \pi_{III}$) and the competitive ($\pi_{III} \geq \pi_{II} \geq \pi_{IV}$) principle. In addition, it corresponds to the order predicted by the inequality aversion principle ($\pi_{II} \geq \pi_{IV}$). Below, we examine the statistical significance of these observations.

Comparing conditions. Taking Condition I as the reference category in the logistic regression reveals that Condition II, Condition III, and Condition IV do not differ from Condition I (coefficient = −0.714, S.E. = 0.517, p = 0.167, coefficient = −0.145, S.E. = 0.475, p = 0.760, and coefficient = 0.398, S.E. = 0.457, p = 0.383, respectively; see Table 3). The effect of gender is not significant in any of the analyses. Thus, although the descriptive statistics suggested otherwise, we must refute our main hypothesis.

To more closely evaluate the predictions of the motivational principles and compare the conditions that contain third parties, we subsequently take Conditions II and III as reference categories (see Table 3).

Conditions II and III differ significantly from each other. When the third party benefits from defection, participants defect more often than when this party is indifferent ($\pi_{II} \geq \pi_{III}$; coefficient = −1.112, S.E. = 0.505; p = 0.028). This finding supports one part of the order predicted by the altruistic motivational principle.

Conditions II and IV do not differ significantly. When the third party benefits from cooperation participants cooperate as much as when this party is indifferent ($\pi_{II} = \pi_{IV}$; coefficient = −0.544, S.E. = 0.464; p = 0.241). Hence, we find no significant support for this part of the order predicted by either the competitive, altruistic, or inequality averse motivational principle.

Conditions III and IV do not differ significantly ($\pi_{III} = \pi_{IV}$; coefficient = −0.569, S.E. = 0.513, p = 0.268). This finding invalidates the orders predicted by the competitive as well as the altruistic motivational principle. Thus, though part of the observed order is in concord with both the altruistic and the competitive motivational principle (i.e. $(\pi_{II} \geq \pi_{III})$ and $(\pi_{IV} \geq \pi_{III})$ for altruism; $(\pi_{II} \geq \pi_{IV})$ for competition), only one part of the altruistic motivational principle (i.e. $(\pi_{II} \geq \pi_{III})$) finds significant support in the data.

Robustness of decisions. In Conditions II, III, and IV participants made a second (hypothetical) choice between cooperation and defection, given an altered payoff matrix. Over all three conditions 21 subjects (15.8%) changed their decisions: 14 participants changed to cooperation, 7 changed to defection. There are no significant differences between the first and second choices for any group.

Discussion: Study 1

Based on the results of Study 1 we have to reject our main hypothesis that externalities affect behavior. The data do suggest that making participants defect more by including a third party that benefits from defection is easier than making participants cooperate more by including a third party that benefits from cooperation.

This finding may result from the truly 'social' dilemma that participants face when a third party benefits from defection. They have to make a choice between benefiting the third party or benefiting the other actor. Previous research shows that people are more inclined to defect when defection can be interpreted as a fair move (e.g. Pillutla and Murnighan, 1995; Van Dijk *et al.,* 2004). Since the interest of the third coincides with participants' self-interest, the third may function as an excuse to defect.

The results of Study 1 contradict expectations based on social preference considerations. Since the effects of such preferences are well-documented in the behavioral economics and social science literatures, and since the general explanatory power of our statistical model is low (R^2_L = 0.03),[2] we propose that important variables have been omitted in Study 1. In particular, we made no attempt to measure individual social preferences in Study 1.

Aggregate behavior may not be adequately modeled by just one motivational principle. This conclusion is supported by the large body of research on social value orientation (e.g. Simpson, 2004; Van Lange, 1992, 1999). Therefore we conduct Study 2, in which we measure participants' motivations using the Triple-Dominance Scale of Social Values. In addition, Study

2 uses a within-subjects design to better control for individual differences in social values or preferences.

Study 2

In social psychology the concept of *social value orientation* (SVO) has been developed to capture individuals' preferences for particular distributions of payoffs for self and others (e.g., Van Lange *et al.,* 1997). Typically, three orientations are distinguished: (i) the prosocial orientation, characteristic of individuals (*prosocials*) seeking to enhance joint outcomes and equality, (ii) the individualistic orientation, characteristic of individuals (*individualists*) seeking to enhance their own outcomes regardless of the outcomes for others, and (iii) the competitive orientation, characteristic of individuals (*competitors*) seeking to increase their relative payoff advantage over others (e.g., Van Lange *et al.,* 2007).

