Reciprocity as an Individual Difference

Kurt A. Ackermann[1], Jürgen Fleiß[2], and Ryan O. Murphy[1]

Abstract

There is accumulating evidence that decision makers (DMs) are sensitive to the distribution of resources among themselves and others, beyond what is expected from the predictions of narrow self-interest. These social preferences are typically conceptualized as being static and existing independently of information about the other people influenced by a DM's allocation choices. In this article, we consider the reactivity of a DM's social preferences in response to information about the intentions or past behavior of the person to be affected by the DM's allocation choices (i.e., how do social preferences change in relation to the other's type). This article offers a conceptual framework for characterizing the link between distributive preferences and reciprocity, and reports on experiments in which these two constructs are disentangled and the relation between the two is characterized.

Keywords

social value orientation (SVO), social preferences, reciprocity, individual differences, preference dynamics

It has been shown in many different studies that decision makers (DMs) generally take into account other peoples' payoffs when making decisions in situations of interdependence and thus typically do not behave according to the predictions of

[1]ETH Zürich, Chair of Decision Theory and Behavioral Game Theory, Zürich, Switzerland
[2]University of Graz, Department of Statistics and Operations Research, Graz, Austria

Corresponding Author:
Kurt A. Ackermann, ETH Zürich, Chair of Decision Theory and Behavioral Game Theory, Clausiusstrasse 50, 8092, Zürich, Switzerland.
Email: kurt.ackermann@gess.ethz.ch

neoclassic economic theory. For instance, there is a large body of evidence showing that people voluntarily forgo some of their own gains for the benefit (or detriment) of others in experimental games, even if the situation is one-shot and anonymous (for reviews, see, for instance, Andreoni, Harbaugh, and Vesterlund 2008; Camerer 2003; Camerer and Thaler 1995; Fehr and Fischbacher 2003). Behavior deviating from simple own payoff maximization has been attributed in part to motivations referred to as social preferences (a.k.a. other-regarding preferences, social value orientation [SVO], altruism, or welfare trade-off ratios). Primarily, social preferences have been understood in terms of simple distributive concerns, where a DM's utility is modeled as a joint function of the DM's own outcome, and also of others' outcomes (see Bolton and Ockenfels 2000; Fehr and Schmidt 1999; Grzelak 1982; Loewenstein, Bazerman, and Thompson 1989; Messick and Thorngate 1967; Radzicki 1976; Wyer 1969). That is, a DM with social preferences derives non-zero utility from benefiting or harming another person, even when information about this other person is absent, i.e., under complete anonymity and lack of information about any of this other person's characteristics, past actions, or type (e.g., is the other person altruistic, prosocial, individualistic, or competitive).

The measurement and investigation of this *unconditional* type of social preferences has a long history in psychology, where this kind of motivation is referred to as Social Value Orientation (for reviews of SVO, see, e.g., Au and Kwong 2004; Bogaert, Boone, and Declerck 2008; McClintock and Van Avermaet 1982; Murphy and Ackermann 2014b). A person's SVO can be understood as the general motivational goal a person has when allocating resources between herself and another person. Different types of SVOs are commonly considered and discussed in the literature. For instance, a person may be motivated to simply maximize his own payoff (individualism) as standard economic theory would predict, maximize the sum of all payoffs (prosociality), minimize the difference between payoffs (another kind of prosociality called inequality aversion [IA]), maximize the other's payoff indifferent to his own (altruism), or maximize his own relative gain (competitiveness), and so on. These archetypes are useful when first considering social preferences, but SVO is a continuous construct and can be defined generally by the weights a DM attaches to his or her own outcomes and to other person's outcomes. These kinds of social preferences can be represented by a joint utility function, that is, a utility function that involves separate terms corresponding to outcomes for the self, the other, and arithmetic combinations thereof.[1]

Empirical work has shown the reliability of SVO as an individual difference (e.g., Au and Kwong 2004; Murphy, Ackermann, and Handgraaf 2011); the distribution of SVOs across individuals has been estimated, revealing substantial reliable differences across individuals and a bimodal pattern of preferences as well. Moreover, SVO has been shown to be a valid predictor of choice behavior in strategic situations, such as social dilemmas (see Balliet, Parks, and Joireman 2009; Murphy and Ackermann 2014a). However, these examples use SVO as

a static construct, not attending to the systematic reactivity of preferences and specifically how DMs' SVOs change in response to information about other people's SVOs and behavior.

There is support for the notion that DMs do sometimes condition their preferences and choices on characteristics of their interaction partners in experimental situations. Early research on this issue has shown that people exhibit dramatically different SVOs depending on whether the recipient in an own-other resource allocation task is described as a ''friend'', an ''unknown'', or as an ''antagonist'' (Sawyer 1966). These results suggest that DMs' distributive preferences, with respect to how resources are allocated between themselves and another person, are sensitive to revealed characteristics of the corresponding other person. Moreover, it has been shown that people may change their distributive preferences depending on the (potential) interaction partners' observed past behavior (e.g., Murphy and Ackermann 2014a; Sonnemans, van Dijk, and van Winden 2006). That is, people may become less prosocial, or even hostile (for evidence on costly punishing behavior, see, e.g., Fehr and Gächter 2000a), after observing noncooperative behavior by their interaction partner (this pattern is sometimes referred to as *behavioral assimilation*, a term coined by Kelley and Stahelski 1970), or DMs may become (more) prosocial after observing a cooperative move from the interaction partner (e.g., Berg, Dickhaut, and McCabe 1995; Diekmann 2004). These two kinds of behavioral responses are often referred to as negative and positive reciprocity, respectively (see Fehr and Gächter 2000b). In addition to observed actions taken by the interaction partner, behavior can be shaped by more complicated expectations of, and beliefs about, the interaction partner. For example, Schubert and Lambsdorff (2014) performed ultimatum game experiments in which Palestinians where confronted with offers from either Palestinian or Israeli proposers. The experimental results showed that Palestinian responders indeed responded differently to Israelis, as shown by the significantly higher rejection rate as compared to the situation where they are matched with Palestinian proposers.

Results suggest that what people respond to when exhibiting positive or negative reciprocity is not merely the outcomes resulting from the interaction partners' behavior (i.e., its consequences), but the interaction partners' *intentions* as revealed by the interaction partners' behavior (see, e.g., Charness and Rabin 2002; Cox, Friedman, and Gjerstad 2007; Dufwenberg and Kirchsteiger 2004; Falk and Fischbacher 2006; Falk, Fehr, and Fischbacher 2008; Levine 1998; Rabin 1993; Stanca, Bruni, and Corazzini 2009). It is important to notice that—at least in the context of experimental games—what has been termed the interaction partners' ''intentions'' is simply a synonym for the interaction partners' ''unconditional social preferences'', their baseline SVO, or their type (e.g., altruistic, prosocial, individualistic, and competitive). That is, the basic idea behind intention-based reciprocity models has been that DMs' SVOs may change depending on the interaction partners' assumed or observed SVO. In other words, people condition their social preferences upon what they believe the other person is like.

We follow this line of reasoning and apply a definition of reciprocity similar to the one provided by Cox (2004, 263) to address these changes in SVO. Concretely, we define reciprocity as *the change in a decision maker's distributive social preferences (i.e., SVO) in response to information about the interaction partner, compared to the decision maker's unconditional distributive social preferences* that are expressed when information about the interaction partner is absent. It is important to note that this definition of reciprocity is a difference score: it is the difference between unconditional (i.e., baseline) SVO and conditional SVO. Operationally then, a researcher would need to measure social preferences twice on the same DM in order to gage the magnitude of reciprocity an individual exhibits. This reactivity, or the dynamics of social preferences, is what concerns us here in this article.

In this study, we investigate how information about an interaction partner's type—that is, the interaction partner's revealed SVO—changes people's distributive social preferences (the DMs' SVOs) in a non-strategic situation. This design allows us to disentangle strategic concerns (anticipated benefits or reprisals) and isolate the effect of changes in SVO. For a similar argument in the context of the sequential prisoner's dilemma, see Clark and Sefton (2001, 55) and Dufwenberg and Kirchsteiger (2004). Hence, we can measure reciprocity alone, without the potential confounding effects beliefs may have on a change in behavior.

There is already clear evidence that people exhibit reciprocity in one-shot situations[2] by showing a willingness to forgo their own gains in order to respond prosocially to observed prosocial behavior in a sequential dictator game where no subsequent interaction will take place (Diekmann 2004). Diekmann also showed that the degree of kindness in reciprocating depended on the degree of kindness of the behavior being reciprocated, at the aggregate level. Our experimental design allows us to qualitatively replicate these findings and extend them in several important ways. First, rather than use a between-subjects design, we employed a within-subjects design and elicited complete reciprocity profiles of individuals across a set of different others. Second, subjects made a series of decisions in dictator games with varying tradeoff slopes (i.e., a mix of constant-sum and non-constant sum dictator games with varying rates of marginal substitution), which allows a more fine-grained assessment of the willingness to pay for the benefit or detriment of the person whose previously observed behavior is being responded to. And third, our experimental design rules out the possibility that the previously observed behavior being responded to might be perceived as a mere strategic gambit. The subjects responding to previous dictator game decisions will be truthfully informed that the person who had made the dictator game decision was not informed that her decision would be revealed to the DM. This ensures that subjects are responding to—that is, the object of reciprocation is—the interaction partner's honest revealed social preferences. This means that the DM does not have to try to glean strategic considerations on behalf of the other, such as acting prosocially only for the sake of inducing positive reciprocity.

