Strong, Bold, and Kind: Self-Control and Cooperation in Social Dilemmas*

Martin G. Kocher[A,B,D]

Peter Martinsson[B]

Kristian Ove R. Myrseth[C]

Conny E. Wollbrant[B]

Abstract: We develop a model that relates self-control to cooperation patterns in social dilemmas, and we test the model in a laboratory public goods experiment. As predicted, we find a robust association between stronger self-control and higher levels of cooperation, and the association is at its strongest when the decision maker's risk aversion is low and the cooperation levels of others high. We interpret the pattern as evidence for the notion that individuals may experience an impulse to act in self-interest—and that cooperative behavior benefits from self-control. Free-riders differ from other contributor types only in their tendency not to have identified a self-control conflict in the first place.

JEL: C91, D03, H40.

Keywords: experiment, public good, self-control, cooperation, risk

---

* Financial support from the Swedish Research Council (*Vetenskapsrådet*), from Formas through the program Human Cooperation to Manage Natural Resources (COMMONS), and the *Ideenfonds* of the University of Munich is gratefully acknowledged. For valuable comments and suggestions, we thank Enrique Fatas, Amrei Marie Lahno, Kei Tsutsui, Lise Vesterlund, two anonymous referees, and seminar participants at the EWEBE Workshop 2011 in Munich; the CESifo Behavioral Economics Area Conference in Munich 2012; the Economic Science Association World Meeting in New York 2012; Victoria University Wellington; Queensland University of Technology, Brisbane; Zeppelin University; the Science of Philanthropy Conference 2013, Chicago; the University of Stavanger; Aarhus University; the Econometric Society Meeting 2014, Philadelphia; and the Frankfurt School of Finance and Management.

[A] Department of Economics, University of Munich, Geschwister-Scholl-Platz 1, D-80539 Munich, Germany. E-mail: martin.kocher@lrz.uni-muenchen.de.

[B] Department of Economics, University of Gothenburg, Box 640, SE-40530 Gothenburg, Sweden. E-mail: peter.martinsson@economics.gu.se and conny.wollbrant@economics.gu.se.

[C] School of Management, University of St Andrews, St Andrews, KY16 9RJ UK. E-mail: kom@st-andrews.ac.uk.

[D] School of Economics and Finance, Queensland University of Technology, Gardens Point Campus, 4000 Brisbane, Australia.

1. Introduction

A substantial body of evidence from laboratory and field studies documents that people cooperate more in public goods games than implied by the selfish free-riding equilibrium (for surveys, see e.g., Ledyard, 1995; Zelmer, 2003; Gächter, 2007; Chaudhuri, 2011). Over the past 25 years, multiple explanations have been proposed for the observed levels of cooperation in social dilemmas—among them, confusion, altruism, warm-glow, inequity aversion, efficiency preferences, and reciprocity (see, for instance, Andreoni, 1990, 1995; Palfrey and Prisbrey, 1997; Anderson *et al.*, 1998; Houser and Kurzban, 2002). A more recent approach to public goods experiments, pioneered by Fischbacher *et al*. (2001), has focused on classifying individuals as types of contributors (see also Kelley and Stahelski, 1970; Andreoni, 1988; Keser and van Winden, 2000; Fehr and Schmidt, 2006). The most prominent types in such public goods experiments are conditional cooperators, who increase their contribution with the (expected) contribution of other group members; free-riders, who do not contribute at all; and triangle (hump-shaped) contributors, who increase their contributions to the public good up to a certain level of (expected) others' contributions and then reduce them (see, Kocher *et al.*, 2008; Herrmann and Thöni, 2009; Fischbacher and Gächter, 2010; Volk *et al.*, 2012; Martinsson *et. al*., 2013). Despite small differences, the overall distribution of types is surprisingly robust across studies and locations, with conditional cooperators representing the most frequent type (usually around half of the decision makers or more), followed by free-riders (around 20 %-30 %), and triangle contributors.

However, discrepancies still remain between empirical results and existing theoretical frameworks. For instance, a majority of decision makers contribute intermediate amounts, whereas linear models of other-regarding preferences predict corner solutions (e.g., Fehr and Schmidt, 1999). Though one may solve the problem by assuming non-linear forms of other-regarding preferences, many existing models have difficulties in explaining other stylized facts from public goods experiments, such as—in its repeated version—the decay of contributions over time, or the so-called "re-start effect."