A large number of studies have shown that SVO is predictive of cooperative behavior in social dilemmas (see Balliet *et al.* (2009) and Bogaert *et al.* (2008) for recent overviews). Moreover, the types distinguished in the SVO concept map onto our motivational principles in the following way. The prosocial orientation encompasses inequality aversion and altruism, and the individualist and competitive orientations map directly onto egoism and competitiveness, respectively.

Thus, we propose that Study 1 showed no effects of our experimental conditions because no single motivational principle was dominant among participants. We expect however that if we include a measurement of participants' social preferences in terms of SVO, different SVO types will be found to react differently to our conditions. In particular, we expect different SVO types to behave according to the predictions of the motivational principles to which they are associated.

Methods

Participants and materials. 124 undergraduate students from the University of Groningen in The Netherlands participated in the experiment of Study 2. By participating, students were granted a course credit point. Participants in Study 2 earned no additional money and all payoffs were thus hypothetical. Like Study 1, the experiment in Study 2 was conducted with paper and pen.

In Study 2, participants played the same four experimental games of Study 1 and in addition completed the Triple Dominance Measure of Social Values. In both tasks participants were asked to imagine that the points were

valuable to themselves and the others involved. The tasks of Study 2 were part of a larger research project in which several researchers asked participants to complete questionnaires. The Triple Dominance Measure of Social Values and the PD game decisions were not adjacent in the complete set of tasks and were printed in a different font and provided with distinct cover sheets so as to dissuade participants from thinking that the tasks were related.

The Triple-Dominance Measure employs the 'decomposed game' approach, in which the participant is given 9 choices between three distributions of valuable 'points' across self and other. At each choice, each of the three distributions is associated with one of the social value orientations. The first choice in the Triple-Dominance Measure is for instance between the following three distributions:

SVO Choice 1

Distribution A: 480 points for self and 80 points for other
Distribution B: 540 points for self and 280 points for other
Distribution C: 480 points for self and 480 points for other

In SVO Choice 1, distribution A is the competitive choice (maximizing the relative advantage of self over other), distribution B is the individualistic choice (maximizing the absolute payoff for self), and distribution C is the prosocial choice (equal payoffs for self and other with maximum total number of points). For a description of all 9 choices see for instance Van Lange *et al.* (2007).

Procedure

The procedure of Study 2 differed in a number of respects from that of Study 1. All participants completed both experimental tasks sitting together in an ordinary lecture room. Seating was such that decisions were private. All participants completed the Triple-Dominance Measure first and made their PD decisions second.

For the Triple-Dominance Measure we used the standard instructions of this instrument. The instructions for the PD decisions were identical to those in Study 1, except for the fact that in Study 2 participants were asked to imagine that points in the game were valuable for all individuals involved. After reading the PD instructions, participants answered a quiz to verify their understanding of the rules of the experimental games, after which they made their four decisions. Contrary to Study 1, participants' age and gender was not registered.

Design

Study 2 had a within-subjects design: each participant made a decision in all four PD conditions described in Study 1. Condition I (i.e., the standard PD) was always the first decision made. The order of the other three conditions was balanced, yielding 6 different orders to which participants were randomly assigned.

Results

Since we have a repeated measures design with a dichotomous dependent variable (i.e., the decisions to cooperate or defect), we used (random intercept) multilevel logistic regression to analyze the data. Participants' decisions in the PD conditions constitute level 1, and participants constitute level 2. The experimental conditions are thus level-1 variables and participants' social value orientations as measured with the Triple-Dominance Measure are level-2 variables.

We use Markov Chain Monte Carlo (MCMC) estimation, since this procedure yields less biased estimates than standard approximate estimation techniques. In addition, MCMC estimation yields the deviance statistic, which allows us to compare nested statistical models. The difference in the deviance statistics of two nested models approximately follows a chi-square distribution with degrees of freedom equal to the difference in the number of parameters of the models compared. Significant values indicate that the model with the lower deviance is an improvement over the model with the higher deviance.