The second goal of this article is to investigate the extent to which SVO is related to reciprocity. Studies have previously shown that people with different SVOs follow different reasoning when engaged in experimental games. For example, Boone, Declerck, and Kiyonari (2010) found that cooperative behavior of subjects with an individualistic value orientation tends to depend on external incentives, while the cooperation of prosocial subjects tends to depend on trust. There exists a fair amount of evidence suggesting that people with prosocial distributive preferences are more likely to reciprocate a corresponding interaction partner's prosocial choices compared to people with individualistic (i.e., goal to maximize own gain) or competitive (i.e., goal to maximize relative gain) distributive preferences (e.g., De Cremer and Van Lange 2001; Kanagaretnam et al. 2009; Van Lange 1999; Van Lange and Semin-Goossens 1998), although the latter two SVO types have also been shown to reciprocate significantly under certain conditions as well (e.g., Parks and Rumble 2001; Sheldon 1999). There is also evidence from researchers applying questionnaire methods, which supports the conclusion of a positive relationship between SVO and reciprocity (Perugini and Gallucci 2001; Perugini et al. 2003), although results are somewhat mixed in these studies.

In any case, the existent literature regarding the relation between SVO and reciprocity leaves room for improvement. First, the measures that have been used to assess SVO in these studies yield categorical data, that is, data on the lowest scale level of measurement (Stevens 1946), which results in low statistical power due to restricted variance. Concretely, subjects have typically been categorized as prosocial, individualistic, or competitive, or even dichotomized as prosocial versus proself. This severely limits statistical power and may yield null results erroneously. Second, some of the cited studies asked subjects to make hypothetical choices rather than decisions with real consequences, which complicates the interpretation of results since a person's reported intention of how to behave in a particular situation is not necessarily consistent with that person's real behavior in the respective situation (see, e.g., Ajzen, Brown, and Carvajal 2004; Sheeran 2002). Third, and most important, no study so far has assessed reciprocity profiles at the individual level. That is to say that no study to our knowledge has assessed how individuals differing in SVO responded differently to a set of distinct previously observed choices from different others. This study addresses these three issues by (a) measuring SVO on a continuous scale by means of the SVO Slider Measure (Murphy, Ackermann, and Handgraaf 2011), (b) implementing complete incentive compatibility, and (c) requiring that each subject responds to a set of empirically observed previous decisions made by matched real interaction partners. Consequently, the data obtained by the current study allow for a fine-grained and comprehensive assessment of reciprocity as an individual difference variable and support a more detailed analysis of its relation with SVO compared to previous studies.

The conjecture that SVO is to some extent dependent on situational factors raises an important question, namely, whether SVO is a responsive motivational state rather than a purely stable personality trait, as has often been tacitly assumed

in previous SVO research. The answer to this question is fundamentally connected to a long-standing discussion in personality and social psychology research, namely, the person–situation debate. It has long been acknowledged in psychology that some personality factors can be conceptualized as both state and trait. The most prominent exemplars are anxiety (see, e.g., Catell 1966; Spielberger 2010) and anger (e.g., Spielberger et al. 1999), but other personality factors show substantial within-person variability as well (Fleeson 2001). The apparent contradiction that the same personality variable can be both stable over time and variable within a person has evoked a considerable and sometimes heated debate in psychology. Some researchers went so far as to deny the usefulness of a state-trait distinction (Allen and Potkay 1981), while others decidedly expressed a contrary opinion (e.g., Zuckerman 1983). However, many researchers appear to have adopted the position that the distinction between state and trait is not arbitrary, and that the two concepts are to some extent dynamically intertwined, and often dependent on each other (e.g., Chaplin, John, and Goldberg 1988; Endler and Kocovski 2001; Fridhandler 1986; Mischel and Shoda 1998; Steyer, Ferring, and Schmitt 1992). To quote Fleeson (2004, 83), ''The person-situation debate is coming to an end because both sides of the debate have turned out to be right.'' With respect to SVO, previous results show a high test–retest reliability of $r = .915$ over a one-week period (Murphy, Ackermann, and Handgraaf 2011). This may lead some to conclude that SVO is only a static trait, but we would not support this simplistic conclusion. As we show in the remainder of this article, there is substantial systematic variability in how people's social preferences change in regard to updated information about the other person, and we posit that it is worth paying attention to the situation (i.e., state) side of the SVO construct as well as the person (i.e., trait) side.

Method

In order to address the research questions outlined previously, we use an experimental setup that consists of two parts, A and B, where the data collected in experiment A serve as stimuli for experiment B. Concretely, experiment A allows us to conduct experiment B without using deception, such that subjects in both experiments A and B make real decisions with real monetary consequences, resulting in a fully incentive compatible experimental design. The procedure we employed is explained in the next sections.

Experiment A: Collecting Stimuli for Experiment B

Experiment A was conducted in terms of a paper-pen choice task, where a total of 148 subjects from various disciplines made just a single decision, namely, to choose one out of four options (A, B, C, or D) of how to distribute money between themselves and a mutually anonymous other person. These four distribution

Table 1. Distribution Options in the Pen and Paper Task with Corresponding Choice Frequencies.

SVO Type	Decision options			Choice frequency	
	Option label	Payoff for			
		Self	Other	#	Percent
Competitive	A	85	15	6	4.1
Individualistic	B	100	50	25	16.9
Prosocial	C	85	85	109	73.6
Altruistic	D	50	100	8	5.4

options are shown in Table 1. Each option dominates[3] the other three options with respect to a particular motivational goal, namely the goal to maximize relative gain (option A dominates), the goal to maximize the own payoff (option B dominates), the goal to maximize the sum of payoffs (option C dominates), and the goal to maximize the other's payoff (option D dominates). These four goals represent four prototypical SVOs: competitiveness, individualism, prosociality, and altruism, respectively. Hence, when a person chooses one of the four options, her "type" is revealed and her primary social preference can be identified. The purpose of experiment A was to obtain choice results that would serve as stimuli for experiment B. The reason for this two-part design is to avoid using deception as a research practice and moreover to maintain incentive compatibility for the participants.

Data collection for experiment A was as follows: the experimenter was introduced to the subjects by a university instructor who then announced that a brief decision-making task would be handed out. The experimenter then thanked the group for their willingness to participate in the experiment and told the subjects that participation is voluntary and participants would be compensated. Each subject was then handed the decision sheet with the four distribution options and an envelope. Furthermore, subjects were informed that decisions of the same kind would be made by other people in the future and that they will be randomly matched in pairs with one of these future DMs, such that final payoffs would be determined. Importantly, the subjects were not informed, however, that their decisions would be revealed to the future DMs they could then be matched with. After all the subjects had made their decisions and put their marked decision sheet in their envelope, the experimenter collected the envelopes. Also an e-mail list was distributed, so that subjects could be contacted for payment once the future DMs had made their choices. Concretely, the subjects' student ID numbers were used to match subjects with their corresponding payment that they then received once experiment B (see the following section) was completed. On average, the

pen-and-paper task took ten minutes to complete and subjects earned an average amount of 4.1 euros (min = 2.5, max = 4.8). The choices subjects made in experiment A are reported in Table 1. To be clear again, experiment A was conducted for the purpose of generating stimuli for experiment B.

Experiment B: Assessing Reciprocity Profiles

Experiment B was carried out over twelve experimental sessions with a total of 148 subjects (same sample size as in experiment A, but different subjects) conducted at the Max Jung laboratory at the University of Graz. The experiment was programmed using z-Tree (Fischbacher 2007) and subjects were recruited using the Online Recruitment System for Economic Experiments (ORSEE; Greiner 2004), which ensured that each subject only participated in the experiment once. In the experiment, subjects made decisions in terms of allocating points which were then exchanged at a conversion rate of 100 points = 2.50 euros. Subjects were paid according their decisions and the decisions of a corresponding interaction partner in one randomly selected Slider Measure item from the baseline condition and also one randomly selected Slider Measure item from one randomly chosen reciprocity condition (explicated subsequently). In the latter case, the interaction partner was a subject from experiment A. This remuneration scheme was common knowledge and is incentive compatible. On average, subjects earned 11.10 euros (min = 8.50 and max = 12.80) including a show up fee of 3 euros. The average duration of a session was about sixty minutes.

Phase 1: Measuring baseline SVO. Upon arrival to the research laboratory, subjects were welcomed by the experimenter and each participant drew shuffled cards with numbers. These numbers corresponded to the workstation numbers inside the laboratory. Subjects were then led into the laboratory where they first read the instructions regarding Phase 1 of the experiment and thereafter were given the opportunity to ask questions. The instructions informed subjects that in Phase 1 of the experiment they will be making a series of 15 decisions about how to allocate monetary resources between themselves and an unspecified anonymous other person who would remain unknown to them. After all subjects had read and indicated they understood the instructions, Phase 1 of the experiment began. All subjects then completed the 15 items of the SVO Slider Measure (Murphy, Ackermann, and Handgraaf 2011) in order to assess their individual SVO. This was implemented in terms of a z-Tree SVO module (Crosetto, Weisel, and Winter 2012). Phase 1 therefore served as the baseline condition, where the subjects' unconditional distributive preferences were assessed.