The aim of our paper is to contribute to the literature on motives for cooperation in social dilemmas. We present a model of rational self-control, which captures the conflict between pro-social and self-interested behavior—and which lends itself to straightforward application in social dilemmas. The model relies on two main ingredients: other-regarding preferences and a self-control cost, based on the dual-self model of Fudenberg and Levine (2006). Furthermore, we model the conflict between free-riding and contributing to the public good as a two-stage

cognitive problem, with an identification stage and a contribution stage, at which, willpower and the average contribution of other group members jointly determine the individual's contribution level. Our model captures the notion that individuals may feel tempted to act in self-interest, while simultaneously holding a "better judgment" to cooperate. This could be thought of as a conflict between an impulse of greed and a commitment to abide by a social norm. Accordingly, we generate two propositions: (1) decision makers, who identify self-control conflict, cooperate more with higher levels of self-control, and (2) this relationship is attenuated for higher levels of risk aversion.

We test our model in the laboratory by implementing a linear public goods game, after which we measure trait self-control, the perception of conflict, and risk preferences. In line with our predictions, players who reported that they experienced conflict contributed significantly more if their level of self-control was high. Moreover, controlling for self-control levels, a higher level of risk aversion is associated with lower levels of contributions. That is, more risk-averse individuals were more likely to avoid incurring the costs of self-control effort to behave cooperatively. Finally, free-riders were much less likely to experience conflict than were conditional cooperators.

The remainder of the paper is organized as follows. Section 2 discusses models and motives of cooperation in public good games, and Section 3 relates the concept of self-control to cooperation. Section 4 introduces our model, and Section 5 describes our experimental design. We present in Section 6 our experimental results. Section 7 discusses our findings and concludes the paper.

2. Explaining cooperation in social dilemmas

A rudimentary overview of the motivations for cooperation in social dilemmas may be organized along three lines: (i) other-regarding preferences about outcomes; (ii) other-regarding preferences about intentions (reciprocity) and strategic motivations; and (iii) bounded-rationality. Notably, existing models struggle to account for all empirical regularities.

Other-regarding preferences about outcomes—such as altruism, inequity aversion or efficiency concerns—were among the first explanations of positive contributions in public goods games, together with warm glow, which is an explanation in itself (for the early literature, see Andreoni, 1990, 1995; Palfrey and Prisbrey, 1997; Anderson *et al.*, 1998). However, existing models of other-regarding preferences about outcomes—due to their linearity—struggle to

explain intermediate contribution levels. Thus, they cannot easily account for decay of cooperation over time, in repeated games.

Other-regarding preferences about intentions have become popular. Essentially, there are two strands to this literature. One, going back to Kreps *et al.* (1982), addresses strategic reputation. The basic notion is that players seek to establish a cooperative reputation by providing positive contributions in early rounds, and so self-interested players have an incentive to manage the beliefs of cooperative group members. This mechanism applies in repeated interaction. A more recent framework assumes heterogeneity in other-regarding preferences. Players want to reciprocate the expected positive contributions of others (e.g., Keser and van Winden, 2000; Fischbacher *et al.*, 2001; Croson, 2007). For example, Croson (2007) finds evidence supporting impure altruism (e.g., Andreoni, 1990) and (simultaneous) reciprocity (e.g., Sugden, 1984) over models of (Kantian) commitment (e.g., Laffont, 1975; Harsanyi, 1980). Specifically, she finds that a given player's contribution correlates positively with the amount contributed by other players, ruling in favor of a reciprocity model. However, her regressions also yield a positive intercept, which points toward an altruism model. Such reasoning works both in a one-shot and in repeated environments, and Fischbacher and Gächter (2010) show that preferences for cooperation elicited in a one-shot game can predict the dynamics of contributions in a repeated game. That is, the interaction between free riders and conditional cooperators, or even the interaction between conditional cooperators with a self-serving bias (matching others' contributions imperfectly), can account for the decay of contributions over time.

One recent theoretical contribution that combines the two strands is that of Ambrus and Pathak (2011), which integrates the two aforementioned strands of other-regarding preferences about intentions. The model assumes a mixed population of players, who are either self-interested or reciprocal. In repeated games, self-interested players may contribute initially to induce reciprocal players to contribute in return. Thus, mean contributions decay over time, as the end of play approaches and the incentive for selfish players to induce cooperation fades. Restarting the game restores incentives for selfish players to contribute once more. The model, however, requires repeated interaction to allow for cooperation, as does the explanation provided by Kreps *et al.* (1982). Models of strategic cooperation thus fail to account for cooperation in one-shot public goods games—which is widely observed.