Since the payoffs associated with the decisions in Study 2 were hypothetical we were especially concerned with the quality of our data. Therefore, we excluded all participants who had made any errors or had any missing values on the 5 quiz items concerning the instructions. This leaves 73 participants with a perfectly answered quiz. Significantly, none of these participants has any missing values on any of the other experimental tasks.

Descriptive statistics. On average across all four conditions in Study 2, participants made the cooperative choice 32.9% of the time. In Condition I, 35.6% of the participants cooperated. In Condition II this was 36.1%, in Condition III 26.7%, and in Condition IV 33.3%. With respect to the three conditions containing externalities we thus observe the same pattern as in Study 1. $\pi_{II} = 0.361 > \pi_{IV} = 0.333 > \pi_{III} = 0.267$. Again this ordering partly contradicts and partly confirms the orders predicted by both the altruistic ($\pi_{IV} \geq \pi_{II} \geq \pi_{III}$) and the competitive ($\pi_{III} \geq \pi_{II} \geq \pi_{IV}$) principle. It also

corresponds to the order predicted by the inequality aversion principle ($\pi_{II} \geq \pi_{IV}$).

To measure social value orientation typically participants are classified as prosocials, individualists or competitors when they make at least six out of nine choices in the Triple-Dominance Measure consistent with one of the orientations (e.g., Van Lange and Kuhlman, 1994). However, this implies that part of the participants cannot be classified, leading to loss of data. Also, the strict classification in types suppresses variance in the responses that might explain behavior. Finally, classifying participants into distinct types typically leads to overestimation of the presence of the type belonging to the most popular response, which is the prosocial answer. Therefore, we computed two variables labeled Prosocial and Competitive, to count the numbers of times (out of 9) that participants made the prosocial and competitive choices, respectively. Note how (9 – Prosocial – Competitive) equals the number of individualistic choices. Thus, the number of individualistic choices constitutes the 'reference category' for Prosocial and Competitive. The averages of Prosocial and Competitive in Study 2 are 5.60 (S.D. = 3.66) and 0.70 (S.D. = 1.67), respectively.

Comparing conditions. We first computed 5 dummy variables to model the six different orders in which the conditions containing externalities were played. This revealed that there is no effect of the order of conditions on the proportion of cooperation. We thus excluded the order variables from further analyses.

To evaluate our hypothesis and predictions, we first estimated a multilevel logistic regression model with only the experimental conditions as independent variables. In this model, Condition I was the reference category and the other three conditions were represented by dummies. Comparison of this model with the empty model that contains no independent variables reveals that adding the experimental conditions does not improve the fit (Difference in deviances = 5.83, df = 3, $p = 0.12$). Thus, as in Study 1, we conclude that the proportions of cooperation do not differ across conditions and we must reject our main hypothesis.

In the next model that we estimated we added the level-2 variables Prosocial and Competitive. This again did not significantly improve the model (Difference in deviances = 1.07, df = 2, $p = 0.78$). Thus, we conclude that there is no main effect of social value orientation across all conditions.

Finally, we added the interactions of Prosocial and Competitive with the experimental conditions, still having Condition I as the reference category. This yields a model with 6 interaction terms that is a large and significant

Table 4. MCMC estimates of random intercept multilevel logistic regression coefficients for the interactions of experimental conditions with Prosocial and Competitive, with cooperation (= 1; defection = 0) as the dependent variable. Reference category is Condition I.

N = 292	Coefficient (standard error)
Prosocial X	
Condition II	0.190 (0.284)
Condition III	−0.455 (0.324)
Condition IV	0.538* (0.253)
Competitive X	
Condition II	−3.225 (1.770)
Condition III	−4.614* (2.282)
Condition IV	−0.662 (0.984)

* Absolute value of (Coefficient/S.E.) larger than 2.

improvement over the previous model with interactions (Difference in deviances = 18.87, df = 6, $p = 0.004$). All main effects of experimental conditions and social value orientation are still statistically insignificant, so the improvement is wholly due to the added interactions. Table 4 contains the estimated coefficients for the interaction terms. The significance of separate coefficients is evaluated by dividing the coefficient by its standard error and comparing this score to a standard normal distribution. We label a coefficient significant when the absolute value of the ratio of the coefficient over the standard error exceeds 2.