The SVO Slider Measure consists of six primary and nine secondary items. The six primary items allow for the assessment of a person's general SVO on a continuous scale in terms of an angle. An SVO angle of 0° indicates perfect selfishness, while a positive angle indicates the degree of positive concern about the payoff for another person (i.e., increasing prosociality), and a negative angle indicates the

degree of negative concern about the payoff for the another person (i.e., increasing competitiveness).[4] The secondary items of the Slider Measure (items 7–15) allow for distinguishing inequality aversion from joint gain maximization among prosocial DMs. For further details about the SVO Slider Measure, see Murphy, Ackermann, and Handgraaf (2011) and Ackermann and Murphy (2012).

Phase 2: Measuring conditional SVO. When all subjects had completed the baseline condition (i.e., Phase 1) they were given information regarding Phase 2 of the experiment (i.e., the reciprocity conditions) on their computer screens. Subjects were informed that in Phase 2 they would be presented with the choices of four different people who had decided previously (i.e., in experiment A) how to allocate monetary resources between themselves and an anonymous other person by choosing one out of the four options indicated in Table 1. The subjects were then instructed to complete the fifteen items of the Slider Measure with respect to *each* of these four specific and different people. Subjects were informed that they will be randomly matched with one of these four people and that the corresponding decisions made by the subject and the matched other person will be realized for payment.[5] Subjects were further informed that the people they are responding to had only been informed that they would be matched with someone for the determination of final payoffs, but that they had not known that their decisions would be revealed to this other person they would be matched with. We informed subjects about this in order to make clear to them that the choices revealed to them are direct indicators of the other persons' baseline preferences (or their *type*), rather than potentially the result of strategic thinking or misrepresentation. Furthermore, before subjects began to complete the Slider Measure in Phase 2 with respect to each of the four others, the experimenter walked around the laboratory showing the subjects the stack of decision sheets from experiment A to assure them that they were going to respond to real decisions made previously by real people. However, subjects were not given information about the distribution of ''types'' observed in experiment A. They were only informed that ''one decision concerning one of these four other persons will be randomly drawn at the end of the experiment'' and that this ''randomly drawn decision will determine their own payoff and the payoff of this specific other person.''

In this second phase of experiment B, each subject was presented with each of the four possible choices people from experiment A had made. Hence, the experiment is a within-subjects design where the subjects were sequentially responding to one person who had chosen option A (competitive condition), to one person who had chosen option B (individualistic condition), to one person who had chosen option C (prosocial condition), and to one person who had chosen option D (altruistic condition) in experiment A. The sequential order of the four conditions was randomized across subjects in order to control for order effects. In each of the four conditions, subjects were first presented with the four options indicated in Table 1 and the corresponding choice made by the other person in experiment A (competitive, individualistic, prosocial, or altruistic) and were then asked to

indicate in words what they think of this other person and the choice this other person made in an open ended text box. In addition, subjects had to indicate how much the other person had allocated to themselves and to the other and could only proceed in the experiment when the response was verified as correct. This additional comprehension check was implemented to mitigate any possible confusion between the amount allocated to themselves and the amount allocated to the other. The written statements were elicited from subjects for two reasons. First, the statements allow for verifying whether the decision made by the ''other'' was interpreted in a sensible way by the subject. If, for example, a subject confronted with the altruistic choice indicated a statement such as ''This person was obviously motivated to minimize the payoff the other person—in this case me—would receive'', this would indicate that the subject misinterpreted the altruistic choice by confusing it with a competitive choice; these confused responses can be flagged and removed from subsequent analyses. Second, the elicited statements allow for a qualitative analysis of the cognitive and emotional reactions subjects express when confronted with other peoples' revealed social preferences. After the subjects indicated their opinion about the interaction partner and the interaction partner's decision, they were then asked to complete the fifteen items of the SVO Slider Measure in response to this particular interaction partner whose distributive social preferences have been revealed. This procedure was the same under each of the four conditions all subjects went through during the experiment. The procedure of presenting each subject with each possible decision made by people in experiment A is analogous to applying the strategy method developed by Selten (1967). As in other cases, this data collection method yields a rich data set for analysis and modeling.

Subjects then filled out the HEXACO personality questionnaire (Ashton and Lee 2009; Ashton et al. 2004) and answered some sociodemographic questions. The HEXACO measure of personality was chosen because it is claimed to allow for a more fine-grained analysis of prosocial behavior (Hilbig et al. 2013). We used the sixty-item German HEXACO version (Moshagen, Hilbig, and Zettler 2014). Subjects who had finished were asked to step outside the laboratory and wait until all subjects had finished answering the questions. Finally, the experimenter asked subjects to step into the laboratory one by one to privately collect their payment.

In the Results section, one can note that sample size varies slightly across different conditions and analyses. The reason for this is that we include subjects for a particular analysis only if they satisfy both of the following two conditions. First, a subject must show transitive choice patterns in the Slider Measure under the experimental conditions that are analyzed. Intransitive choice patterns in the Slider Measure are indicators of random responding (see Murphy, Ackermann, and Handgraaf 2011) and were observed here only rarely (i.e., 2.7 percent in the baseline condition). Second, there must not be written statements collected under the reciprocity conditions indicating that a subject misinterpreted the choice made by the corresponding other person. If, for example, a subject had written a

Table 2. Descriptive Statistics on SVO and Changes in SVO per Condition.

Condition	Mean SVO	N_{SVO}	Mean ΔSVO	$N_{\Delta SVO}$	Negative shift (%)	No shift (%)	Positive shift (%)
Baseline	21.1	144					
Competitive	12.4	130	−9.4	127	59.8	16.5	23.7
Individualistic	17.7	141	−4.0	137	47.5	22.6	29.9
Prosocial	28.7	147	7.3	144	15.3	15.3	69.4
Altruistic	28.5	142	7.3	139	15.8	18.7	65.5

Note: The column *Mean SVO* indicates the average SVO angle as obtained by the SVO Slider Measure per condition, while the subsequent column (N_{SVO}) provides the corresponding sample sizes. The column *Mean ΔSVO* indicates the average difference between the SVO angles as obtained under the respective condition and those obtained under the baseline condition, while the subsequent column ($N_{\Delta SVO}$) provides the corresponding sample sizes. The numbers in column $N_{\Delta SVO}$ are smaller than the values in column N_{SVO} because subjects had to show a transitive choice pattern in both the baseline condition and the corresponding reciprocity condition in order to be included for the analysis of difference scores (i.e., ΔSVO). The percentages of subjects showing a negative, zero, or positive shift in their SVO angle per condition are reported in columns 6, 7, and 8, respectively.

statement such as "The other person made a choice that maximized his own payoff" in the altruism condition, this subject would be excluded from all analyses involving the altruism condition. Because of these exclusion criteria the sample sizes are not constant across analyses. An overview on the respective sample sizes in the different conditions can be found in Table 2. Less stringent exclusion criteria were also considered, and when implemented the pattern of results remained consistent with what is reported here. Raw data are publicly available for download.

Results

Primary Results

A summary of the main results is shown in Figure 1. The distributions of SVO angles as produced in response to the choices made by the competitive, individualistic, prosocial, and altruistic "others" each differ significantly from the baseline SVO angle distribution as indicated by Kolmogorov–Smirnov (K–S) tests.[6] This is clear evidence that a sufficient number of subjects altered their SVO angles in response to the choices made by the corresponding "others" *in particular directions*. Furthermore, the comparisons[7] among the SVO reciprocity distributions indicate that they are all significantly different from each other as well, except for the comparison between SVO in response to prosociality and in response to altruism (Wilcoxon signed-rank test, $p = .89$; K–S test, $D = 0.08$, $p = .66$). This means that subjects respond differently to competitiveness than to individualism, prosociality, and altruism and differently to individualism than to prosociality and altruism, but respond to prosociality and altruism in generally the same way. To
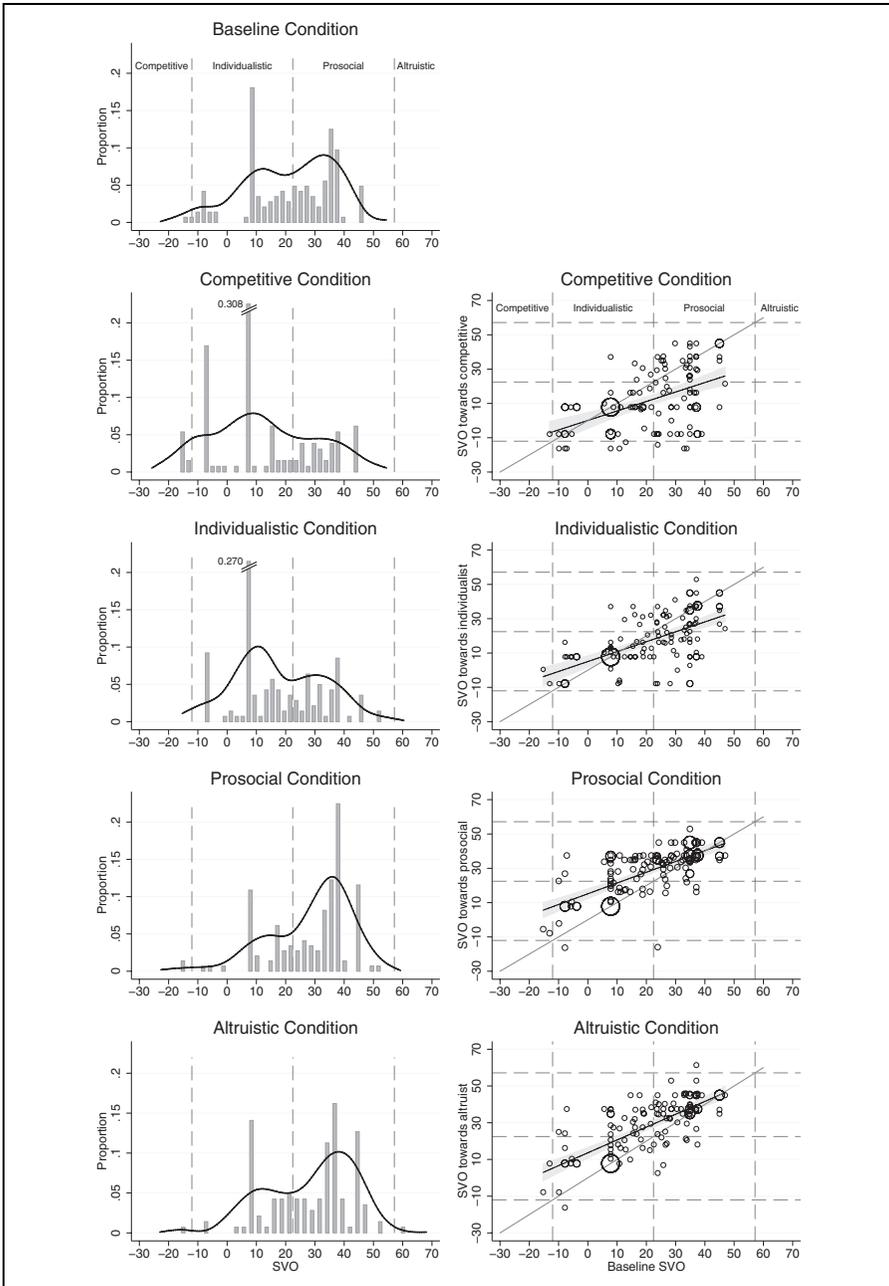
Figure 1. SVO and reciprocity.