The third class of models requires some level of bounded rationality. The early contributions were based on the concept of "confusion." Participants in public goods games were thought to be confused in early rounds, resulting in over-contributions, but they would learn the optimal strategy as play progressed. Thus, their contributions would drop with rounds played.

This explanation, however, is hard to reconcile with the oft-observed slow decay. It is also inconsistent with the so-called "restart effect;" after contributions have decayed over a pre-announced number of rounds—whereupon subjects are unwittingly invited to play the game once more—the contribution pattern repeats itself (e.g., Andreoni, 1988; Andreoni, 1995; Houser and Kurzban, 2002; Neugebauer *et al.*, 2009).

Several additional recent models take "confusion" into account by explicitly modeling the extent of bounded rationality. In Kandori (2002), for example, the individual gains utility over material payoffs, as well as a "psychological utility," which arises from a desire to follow the norm, captured by the median behavior of others. In addition, individual behavior is subject to random shocks, which account for the decay of contributions over time. In Figuières *et al.* (2013), players have in mind an ideal contribution level (as in, Nyborg, 2000; Brekke *et al.*, 2013), but as play progresses over the rounds, they adjust contributions downwards, towards the mean contributions of others. In both cases, the extent of bounded rationality—random shocks and non-equilibrium beliefs—is relatively small. Klumpp (2012) allows a greater deviation from classical rationality, by assuming that players are satisficers, content with reaching a contribution level slightly below their optimum. The model consists of an additively separable utility function, with "material utility" and "psychological utility," the latter of which depends on the average contribution of others. Thus, the model accounts for the decay of cooperation over time and the re-start effect.[1]

Similar to many of the aforementioned approaches, our model features a utility function with an other-regarding preference component, combined with a specific form of bounded rationality: the self-control problem. More specifically, our model contains a cost of deviating from the mean contribution of others. One may interpret this term as a preference for reciprocity, or a desire to match a norm. The sensitivity to this cost is measured by a parameter that is interpreted as one dimension of a decision maker's type. We assume that individuals are tempted to behave in a self-interested manner and may therefore not fully match the contributions of others. How closely players match the mean contribution of others also depends on another type dimension, namely willpower. Willpower is captured by a parameter that measures how costly it is for the individual to overcome selfish urges. The combination in our model, of other-regarding preferences and the self-control problem, can account for all aforementioned empirical regularities observed in public goods games. Although we do not apply our model to the repeated

[1] There is also an extensive literature on neurophysiological foundations of cooperation and on emotions and punishment (e.g., Joffily *et al.*, 2014; Boyce *et al.*, 2015; Dickinson and Masclet, 2015).

setting, it would be straightforward to extend it to account for decay of contributions and the re-start effect.

3. Self-control and cooperation

There is a growing empirical literature on the relationship between cooperation and constructs closely related to self-control. Roughly speaking, we can organize this literature according to three types of psychological constructs studied: (i) time preferences; (ii) intuitive versus reflective responses; and (iii) trait self-control. Each of the constructs captures important aspects of generally accepted conceptualizations of the self-control problem. A typical conceptualization—adopted in this paper—views the self-control problem as an intra-personal conflict between "better judgment" and "temptation" (e.g., Thaler and Shefrin, 1981; Schelling, 1984; Loewenstein, 1996). This view is consistent with a variety of common modeling approaches (e.g., Ainslie, 1992; Laibson, 1997; Gul and Pesendorfer, 2001; Bénabou and Tirole, 2002; Bodner and Prelec, 2003; Bénabou and Tirole, 2004; Battaglini *et al.*, 2005; Fudenberg and Levine, 2006; Myrseth and Wollbrant, 2013).

An early paper that addresses the relationship between cooperation and self-control considered time preferences. Curry *et al.* (2008) find that individuals' discount rates are negatively associated with their contributions to the public account. That is, more impatient individuals contribute less to the public good than do patient ones. Fehr and Leibbrandt (2011) combine laboratory data on time preferences, as well as extraction in a common pool resource problem, with field data on the catches of fishermen in Brazil. Their data indicate that those in the experiment who exhibited less impatient behavior were in the field less likely to over-exploit the common pool resource, but in a laboratory study no more or less likely to cooperate. Consistent with the aforementioned results, Burks *et al.* (2009) report a positive association between "short-term" patience—the β in the β-δ model—and cooperative behavior in a sequential prisoner's dilemma. Houser *et al.* (2012) subjected children to a common pool resource problem, in which a delay-of-gratification task (an analogue to the classic "marshmallow" problem) represented the resource extraction. They find that younger children were more likely to extract the resource than were older children, who are presumed more able to exercise self-control (e.g., Mischel and Metzner, 1962). Taken together, the existing literature on cooperation and time preferences would be consistent with the notion that self-control benefits cooperation.