Table 4 shows that prosocial responses to the Triple-Dominance Scale are associated with significantly more cooperation in Condition IV, compared to Condition I. Thus, as we move from Condition I to Condition IV participants with a higher score on Prosocial increase their cooperation more, than do participants with a lower score on Prosocial. No such effect exists for Competitive. Competitive responses to the Triple-Dominance Scale however, are associated with significantly less cooperation in Condition III, compared to Condition I. Thus, as we move from Condition I to Condition III participants with a higher score on Competitive decrease their cooperation more, than do participants with a lower score on Competitive. A similar tendency exists for Prosocial, but this effect does not reach statistical significance. Note that, since the conditions have no main effect, participants with a tendency to give the individualist response to the Triple-Dominance Scale (i.e., participants with low scores on both Prosocial and Competitive) do not significantly change their behavior when moving from Condition I to any of the other conditions.

We ran similar analyses with Condition II as the reference category. Again, the model including the interactions of experimental conditions with Prosocial and Competitive was the best model and provided a significant improvement over the model with only main effects of conditions and Prosocial and Competitive (Difference in Deviances = 20.263 df = 6, $p = 0.002$). When evaluating the significance of individual interaction effects, the only significant effect is associated with the interaction between Condition III and Prosocial (Coefficient = −0.618, S.E. = 0.301). Thus, when moving from Condition II to Condition III in this model, participants with a higher score on Prosocial decrease their cooperation more than do participants with a lower score on Prosocial. The interactions between Condition IV and Prosocial, and between Condition III and Competitive had the same signs as in the previous analyses (Coefficient = 0.339, S.E. = 0.24, and Coefficient = −1.151, S.E. = 2.415), but did not reach significance. The interaction between Condition IV and Competitive was positive but also insignificant (Coefficient = 2.379, S.E. = 1.726). This model also yielded a main effect of Prosocial (Coefficient = 0.534, S.E. = 0.223), implying that participants with a higher score on Prosocial generally cooperate more.

Finally, we also ran an analysis with Condition III as the reference category. As before the model including the interactions of experimental conditions with Prosocial and Competitive was the best model, and provided a significant improvement over the model with only main effects of conditions and Prosocial and Competitive (Difference in deviances = 20.521 df = 6, $p = 0.002$). The interactions between Prosocial and Condition IV (Coefficient = 0.932, S.E. = 0.294), and between Competitive and Condition I (Coefficient = 3.868, S.E. = 1.754) were significant in this model. Thus, when moving from Condition III to Condition IV in this model, participants with a higher score on Prosocial increase their cooperation more than do participants with a lower score on Prosocial. Similarly, when moving from Condition I to Condition III participants with a higher score on Competitive decrease their cooperation more than do participants with a lower score on Competitive.

Interestingly, this model also yields significant main effects of Competitive (Coefficient = −4.036, S.E. = 1.772) and Condition IV (Coefficient = −6.055, S.E. = 2.324). The former effect is in the expected direction (competitive participants cooperating less). The latter effect implies that participants with a tendency to give the individualist response to the Triple-Dominance Scale (i.e., participants with low scores on both Prosocial and Competitive) significantly *decrease* their cooperation when moving from Condition III to Condition IV.

Discussion: Study 2

The results of Study 2 confirm the results of Study 1 in that there is no main effect of the inclusion of externalities for a third party on the proportion of cooperation in the PD. However, inclusion of a measurement of participants' social value orientation showed that participants' responses in the different experimental conditions are moderated by their social value orientation. Thus, while *aggregated* behavior is not affected by externalities, the behavior of *individual actors* is. This result provides partial support for our main hypothesis.

The data of Study 2 suggest that a more prosocial value orientation is associated with a response pattern that matches the altruism prediction derived in the theory section: $\pi_{IV} \geq \pi_{II} \geq \pi_{III}$. The second inequality ($\pi_{II} \geq \pi_{III}$) was found to be strict and significant in the analyses with Condition II as reference category. Interestingly, a more competitive value orientation is associated with the same response pattern, although none of the effects are then significant. In the first two models we estimated an individualistic social value orientation was associated with no differences in behavior across the conditions. This is in line with an interpretation of individualists behaving as predicted by the rational egoist motivational principle. In the last model (with Condition III as reference category) however, an individualistic orientation was associated with the prediction of the *competitive* motivational principle.