have an impression of what the reciprocal reactions mean on the aggregate level in monetary terms, we can compute an approximation of subjects' willingness to pay for an increase of one monetary unit on the side of the interaction partner by taking the tangent of subjects' SVO angles. In the baseline measurement, subjects are on average willing to pay 0.42 monetary units to increase the "other's" payoff by one unit. However, when the "other" is known to have made a competitive choice, the average willingness to pay drops by 40 percent to 0.25 monetary units. In response to an individualistic person, the average willingness to pay drops by 17 percent to 0.35. In contrast, when the "other" is known to be prosocial, the average willingness to pay for a one-unit increase in the payoff for the "other" increases by 38 percent to 0.58 compared to the baseline willingness to pay. Finally, when the "other" is known to have made an altruistic choice, the average willingness to pay increases by 41 percent to 0.59 monetary units compared to the baseline.

The scatter plots in Figure 1 summarize the entirety of experiment B. They show how subjects' SVO angles from the baseline condition relate to their SVO angles as produced in response to each of the four other types. Observations on the diagonal line indicate no shift in SVO (i.e., no reciprocity), while observations above the diagonal indicate upward shifts (i.e., increase in SVO angle indicating positive reciprocity) and observations below the diagonal indicate downward shifts (i.e., decrease in SVO angle indicating negative reciprocity). Table 2 shows the mean shifts in SVO angles as well as the percentage of subjects shifting and not shifting in a particular direction for each condition on the aggregate level. As can be seen in Figure 1, shifts in SVO on the individual level predominantly occur when the interaction partner's SVO diverges from the DM's own SVO. A competitive interaction partner evokes negative reciprocity predominantly among subjects who tend to be prosocial, while subjects who themselves tend to be competitive do not show much of a shift in their SVO angles. In contrast, a prosocial or even an altruistic interaction partner evokes positive reciprocity predominantly among subjects who tend to be individualistic or competitive, while subjects who themselves tend to be prosocial do not show much of a shift, but rather just express their baseline prosociality toward the prosocial interaction partner. It is interesting to see that even subjects who tend to be competitive, as indicated by negative baseline SVO angles, can show considerable positive reciprocity in response to a prosocial or altruistic interaction partner.

However, there are large individual differences in the degree to which individuals are reciprocal. About 10 percent of subjects are not reciprocal at all, as evidenced by identical SVOs in response to the revealed actions of others as compared to their baseline SVO.[8] However, the vast majority of subjects did show at least some degree of reciprocity, and many of them show considerable degrees of reciprocity. Figure 2 shows the distribution of the absolute average shifts in SVO angle. One thing worth noting here, however, is that the degree of reciprocity when computed as the absolute average shift in SVO angle does not take into account that some subjects showed negative reciprocity toward prosociality and/or altruism,
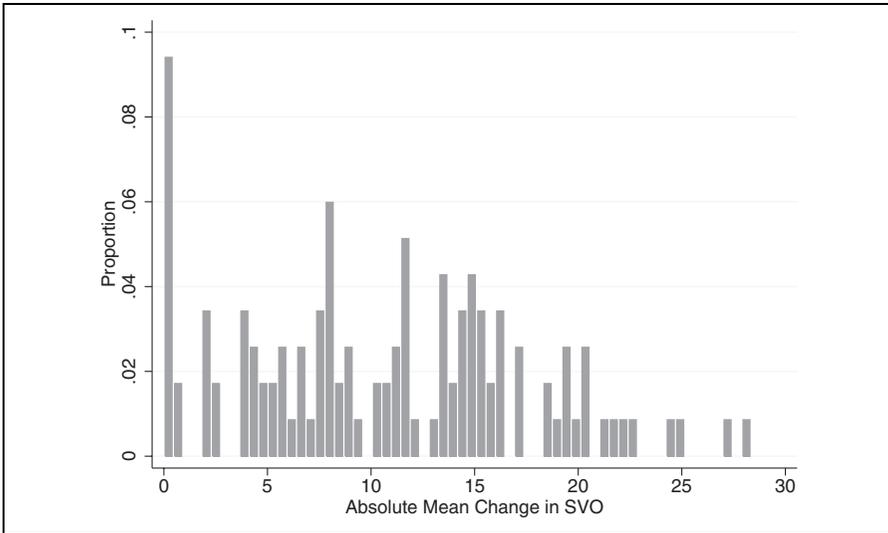
Figure 2. Aggregate absolute reciprocity as indicated by the distribution of average absolute shifts in social value orientation (i.e., $|\Delta SVO|$) across the four conditions(n = 117).

some other subjects showed positive reciprocity toward individualism and/or competitiveness, and yet some other subjects showed consistent positive or negative reciprocity across all conditions. Although these patterns are difficult to rationalize, we are hesitant to exclude these cases from the analysis since we cannot rule out that these patterns are consequences of the subjects' real preferences rather than just noise. Figure 3 shows examples of eight different general patterns of reciprocity profiles, each with an indication of the proportion of subjects showing the corresponding type of pattern. Subject 14 is an example of those 9.4 percent of subjects who do not react in any way to the revealed preferences of "others", while subject 54 is an example of subjects who show negative reciprocity toward competitiveness and individualism and positive reciprocity toward prosociality and altruism. This pattern of both positive reciprocity when matched with a prosocial and altruistic other and negative reciprocity when matched with a competitive and individualistic other is the most common pattern (23.9 percent). Subject 113 is an example of an individualistic subject who does not react to competitiveness nor individualism but shows positive reciprocity toward prosociality and altruism. Subject 111 is an example of a prosocial subject who does not react to prosociality nor altruism but shows negative reciprocity toward competitiveness and individualism only. Note that these patterns are exhibited by only 4.3 percent and 1.7 percent of subjects, respectively. In addition, there are a number of subjects (22.2 percent) who only show partial negative and/or positive reciprocity. An exemplar of this response type is subject 109 who only reacts with a decrease in SVO when matched with a competitive interaction partner and an increase in SVO when matched with an altruistic interaction partner.
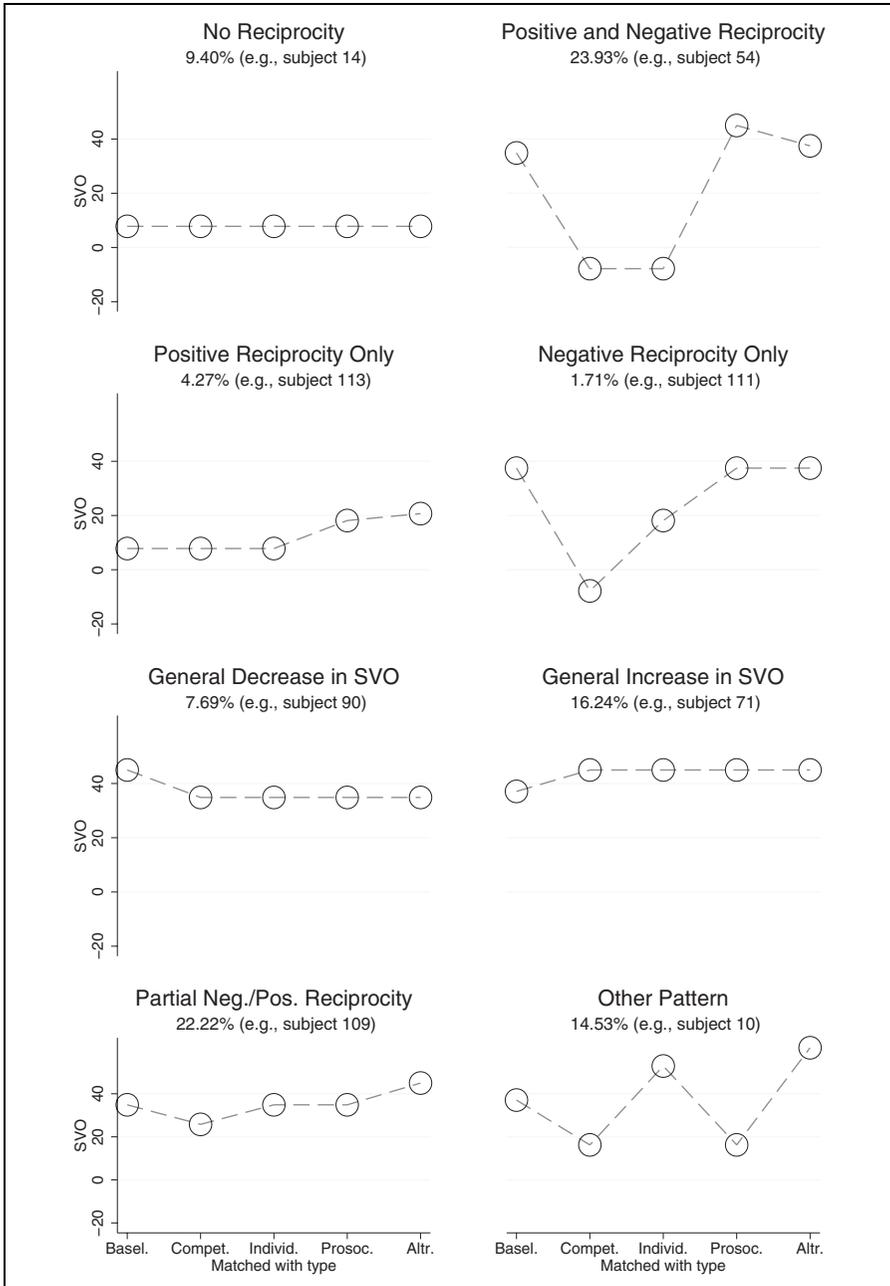
Figure 3. Examples of individual reciprocity profiles (*n* = 117).