Considering the main effects of SVO, we found (in the model with Condition II as reference category) that a prosocial orientation is associated with more cooperation and (in the model with Condition III as reference category) that a competitive orientation is associated with less cooperation.

Comparing the results of Study 1 and the analyses reported in Table 4 reveals a very informative result. Study 1 already suggested that making participants defect more by including a third party that benefits from defection is easier than making participants cooperate more by including a third party that benefits from cooperation. Table 4 shows how this depends on an individual's social value orientation.

On the one hand, individuals with a more prosocial value orientation are persuaded to cooperate more when a third party benefits from cooperation. They do not defect more, however, when the third party benefits from defection. Individuals with a more competitive value orientation on the other hand, are persuaded to defect more in this latter instance. They, however, do not cooperate more when the third party benefits from cooperation.

General discussion

Social dilemmas, although identical in their payoff structures for actors, may greatly differ in their externalities for third parties. These differences affect behavior of actors. But not every actor is affected in the same way. Our main results are that prosocial actors can be instigated to cooperate more than in a standard one-shot, 2-person prisoner's dilemma game (the 'standard PD') when a third party benefits from cooperation, but will not defect more when a third party benefits from defection. Competitive actors, on the other hand, will defect more than in a standard PD when a third party benefits from defection, but cannot be persuaded to cooperate more when a third benefits from cooperation. Consequently, we assert that *both* externalities for third parties *and* social value orientation of actors have to be taken into account when predicting behavior in real-life social dilemmas.

In the experimental game used in both of our studies, we examined the effects of the presence of a third party that was (i) indifferent with respect to the dilemma outcome, (ii) benefited by cooperation, and (iii) benefited by defection. Predictions concerning the expected rate of cooperation were derived from four motivational principles: egoism, altruism, inequality aversion, and competition. To test the predictions, we incorporated externalities into a standard PD with positive payoffs. The payoff structure for the actors was identical across the four games, the externalities for the third party were not. While in Study 1 we observed some differences in cooperation across conditions, no motivational principle could account for all our results in a satisfactory manner. The observed ordering of conditions in amount of cooperation simultaneously corresponded to parts of the orderings predicted by the altruistic, inequality averse, and competitive motivation. Only part of the altruistic motivation was supported significantly: subjects defected more when a third party benefited from defection than when this party was indifferent. However, since altruism cannot be switched on and off, concluding that altruism 'partly holds' on the basis of partial support for the altruism principle would be incorrect.[3]

The data did suggest that it is easier to make people defect than to make them cooperate by including a third party that benefits from defection or cooperation respectively. In all our PD games, except for the one where a third party benefits from defection, it is evident from the structure of the dilemma what the most social choice is: cooperation. But when a third benefits from defection, no single choice is definitively social. We therefore propose that this finding can be partly explained by the lack of a *dominant social strategy* when a third benefits from defection.

The results from Study 2 suggested that behavior in identical social dilemmas with different externalities is moderated by social value orientation. A prosocial orientation leads people to cooperate more to benefit a third party. Compared to their cooperation rate in a standard PD, prosocials are not easily enticed to defect more on the other actor when a third party benefits from defection. Competitors show the complementary pattern: compared to their cooperation rate in a standard PD they will defect more often when a third party is benefited, but will not cooperate more when the third party profits from cooperation. Finally, compared to their cooperation rate in a standard PD, individualists show no tendency to change their behavior in response to the addition of a third party receiving externalities. Study 2 also showed that in our experimental setting, *aggregated* cooperation rates are not influenced by the presence of third parties, but individual cooperation is. Our main hypothesis is therefore partially supported.

Study 2 has interesting implications for social dilemmas outside of the lab. A social dilemma containing externalities with competitive business leaders as actors may have a very different outcome than the exact same dilemma among prosocial commune members. The differences in cooperation between these actors may partly be attributed to attitudinal differences (as we found main effects of SVO in two of the three models we estimated in Study 2). However, all three models that we estimated in Study 2 suggest that these differences in cooperation are also caused by an interaction between SVO and externalities.