Subjects 90 and 71 are examples of subjects who show a reciprocal reaction to others but do not differentiate between these revealed types, resulting in a general decrease (7.7 percent) or increase (16.2 percent) in SVO across conditions. As mentioned earlier, patterns of that kind are difficult to explain. Perhaps the simplest explanation for such patterns would be that these subjects are not reciprocal at all, but made a mistake in one or a few items in the Slider Measure under the baseline condition which they then corrected under the reciprocity conditions. Another explanation might be that these subjects really have different distributive social preferences when the other is a specified particular other person compared to an unspecified "someone" who will be randomly selected after the decisions have been made. However, we cannot shed further light on the rationale behind patterns of that kind by means of the data available to us and can only speculate about potential causes. Other patterns that cannot be assigned to one of the aforementioned categories are exhibited by 14.5 percent of the subjects. These examples make clear that there is substantial heterogeneity in the patterns of how subjects react to various others when information about them is revealed.

One question would be to consider if different types of DMs are more reactive than others. For example, it could be conjectured that people with prosocial SVOs are more reactive than are people with individualistic motivations. However, the bivariate correlation between baseline SVO and degree of reciprocity (i.e., absolute shifts in SVO across the four conditions) is not significantly different from zero, indicating that there is no significant relation between SVO and reciprocity (Figure A1 of the Online Appendix shows the corresponding scatter plot). The absolute degree of reactivity in people's social preferences is not conditioned on their baseline preferences.

Secondary Results

SVO and its relation to positive and negative reciprocity. The mean positive change in SVO angles over all four conditions captures the strength of subjects' positive reciprocal reactions, while the absolute mean negative change indicates the strength of their negative reciprocal reactions. On the aggregate level, subjects showed an average positive change in SVO of 5.3 degrees and an average negative change in SVO of 5.4 degrees over all four conditions.[9] We observe that 26.5 percent of all subjects exhibit no positive reciprocity, and 35.0 percent do not show negative reciprocity. Overall, we do not find a significant difference between the two distributions (K–S test, $D = 0.11$, $p = .407$), indicating that the effects of positive and negative reciprocity are about the same on the aggregate.

Figure 4a shows a scatter plot indicating how positive and negative reciprocity are jointly distributed on the individual level. We observe some cases near the diagonal that exhibit a similar degree of positive and negative reciprocity. Cases below the diagonal show stronger negative than positive reciprocity including cases who exhibit negative reciprocity only. The opposite is true for cases above
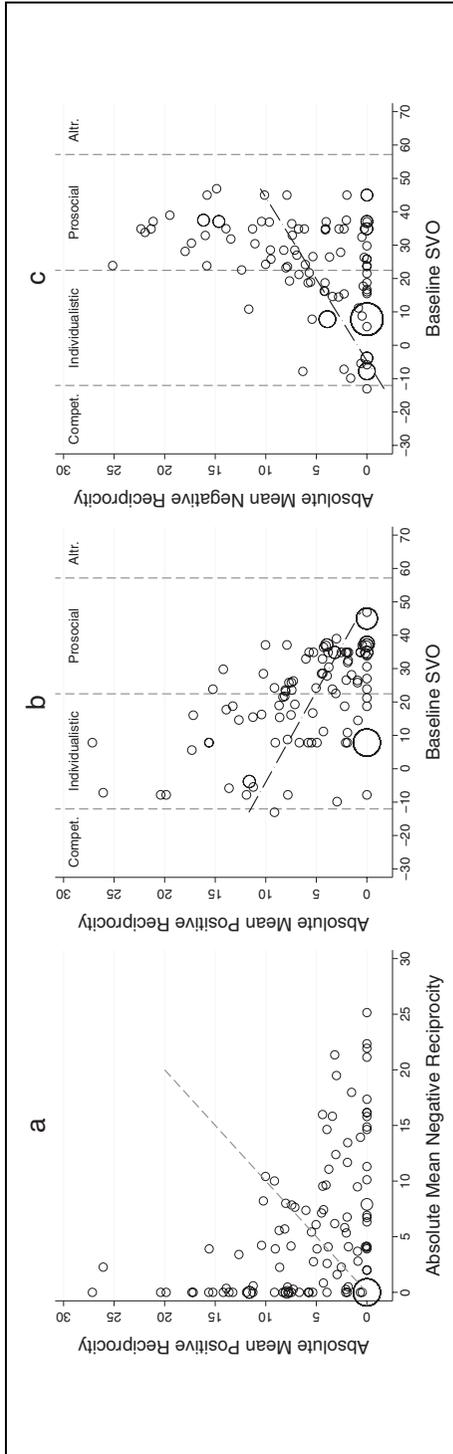
Figure 4. Relations between positive and negative reciprocity and baseline SVO ($n = 117$).

the diagonal where either less negative reciprocity as compared to positive reciprocity is observed or where negative reciprocity is completely absent. We can identify several factors that are partially responsible for the pattern shown in Figure 4a. First of all, we have seen that individual reciprocity profiles are dependent on SVO in that subjects with higher SVO angles predominantly show negative reciprocal reactions toward competitive and individualistic interaction partners while showing no or only little positive reciprocity toward prosociality and altruism, while the opposite holds for subjects with lower SVO angles. Furthermore, we have seen that some subjects appear to show only positive or only negative reciprocity across all conditions. And finally, there are subjects—predominantly with intermediate baseline SVO angles—who show as much negative reciprocity toward competitiveness and individualism as they show positive reciprocity toward prosociality and altruism. As mentioned earlier, Figure 3 shows examples for each of these patterns. The combination of these factors together, we argue, shape the triangle form of observations shown in Figure 4a and lead to a negative correlation of $r = -0.367$ ($p < .001$) between positive and negative reciprocity. This interpretation of the results is also supported by the observation of a negative correlation of $r = -0.474$ ($p < .001$) between SVO (baseline) and positive reciprocity, and a positive correlation of $r = 0.479$ ($p < .001$) between SVO (baseline) and negative reciprocity as visualized in Figures 4b and 4c, respectively. Since there are two separate effects in opposing directions for positive and negative reciprocity, this explains why we observe no correlation between baseline SVO and overall reciprocity.

HEXACO. All factors of the HEXACO questionnaire show acceptable internal reliability.[10] The correlation between the Honesty-Humility factor and baseline SVO is $r = 0.25$ ($p < .01$). The direction and magnitude of this relation are in accordance with previous findings (e.g., Hilbig and Zettler 2009; Hilbig et al. 2012). Also, this relation is the only one between baseline SVO and HEXACO scales that remains significant after Bonferroni correction. Without Bonferroni correction, the relation between baseline SVO and the HEXACO scale "Openness" is significant as well with $r = 0.17$ ($p < .05$).

We find no relationship of any of the HEXACO scales with overall reciprocity. However, when SVO (baseline) is statistically controlled, we find a significant positive relation between the Honest-Humility factor and average *positive* reciprocity ($r_{partial} = .23$, $p < .05$), and a significant negative relation between Honesty-Humility and average absolute *negative* reciprocity ($r_{partial} = -.20$, $p < .05$).