Additionally, previous research indicates that groups are generally more antagonistic than individuals (i.e., the interindividual–intergroup discontinuity – Insko *et al.,* 1988*,* 1994; Schopler *et al.,* 2001; Wildschut *et al.,* 2003). Insko *et al.* (1994), for instance, found that groups in public goods games and prisoner's dilemmas act much more competitively than individuals. As we found an interaction between competitiveness and externalities for defection, it thus seems likely that defection would proliferate in an environment *with positive externalities for defection* where groups, rather than individuals, make the decisions.

Our studies pose questions for future research. We did, for instance, not look at negative externalities. It seems nevertheless likely that negative externalities will elicit different behavior from actors than positive externalities. Van Beest *et al.* (2005) for instance, find that people do not mind taking gains for themselves, but that they dislike imposing losses on others. In case of externalities, this would imply that people are more reluctant to cooperate or defect when a third party is damaged by this move.

Additionally, we found that one single motivational principle cannot explain behavior in dilemmas with externalities. Behavior could be explained,

however, by taking into account people's different social values. SVO nevertheless assumes that people have the same social values towards all individuals. People may nevertheless hold different attitudes towards different others (i.e. actors and thirds). Attitudes towards thirds and actors should be investigated in future studies.

This paper has taken the first step in investigating the influence of externalities on behavior in social dilemmas. Whereas individualists seem largely unaffected, third parties can get the best out of prosocials and the worst out of competitors. Our paper shows that externalities are able to fundamentally affect individual behavior.

Appendix: Mathematical derivation of predictions of motivational principles

Let $V_{S,E}$ be an actor's expected utility of playing strategy S in Condition E, let $u(x)$ denote an actor's utility for payoff vector x, and let p_E be the subjective probability (i.e. the actor's assessment) that the other actor cooperates in Condition E. In a payoff vector the first and second dollar amounts indicate the payoffs for actors A and B, respectively. The third dollar amount indicates the externality for the third party. Then the following holds:

$$\begin{aligned}
V_{C,II} &= p_{II}u(\$3,\$3,\$2) + (1-p_{II})u(\$0,\$5,\$2)\\
V_{D,II} &= p_{II}u(\$5,\$0,\$2) + (1-p_{II})u(\$1,\$1,\$2)\\
V_{C,III} &= p_{III}u(\$3,\$3,\$0) + (1-p_{III})u(\$0,\$5,\$2)\\
V_{D,III} &= p_{III}u(\$5,\$0,\$2) + (1-p_{III})u(\$1,\$1,\$5)\\
V_{C,IV} &= p_{IV}u(\$3,\$3,\$5) + (1-p_{IV})u(\$0,\$5,\$2)\\
V_{D,IV} &= p_{IV}u(\$5,\$0,\$2) + (1-p_{IV})u(\$1,\$1,\$0)\,.
\end{aligned} \tag{1}$$

Let π_E be the predicted proportion of cooperative choices in Condition E. Since egoism is not affected by other actors' payoffs, the *egoism prediction* is: $\pi_{II} = \pi_{III} = \pi_{IV}$. To derive predictions for the other motivational principles we make the auxiliary assumption that $p_{II} = p_{III} = p_{IV}$ (i.e. an actor's subjective probability of the other actor cooperating is constant across conditions).

This way, we do not make additional assumptions about the shape of actors' utility functions. Varying subjective probabilities would prevent us from testing motivational principles more generally.

The altruism assumption A3.1 now yields $V_{D,III} \geq V_{D,II}$, $V_{C,II} \geq V_{C,III}$, $V_{D,II} \geq V_{D,IV}$ and $V_{C,IV} \geq V_{C,II}$. Thus, any actor defecting in Condition II

($V_{D,II} \geq V_{C,II}$) defects in Condition III ($V_{D,III} \geq V_{C,III}$), and any actor cooperating in Condition III, cooperates in Condition II. Furthermore, any actor cooperating in Condition II ($V_{C,II} \geq V_{D,II}$) cooperates in Condition IV ($V_{C,IV} \geq V_{D,IV}$), and any actor defecting in Condition IV defects in Condition II. This yields the *altruism prediction*: $\pi_{IV} \geq \pi_{II} \geq \pi_{III}$.