Honesty-Humility is described as a "tendency to be fair and genuine in dealing with others, in the sense of cooperating with others even when one might exploit them without suffering retaliation" (Ashton and Lee 2007, 156). So it makes sense that people who exhibit this trait show less negative and more positive reciprocity than those who do not. These findings add further support to the argument of separating Honesty-Humility from other personality traits (Hilbig et al. 2013), since it

appears to have separate effects on cooperative behavior in different situations. Nevertheless, the overall pattern of our results also suggests that tendencies toward positive and negative reciprocity may indeed be considered relatively independent motivational inclinations rather than collinear ones, which would be in support of other findings (e.g., Eisenberger et al. 2004; Egloff, Richter, and Schmukle 2013; Yamagishi et al. 2012). Importantly, however, our results are exploratory rather than confirmatory in this respect, and a new experiment with a potentially larger sample size would be required to judge the robustness of the effects concerning the relationship between personality factors and reciprocity that we found.

Inequality aversion. An alternative explanation for shifts in SVO angles may be that people do not change their baseline distributive social preferences, but simply always express the same distributive social preference, namely, inequality aversion, when responding to advance payoff allocations. If a person wanted to minimize the difference between *final* payoffs for both DMs, then this person would be expected to show a behavioral pattern that—by itself—is indistinguishable from perfect reciprocity. That is, such a person would respond competitively to a competitor, individualistically to an individualist, prosocially to a prosocial, and altruistically to an altruist, since these responses would guarantee that both DMs receive equal payoffs in the end. We can test whether this alternative explanation holds. The secondary items of the Slider Measure allow for the assessment of the degree of IA among prosocial people (i.e., subjects with an SVO angle between $22.45°$ and $57.15°$) given that their choices in the secondary items, too, are consistent with a prosocial orientation (i.e., IA or joint gain maximization) rather than an individualistic or altruistic orientation.[11] Concretely, for subjects who meet these requirements, we can compute an IA index ranging from zero to one, where an index of zero indicates that a person is perfectly inequality averse (i.e., showing a choice pattern in the secondary items of the Slider Measure that is in perfect accordance with minimizing differences in payoffs), and an index of one indicates that the person is perfectly joint gain maximizing (i.e., showing a choice pattern in the secondary items of the Slider Measure that is in perfect accordance with maximizing the sum of payoffs). Among the 117 subjects who showed transitive choice patterns in the baseline SVO measurement and correct interpretation of the interaction partner's type across all conditions, there are 51 prosocial subjects (of the 64 prosocials in total) for which an IA index can be computed. Among these 51 prosocial subjects, 33 (i.e., 28.21 percent of all 117 subjects under consideration) tend to be inequality averse as indicated by an IA index of less than 0.5. Only 1 of the 117 subjects under consideration is perfectly inequality averse as indicated by an IA index of exactly zero (Figure A3 of the Online Appendix shows the distribution of the IA index). If the alternative explanation for the shifts in SVO angles holds, then we would expect that the vast majority of these 33 subjects who tend to be inequality averse respond competitively to a competitive interaction partner. However, this is not what we find. In fact, only 1 of these 33 inequality averters

Table 3. Proportion of Inequality Averse, Joint Max, and Other Subjects as well as the Median IA Index per Condition.

		Conditions				
		Baseline	Competitive	Individualistic	Prosocial	Altruistic
Other types		45.3	71.8	64.1	31.6	33.3
	IneqAvers	28.2	13.7	15.4	35.9	27.4
Prosocial types	JointMax	15.4	9.4	12.8	23.9	30.8
	n/a	11.1	5.1	7.7	8.5	8.5
IA median		.35	.33	.38	.34	.53

Note: Numbers in row 1 through row 4 are percentages of subjects categorized in the respective class per condition. These proportions are based on $n = 117$ subjects. The numbers in row 5 are the median values of the inequality aversion (IA) index from those subjects who are categorized as prosocial and show a choice pattern in the Slider Measure's secondary items consistent with a prosocial orientation. The inequality aversion (IA) index can range between 0 (i.e., perfect inequality aversion) and 1 (i.e., perfect joint gain maximizing).

responded competitively to a competitor, thereby equalizing final payoffs (the distribution of these 33 subjects' SVO angles when matched with a competitive other is shown in Figure A4 of the Online Appendix). The modal response was individualistic with an SVO angle of $-7.82°$, which results from a perfectly individualistic choice pattern across all items where benefiting or harming the "other" is costly, and a competitive choice in the item where harming the "other" is free. Clearly, such a choice pattern does not serve the purpose of equalizing final payoffs, but punishing the other when it is free, and being concerned with one's own payoff when punishment is costly. We therefore refute this alternative explanation of inequality aversion as the main driver of shifts in SVO angles.

A further question regarding inequality aversion is whether this particular distributive social preference can be expressed in terms of a reciprocal reaction. Table 3 shows for each condition the percentage of subjects that can be categorized as prosocial (i.e., subjects with an SVO angle between $22.45°$ and $57.15°$) subdivided into three categories as follows: Joint gain maximization (JointMax, i.e., prosocial subjects with an IA index $\geq .5$), IA (IneqAvers, i.e., prosocial subjects with an IA index $< .5$), and prosocials whose choice patterns do not allow for the computation of an IA index (n/a, i.e., prosocial subjects who showed a choice pattern in the secondary items that is neither consistent with joint gain maximization, nor inequality aversion). The fifth row shows the median IA index across all subjects for which an IA index can be computed in the corresponding condition. In order to facilitate comparisons across conditions, only those $n = 117$ subjects are considered who show transitive choice patterns in the Slider Measure and no misperception of the "other's" types across all conditions. As can be seen in Table 3, the percentage of subjects showing an inequality averse choice pattern in the Slider Measure varies substantially across conditions in expected ways. Regarding our question, the

Table 4. Type Comparison between Baseline and Prosocial Condition.

		Prosocial condition			
		Other types	IneqAvers	JointMax	n/a
Baseline	Other types	31	12	4	6
	Inequality aversion	2	24	5	2
	Joint gain maximization	1	3	14	0
	n/a	3	3	5	2

Note: Numbers in the table are counts of subjects ($n = 117$).

comparison between the percentage of inequality averse subjects in the baseline condition and the percentage of inequality averters in response to prosociality is most interesting. There are 7.7 percent more inequality averters among the subjects when they are responding to prosociality compared to the baseline condition. Table 4 informs about how this 7.7 percent increase in total is realized. While 9 subjects who showed inequality aversion in the baseline condition changed their distributive social preferences when responding to prosociality, 18 subjects who did not show concern about equality in the baseline condition responded to prosociality with an inequality averse choice pattern. Hence, there are 18 observations of inequality aversion as a response to prosociality. This is clear evidence that IA can be expressed in terms of a reciprocal reaction.

Discussion

We have defined reciprocity as a change in the distributive social preferences of an individual in response to information about an interaction partner's characteristics. Thus, we have conceptualized reciprocity as an individual difference that can be measured and used as a dependent variable. We show that the vast majority of people do change their distributive social preferences toward an interaction partner when the interaction partner's "type" is revealed, thereby showing reciprocity. Furthermore, we show that there is considerable heterogeneity in the degree to which people are reciprocal and how degrees of reciprocity relate to SVO. These findings have significant implications regarding the study and modeling of preferences, beliefs, and behavior in experimental games that involve the revelation of information about interaction partners. It has been common practice in the economic sciences to infer distributive social preferences from responder behavior in experimental settings such as ultimatum games, for instance. It has been neglected, however, that the responder behavior may not reveal the responder's baseline distributive social preferences, but the responder's reciprocal reaction to revealed information about the proposer's characteristics, namely, the proposer's behavior. In situations of that kind, it has often been assumed that the expression of one particular

distributive social preference, inequality aversion, is responsible for the decision to accept an equal split and refute offers that deviate too much from it in ultimatum games, for instance. We have shown that there is another explanation for the behavioral pattern often observed in these types of experimental games, namely, that distributive social preferences may change in response to information about the interaction partner, such that DMs may assign a different weight—positive, zero, or negative—to the outcome for the interaction partner once information about the interaction partner is revealed compared to when no information is available. Consequently, both punishment and reward behavior can result from reciprocity as an individual difference and social preferences show a reactivity to information about the other.

Our results clearly support the validity of recently developed models of behavior in experimental games that take into account the significance of information about the types of interaction partners (e.g., Charness and Rabin 2002; Cox, Friedman, and Gjerstad 2007; Dufwenberg and Kirchsteiger 2004; Falk and Fischbacher 2006; Falk, Fehr, and Fischbacher 2008; Levine 1998; Rabin 1993; Stanca, Bruni, and Corazzini 2009). However, there is large heterogeneity in both peoples' baseline distributive social preferences and the degree to which people are reciprocal. Our results suggest that apart from people's beliefs, both individual differences in distributive social preferences and individual differences in the degree to which these preferences can change in response to information about the environment—predominantly about interaction partners—are key to understanding behavior in many situations of interdependence and strategic choice.

Finally, this article highlights the general usefulness and broad potential of measuring SVO as a dependent variable. One can readily imagine a host of similarly structured experiments where baseline SVO is first measured and then contrasted against another measure of SVO in regard to a specified other, or in regard to the same other in a different context or with updated information. For example, one could consider to what extent SVO changes when the other person is part of an in-group/out-group? Or how does the reactivity of social preferences relate to the DM's own, and the other's, demographic or other characteristics. Are there interactions between the DM's and the other's characteristics in affecting reciprocal reactions (i.e., homophily effects)? How does the status of the other (earned or otherwise bestowed arbitrarily) influence SVO? In general, to what degree are DM's willing to make costly trade-offs when the features (attributes, identity, or past behavior) about the other are made explicit? And to what extent are these preference dynamics dependent on social context? Our findings demonstrate the flexibility of social preferences, but understanding the general structure of these contingencies at the individual level is an open question and one of great interest across many disciplines in the social sciences.