In a similar vein, inequality aversion assumption A3.2 yields $V_{C,II} \geq V_{C,III}$, $V_{D,II} \geq V_{D,III}$, which implies that actors' choices in Condition II do not imply certain choices in Condition III according to this motivational principle. Therefore, we cannot derive a prediction for the comparison between these conditions for the inequality aversion principle. The fact that the inequality aversion principle does not impose a complete ordering of our experimental conditions does not imply that this principle is in some sense 'irrational'. It simply means that in our particular experimental game, the principle allows both possible orderings of Condition II and Condition III.

The inequality aversion principle also yields $V_{C,II} \geq V_{C,IV}$ and $V_{D,II} = V_{D,IV}$, implying $\pi_{II} \geq \pi_{IV}$, which is the *inequality aversion prediction*. Finally, an analogous analysis shows that the *competition prediction* derivable from assumption A3.3 is $\pi_{III} \geq \pi_{II} \geq \pi_{IV}$.

Notes

1 The phrase ceteris paribus indicates that the assumptions are applicable to the comparison of outcomes (i.e. payoff vectors) that are different with respect to the payoffs of one individual, but are otherwise equal.

2 The formula for R^2_L is $R^2_L = 1 - \frac{\ln(L_M)}{\ln(L_0)}$, where L_M is the value of the likelihood function for the model with all predictors included and L_0 is the corresponding value for the model with the intercept only (Allen and Le, 2008). We use R^2_L as a measure for overall effect size of the model because it is relatively invariant to the base rate. R^2_L can be interpreted as a proportional reduction in error measure, analogous to the regular R-square.

3 Although people could still be altruistic if they expected their opponents to defect more in Condition IV and were thus more afraid to cooperate, this explanation seems unlikely. Most people expect others to behave like themselves in dilemma games (Kanazawa, 2007; Kuhlman and Wimberley, 1976; Ross *et al*., 1977; Van Lange, 1992). People are often too optimistic about behavior of others (Offerman *et al*., 1996), and prosocials are relatively insensitive to expected bad behavior of their partners (Offerman *et al.,* 1996).

References

Allen J and Le H (2008) An additional measure of overall effect size for logistic regression models. *Journal of Educational and Behavioral Statistics* 33: 416–441.

Axelrod R (1984) *The Evolution of Cooperation.* New York: Basic Books.

Balliet D, Parks C, and Joireman J (2009) Social value orientation and cooperation in social dilemmas: A meta-analysis. *Group Processes and Intergroup Relations* 12: 533–547.

Bogaert S, Boone C, and Declerck C (2008) Social value orientation and cooperation in social dilemmas: A review and conceptual model. *British Journal of Social Psychology* 47: 453–480.

Bolton GE and Ockenfels A (2000) ERC: A theory of equity, reciprocity and competition. *American Economic Review* 90: 166–193.

Dawes CT, Fowler JH, Johnson T, McElreath R, and Smirnov O (2007) Egalitarian motives in humans. *Nature* 446: 794–796.

Dawes RM (1980) Social dilemmas. *Annual Review of Psychology* 31: 169–193.

Eek D and Gärling T (2006) Prosocials prefer equal outcomes to maximizing joint outcomes. *British Journal of Social Psychology* 45: 321–337.

Fehr E and Fischbacher U (2002) Why social preferences matter: The impact of non-selfish motives on competition, cooperation and incentives. *The Economic Journal* 112: C1–C33.

Fehr E and Schmidt KM (1999) A theory of fairness, competition, and cooperation. *The Quarterly Journal of Economics* 114: 817–868.

Frank RH (1985) *Choosing the Right Pond: Human Behavior and the Quest for Status*. Oxford: Oxford University Press.

Gintis H, Bowles S, Boyd R, and Fehr E (2003) Explaining altruistic behavior in humans. *Evolution and Human Behavior* 24: 153–172.

Insko CA, Hoyle RH, Pinkley RL, *et al.* (1988) Individual–group discontinuity: The role of a consensus rule. *Journal of Experimental Social Psychology* 24: 505–519.