Authors' Note

Declaration of Conflicting Interests

Funding

Notes

1. The following is an example of a joint utility function that can account for social preferences: $U(\pi_s, \pi_o) = \pi_s + \alpha\pi_o - \beta|\pi_s - \pi_o|$, where π_s is the payoff for the self, π_o is the payoff for the other, and α and β are parameters (weights) attached to the respective terms.
2. This form of reciprocity that is expressed in one-shot situations, where no future interactions are taking place and thus beliefs are inconsequential, is sometimes referred to as ''altruistic reciprocity'' and is in accordance with our definition of reciprocity given previously.
3. In the terminology of Messick and McClintock (1968), the item is a quadruple dominance decomposed game.
4. In essence, the SVO angle is a trigonometric function of parameter α in the utility function $U(\pi_s, \pi_o) = \pi_s + \alpha\pi_o$, where π_s is the payoff for the self and π_o is the payoff for the other person.
5. The relative frequencies of how many times each of the four different options had been chosen in experiment A were taken as the probabilities that a subject would be matched with a person who made a corresponding choice for determining final payoffs. However, subjects in experiment B were neither informed about these relative frequencies nor about the corresponding details of the matching procedure.
6. The Kolmogorov–Smirnov (K–S) statistics from the SVO reciprocity distributions in comparison to the baseline SVO distribution are $D = 0.30$, $p < .001$, for the competitiveness condition; $D = 0.17$, $p < .022$, for the individualism condition; $D = 0.25$, $p < .001$, for the prosociality condition; and $D = 0.27$, $p < .001$, for the altruism condition. Wilcoxon signed rank tests corroborate these results, indicating that subjects' SVO angles changed in each reciprocity condition compared to the baseline condition.
7. Again, both K–S tests and Wilcoxon signed rank tests were used and indicated the same results.
8. The actual share of subjects who do not show significant reciprocity is likely slightly underestimated due to measurement error. There may be subjects whose SVO angles changed slightly across conditions not because they reacted to the others in some way, but because they did not reproduce their choice pattern exactly between conditions. However, due to the Slider Measure's very high reliability (Test–retest $r = .915$; see Murphy, Ackermann, and Handgraaf 2011), measurement errors are expected to be fairly small.

9. The difference is not significant (Wilcoxon signed-rank test, $p = .985$). This difference is increased if we exclude cases that show no positive or negative reciprocity (7.2 and 8.3) but remains insignificant (Wilcoxon signed-rank test, $p = .171$). Figure A2 of the Online Appendix shows the distributions of positive and negative reciprocity separately.

10. Cronbach's α for the factors are as follows: Honesty-Humility: 0.77; Emotionality: 0.84; Extraversion: 0.77; Agreeableness: 0.76; Conscientiousness: 0.81; Openness: 0.78.

11. For details about the IA index, see Murphy, Ackermann, and Handgraaf (2011) and Ackermann and Murphy (2012).

References

Ackermann, K. A., and R. O. Murphy. 2012. "Tutorial on How to Evaluate the SVO Slider Measure's Secondary Items." http://vlab.ethz.ch/svo/SVO_Slider/SVO_Slider_scoring.html. Accessed November 30, 2012.

Ajzen, I., T. Brown, and F. Carvajal. 2004. "Explaining the Discrepancy between Intentions and Actions: The Case of Hypothetical Bias in Contingent Valuation." *Personality and Social Psychology Bulletin* 30 (9): 1108-121.

Allen, B. P., and C. R. Potkay. 1981. "On the Arbitrary Distinction between States and Traits." *Journal of Personality and Social Psychology* 41 (5): 916-28.

Andreoni, J., W. Harbaugh, and L. Vesterlund. 2008. "Altruism in Experiments." In *New Palgrave Dictionary in Economics*, 2nd ed., edited by S. Durlauf and L. Blume. Macmillan.

Ashton, M. C., and K. Lee. 2007. "Empirical, Theoretical, and Practical Advantages of the HEXACO Model of Personality Structure." *Personality and Social Psychology Review* 11 (2): 150-166.

Ashton, M. C., and K. Lee. 2009. "The HEXACO-60: A Short Measure of the Major Dimensions of Personality." *Journal of Personality Assessment* 91 (4): 340-45.

Ashton, M. C., K. Lee, M. Perugini, P. Szarota, R. E. De Vries, L. Di Blas, K. Boies, and B. De Raad. 2004. "A Six-factor Structure of Personality-descriptive Adjectives: Solutions from Psycholexical Studies in Seven Languages." *Journal of Personality and Social Psychology* 86 (2): 356-66.

Au, W., and J. Kwong. 2004. "Measurements and Effects of Social-value Orientation in Social Dilemmas: A Review." In *Contemporary Psychological Research on Social Dilemmas*, edited by R. Suleiman, D. Budescu, I. Fischer, and D. Messick, 71-98. New York: Cambridge University Press.

Balliet, D., C. Parks, and J. Joireman. 2009. "Social Value Orientation and Cooperation in Social Dilemmas: A Meta-analysis." *Group Processes & Intergroup Relations* 12 (4): 533-47.

Berg, J., J. Dickhaut, and K. McCabe. 1995. "Trust, Reciprocity, and Social History." *Games and Economic Behavior* 10 (1): 122-42.

Bogaert, S., C. Boone, and C. Declerck. 2008. "Social Value Orientation and Cooperation in Social Dilemmas: A Review and Conceptual Model." *British Journal of Social Psychology* 47 (3): 453-80.

Bolton, G., and A. Ockenfels. 2000. "ERC: A Theory of Equity, Reciprocity, and Competition." *The American Economic Review* 90 (1): 166-93.

Boone, C., C. Declerck, and T. Kiyonari. 2010. "Inducing Cooperative Behavior among Proselfs Versus Prosocials: The Moderating Role of Incentives and Trust." *Journal of Conflict Resolution* 54 (5): 799-824.

Camerer, C. 2003. *Behavioral Game Theory: Experiments in Strategic Interaction*. Princeton, NJ: Princeton University Press.

Camerer, C., and R. Thaler. 1995. "Anomalies: Ultimatums, Dictators and Manners." *The Journal of Economic Perspectives* 9 (2): 209-19.

Catell, R. B. 1966. "Anxiety and Motivation: Theory and Crucial Experiments." In *Anxiety and Behavior*, edited by C. D. Spielberger, 23-62. New York: Academic Press.

Chaplin, W. F., O. P. John, and L. R. Goldberg. 1988. "Conceptions of States and Traits: Dimensional Attributes with Ideals as Prototypes." *Journal of Personality and Social Psychology* 54 (4): 541-57.

Charness, G., and M. Rabin. 2002. "Understanding Social Preferences with Simple Tests." *Quarterly Journal of Economics* 117 (3): 817-69.

Clark, K., and M. Sefton. 2001. "The Sequential Prisoner's Dilemma: Evidence on Reciprocation." *The Economic Journal* 111 (468): 51-68.

Cox, J. 2004. "How to Identify Trust and Reciprocity." *Games and Economic Behavior* 46 (2): 260-81.

Cox, J., D. Friedman, and S. Gjerstad. 2007. "A Tractable Model of Reciprocity and Fairness." *Games and Economic Behavior* 59 (1): 17-45.

Crosetto, P., O. Weisel, and F. Winter. 2012. "A Flexible z-Tree Implementation of the Social Value Orientation Slider Measure (Murphy et al. 2011) – Manual –." http://ssrn.com/abstract=2195598. Accessed January 2, 2013.

De Cremer, D., and P. Van Lange. 2001. "Why Prosocials Exhibit Greater Cooperation than Proselfs: The Roles of Social Responsibility and Reciprocity." *European Journal of Personality* 15 (S1): 5-18.

Diekmann, A. 2004. "The Power of Reciprocity: Fairness, Reciprocity, and Stakes in Variants of the Dictator Game." *Journal of Conflict Resolution* 48 (4): 487-505.

Dufwenberg, M., and G. Kirchsteiger. 2004. "A Theory of Sequential Reciprocity." *Games and Economic Behavior* 47 (2): 268-98.

Egloff, B., D. Richter, and S. C. Schmukle. 2013. "Need for Conclusive Evidence that Positive and Negative Reciprocity are Unrelated." *Proceedings of the National Academy of Sciences* 110 (9): E786.

Eisenberger, R., P. Lynch, J. Aselage, and S. Rohdieck. 2004. "Who Takes Most Revenge? Individual Differences in Negative Reciprocity Norm Endorsement." *Personality and Social Psychology Bulletin* 30 (6): 787-99.

Endler, N. S., and N. L. Kocovski. 2001. "State and Trait Anxiety Revisited." *Anxiety Disorders* 15 (3): 231-45.

Falk, A., E. Fehr, and U. Fischbacher. 2008. "Testing Theories of Fairness — Intentions Matter." *Games and Economic Behavior* 62 (1): 287-303.

Falk, A., and U. Fischbacher. 2006. "A Theory of Reciprocity." *Games and Economic Behavior* 54 (2): 293-315.