Insko CA, Schopler J, Graetz KA, *et al.* (1994) Individual–intergroup discontinuity in the Prisoner's Dilemma Game. *Journal of Conflict Resolution* 38: 87–116.

Kahneman D and Tversky A (1979) Prospect theory: Analysis of decision under risk. *Econometrica* 47: 263–291.

Kanazawa Y (2007. The promotion and evolution of cooperation through projection: Implications for social dilemmas and trust. *Journal of Mathematical Sociology* 31: 187–204.

Kollock P (1998a) Transforming social dilemmas: Group identity and cooperation. In Danielson, P. (ed), *Modeling Rational and Moral Agents*, pp. 186–210. Oxford: Oxford University Press.

Kollock P (1998b) Social dilemmas: The anatomy of cooperation. *Annual Review of Sociology* 24: 183–214.

Kuhlman MD and Wimberley DL (1976) Expectations of choice behavior held by cooperators, competitors, and individualists across four classes of experimental games. *Journal of Personality and Social Psychology* 34: 69–81.

Levine DK (1998) Modeling altruism and spitefulness in experiments. *Review of Economic Dynamics* 1: 593–622.

Offerman T, Sonnemans J, and Schram A (1996) Value orientations, expectations and voluntary contributions in public goods. *The Economic Journal* 106: 817–845.

Ostrom E, Walker J, and Gardner R (1992) Covenants with and without a sword: Self-governance is possible. *American Political Science Review* 86: 404–417.

Pillutla MM and Murnighan KJ (1995) Being fair or appearing fair: Strategic behavior in ultimatum bargaining. *The Academy of Management Journal* 38: 1408–1426.

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

Ross L, Greene D, and House P (1977) False consensus effect: Egocentric bias in social-perception and attribution processes. *Journal of Experimental Social Psychology* 13: 279–301.

Schopler J, Insko CA, Wieselquist J, *et al*. (2001). When groups are more competitive than individuals: The domain of the discontinuity effect. *Journal of Personality and Social Psychology* 80: 632–644.

Simpson B (2004) Social values, subjective transformations, and cooperation in social dilemmas. *Social Psychology Quarterly* 67: 385–395.

Simpson B and Willer R (2008) Altruism and indirect reciprocity: The interaction of person and situation in prosocial behavior. *Social Psychology Quarterly* 71: 37–52.

Van Beest I, Van Dijk E, De Dreu CKW, and Wilke HAM (2005) Do-no-harm in coalition formation: Why losses inhibit exclusion and promote fairness cognitions. *Journal of Experimental Social Psychology* 41: 609–617.

Van Dijk E, De Cremer D, and Handgraaf MJJ (2004) Social value orientations and the strategic use of fairness in ultimatum bargaining. *Journal of Experimental Social Psychology* 40: 697–707.

Van Lange PAM (1992) Confidence in expectations: A test of the triangle hypothesis. *European Journal of Personality* 6: 371–379.

Van Lange PAM (1999) The pursuit of joint outcomes and equality in outcomes: An integrative model of social value orientation. *Journal of Personality and Social Psychology* 77: 337–347.

Van Lange PAM and Kuhlman DM (1994) Social value orientation and impressions of partner's honesty and intelligence: A test of the might versus morality effect. *Journal of Personality and Social Psychology* 67: 126–144.

Van Lange PAM, Otten W, De Bruin EMN, and Joireman JA (1997) Development of prosocial, individualistic, and competitive orientations: Theory and preliminary evidence. *Journal of Personality and Social Psychology* 73: 733–746.

Van Lange PAM, Bekkers R, Schuyt TNM, and Van Vugt M (2007) From games to giving: Social value orientation predicts donations to noble causes. *Basic and Applied Social Psychology* 29: 375–384.

Wildschut T, Pinter B, Vevea JL, Insko CA, and Schopler J (2003) Beyond the group mind: A quantitative review of the interindividual–intergroup discontinuity effect. *Psychological Bulletin* 129: 698–722.

Yamagishi T (1986) The provision of a sanctioning system as a public good. *Journal of Personality and Social Psychology* 51: 110–116.

Yamagishi T (1995) Social dilemmas. In Cook KS, Fine GA, and House JS *Social Perspectives on Social Psychology,* pp. 311–335. Boston: Allyn and Bason.