Fehr, E., and U. Fischbacher. 2003. "The Nature of Human Altruism." *Nature* 425 (6960): 785-91.

Fehr, E., and S. Gächter. 2000a. "Cooperation and Punishment in Public Goods Experiments." *The American Economic Review* 90 (4): 980-94.

Fehr, E., and S. Gächter. 2000b. "Fairness and Retaliation: The Economics of Reciprocity." *Journal of Economic Perspectives* 14 (3): 159-81.

Fehr, E., and K. Schmidt. 1999. "A Theory of Fairness, Competition, and Cooperation." *Quarterly Journal of Economics* 114 (3): 817-68.

Fischbacher, U. 2007. "z-tree: Zurich Toolbox for Ready-made Economic Experiments." *Experimental Economics* 10 (2): 171-78.

Fleeson, W. 2001. "Toward a Structure- and Process-integrated View of Personality: Traits as Density Distribution of States." *Journal of Personality and Social Psychology* 80 (6): 1011-27.

Fleeson, W. 2004. "Moving Personality beyond the Person-situation Debate: The Challenge and the Opportunity of Within-person Variability." *Current Directions in Psychological Science* 13 (2): 83-87.

Fridhandler, B. M. 1986. "Conceptual Note on State, Trait, and the State-trait Distinction." *Journal of Personality and Social Psychology* 50 (1): 169-74.

Greiner, B. 2004. "An Online Recruitment System for Economic Experiments." In *Forschung und wissenschaftliches Rechnen 2003*, GWDG Bericht 63, edited by K. Kremer and V. Macho, 79-93. Göttingen, Germany: Ges. für Wiss. Daten-verarbeitung.

Grzelak, J. 1982. "Preferences and Cognitive Processes in Interdependence Situations: A Theoretical Analysis of cooperation." In *Cooperation and Helping Behavior*, edited by V. Derlega and J. Grzelak, 95-122. New York: Academic Press.

Hilbig, B. E., and I. Zettler. 2009. "Pillars of Cooperation: Honesty-Humility, Social Value Orientations, and Economic Behavior." *Journal of Research in Personality* 43 (3): 516-19.

Hilbig, B. E., I. Zettler, F. Leist, and T. Heydasch. 2013. "It Takes Two: Honesty–humility and Agreeableness Differentially Predict Active versus Reactive Cooperation." *Personality and Individual Differences* 54 (5): 598-603.

Hilbig, B. E., I. Zettler, M. Moshagen, and T. Heydasch. 2012. "Tracing the Path from Personality—via Cooperativeness—to Conservation." *European Journal of Personality* 27 (4): 319-27.

Kanagaretnam, K., S. Mestelman, K. Nainar, and M. Shehata. 2009. "The Impact of Social Value Orientation and Risk Attitudes on Trust and Reciprocity." *Journal of Economic Psychology* 30 (3): 368-80.

Kelley, H., and A. Stahelski. 1970. "Social Interaction Basis of Cooperators and Competitors Beliefs about Others." *Journal of Personality and Social Psychology* 16 (1): 66-91.

Levine, D. 1998. "Modeling Altruism and Spitefulness in Experiments." *Review of Economic Dynamics* 1 (3): 593-622.

Loewenstein, G., M. Bazerman, and L. Thompson. 1989. "Social Utility and Decision-making in Interpersonal Contexts." *Journal of Personality and Social Psychology* 57 (3): 426-41.

McClintock, C., and E. Van Avermaet. 1982. "Social Values and Rules of Fairness: A Theoretical Perspective." In *Cooperation and Helping Behavior*, edited by V. Derlega and J. Grzelak, 43-71. New York: Academic Press.

Messick, D., and C. McClintock. 1968. "Motivational Bases of Choice in Experimental Games." *Journal of Experimental Social Psychology* 4 (1): 1-25.

Messick, D., and W. Thorngate. 1967. "Relative Gain Maximization in Experimental Games." *Journal of Experimental Social Psychology* 3 (1): 85-101.

Mischel, W., and Y. Shoda. 1998. "Reconciling Processing Dynamics and Personality Dispositions." *Annual Review of Psychology* 49 (1): 229-58.

Moshagen, M., B. E. Hilbig, and I. Zettler. 2014. "Faktorenstruktur, psychometrische Eigenschaften und Messinvarianz der deutschsprachigen Version des 60-Item HEXACO Persönlichkeitsinventars." *Diagnostica* 60 (2): 86-97.

Murphy, R. O., and K. A. Ackermann. 2014a. "Explaining Behavior in Public Goods Games: How Preferences and Beliefs Affect Contribution Levels." Manuscript in preparation, ETH Zürich.

Murphy, R. O., and K. A. Ackermann. 2014b. "Social Value Orientation: Theoretical and Measurement Issues in the Study of Social Preferences." *Personality and Social Psychology Review* 18 (1): 13-41.

Murphy, R. O., K. A. Ackermann, and M. J. J. Handgraaf. 2011. "Measuring Social Value Orientation." *Judgment and Decision Making* 6 (8): 771-81.

Parks, C. D., and A. C. Rumble. 2001. "Elements of Reciprocity and Social Value Orientation." *Personality and Social Psychology Bulletin* 27 (10): 1301-9.

Perugini, M., and M. Gallucci. 2001. "Individual Differences and Social Norms: The Distinction between Reciprocators and Prosocials." *European Journal of Personality* 15 (S1): 19-35.

Perugini, M., M. Gallucci, F. Presaghi, and A. Ercolani. 2003. "The Personal Norm of Reciprocity." *European Journal of Personality* 17 (4): 251-83.

Rabin, M. 1993. "Incorporating Fairness into Game-theory and Economics." *American Economic Review* 83 (5): 1281-302.

Radzicki, J. 1976. "Technique of Conjoint Measurement of Subjective Value of Own and Other's Gains." *Polish Psychological Bulletin* 7 (3): 179-86.

Sawyer, J. 1966. "Altruism Scale—A Measure of Co-operative, Individualistic, and Competitive Interpersonal Orientation." *American Journal of Sociology* 71 (4): 407-16.

Schubert, M., and J. G. Lambsdorff. 2014. "Negative Reciprocity in an Environment of Violent Conflict: Experimental Evidence from the Occupied Palestinian Territories." *Journal of Conflict Resolution* 58 (4): 539-563.

Selten, R. 1967. "Die Strategiemethode zur Erforschung des eingeschränkten rationalen Verhaltens im Rahmen eines Oligopolexperimentes." In *Beiträge zur Experimentellen Wirtschaftsforschung*, edited by H. Sauermann, 136-68. J. C. B. Mohr: Tübingen.

Sheeran, P. 2002. "Intention-behavior Relations: A Conceptual and Empirical Review." *European Review of Social Psychology* 12 (1): 1-36.

Sheldon, K. 1999. "Learning the Lessons of Tit-for-tat: Even Competitors Can Get the Message." *Journal of Personality and Social Psychology* 77 (6): 1245-53.

Sonnemans, J., F. van Dijk, and F. van Winden. 2006. "On the Dynamics of Social Ties Structures in Groups." *Journal of Economic Psychology* 27 (2): 187-204.

Spielberger, C. D. 2010. "State-trait Anxiety Inventory." In *The Corsini Encyclopedia of Psychology*, edited by I. B. Weiner and E. E. Craighead, 1698-99. Hoboken, NJ: John Wiley & Sons.

Spielberger, C. D., S. J. Sydeman, A. E. Owen, and B. J. Marsh. 1999. "Measuring Anxiety and Anger with the State-trait Anxiety Inventory (STAI) and the State-trait Anger Expression Inventory (STAXI)." In *The Use of Psychological Testing for Treatment Planning and Outcomes Assessment*, 2nd ed., edited by M. E. Maruish, 993-1021. Hillsdale, NJ: Lawrence Erlbaum Associates.

Stanca, L., L. Bruni, and L. Corazzini. 2009. "Testing Theories of Reciprocity: Do Motivations Matter?" *Journal of Economic Behavior and Organization* 71 (2): 233-45.

Stevens, S. 1946. "On the Theory of Scales of Measurement." *Science* 103 (2684): 677-80.

Steyer, R., D. Ferring, and M. J. Schmitt. 1992. "States and Traits in Psychological Assessment." *European Journal of Psychological Assessment* 8 (2): 79-98.

Van Lange, P. 1999. "The Pursuit of Joint Outcomes and Equality in Outcomes: An Integrative Model of Social Value Orientation." *Journal of Personality and Social Psychology* 77 (2): 337-49.

Van Lange, P., and A. Semin-Goossens. 1998. "The Boundaries of Reciprocal Cooperation." *European Journal of Social Psychology* 28 (5): 847-54.

Wyer, R. 1969. "Prediction of Behavior in 2-person Games." *Journal of Personality and Social Psychology* 13 (3): 222-38.

Yamagishi, T., Y. Horita, N. Mifune, H. Hashimoto, Y. Li, M. Shinada, A. Miura, K. Inukai, H. Takagishi, and D. Simunovic. 2012. "Rejection of Unfair Offers in the Ultimatum Game is no Evidence of Strong Reciprocity." *Proceedings of the National Academy of Sciences* 109 (50): 20364-68.

Zuckerman, M. 1983. "The Distinction between Trait and State Scales is *not* Arbitrary: Comment on Allen and Potkay's On the Arbitrary Distinction between Traits and States." *Journal of Personality and Social Psychology* 44 (5): 1083-86.