However, studies that explore whether cooperation derives from intuitive or deliberative thought processes—that is, whether cooperation represents a "default" response—paint a conflicting picture. Building on dual process theories (for a review, see Alós-Ferrer and Strack, 2014), these studies typically rely on the measurement or manipulation of decision times, on the assumption that more intuitive responses are quicker. Several papers report negative associations between decision times and cooperation in social dilemmas (Rand *et al.*, 2012; Lotito *et al.*, 2013; Rand *et al.*, 2014), and some also find that inducing people to decide quicker causes them to cooperate more (Rand *et al*., 2012; Rand *et al*., 2014). Furthermore, Nielsen *et al.* (2013) report that free-riders, classified according to the Fischbacher *et al.* (2001) taxonomy, exhibited shorter decision times than did conditional cooperators. However, Tinghög *et al*. (2013) find null effects of time pressure and one negative effect, notably, after excluding participants who failed comprehension but including those who disobeyed the time constraint. Verkoeijen and Bouwmeester (2014) report null effects, and Lohse *et al.* (2014) find that decision times are positively associated with cooperation, both for within- and between-individual comparisons. More recently, Myrseth and Wollbrant (2015a) re-examine the data from Rand *et al.* (2012) and Rand *et al*. (2014). They argue that the vast majority of the cooperation decisions in the two papers are too slow to allow discrimination of intuitive response from deliberative decision.[2]

The empirical part of this paper belongs to a third category, which examines the association between trait self-control and cooperation. Martinsson *et al.* (2014) implemented in a public goods game framing treatments that were intended to raise or lower the likelihood that individuals identify a self-control conflict. The idea was that individuals may or may not perceive a self-control conflict between urges to act in self-interest and better judgment to cooperate—and that they would engage self-control for the purpose of cooperating only if they have identified a self-control conflict in the first place. Consistent with their predictions, from a simplified version of the model presented in this paper, individuals in the treatment intended to raise the likelihood of conflict identification exhibit a positive correlation between trait self-control and cooperation, but those in the treatment intended to lower the likelihood exhibit no correlation. Myrseth *et al.* (2015) manipulated the degree to which the endowment in the public goods game was

[2] The pattern emerging from dictator games largely mirrors the conflicting pattern from public good games. Piovesan and Wengström (2009) find that selfish choices in a repeated dictator game are correlated with lower response times. Studies that manipulate cognitive resources through depletion or load have yielded mixed results, namely both evidence for and against the proposition that giving requires deliberation (e.g., Hauge *et al.*, 2009; Cornellisen *et al.*, 2011; Schulz *et al.*, 2014; Achtziger *et al.*, 2015).

represented in tangible (i.e., cash or tokens) or abstract form (i.e., on a computer screen, as is standard). They find a positive correlation between trait self-control and cooperation when the endowment was represented in tangible form, but no correlation when the endowment was represented abstractly.

4. The Model

Following Myrseth and Fishbach (2009), we propose a two-stage model with a conflict identification stage (perception of conflict) and a contribution stage (resolution of the conflict). In the model, nature decides in the first stage whether or not an agent identifies conflict between the selfish impulse and the better judgment to cooperate (for instance, following a social norm). If conflict is not identified, the decision process ends, and the agent contributes zero.[3] If identified, the agent decides how much effort to invest into self-control effort.

More formally, we assume that the utility function *U* of individual *i* is given by

$$U_i = u(\pi_i) - k - s. \tag{1}$$

Here, π_i is the individual*'s monetary payoff,* which depends on the public good technology, the relative price of the private good, and on the contribution costs to the public good (i.e., $\pi_i = f(c_i, G(\cdot))$ and $\frac{\partial \pi_i}{\partial c_i} < 0$; where *G* denotes the public technology); $c_i \geq 0$ is individual $i's$ contribution to the public good from the available endowment (and the rest is left for the consumption of the private good). With respect to the public good technology, we assume that

$$\frac{\partial G(\cdot)/\partial c_i}{n} < \partial G(\cdot)/\partial c_i < 1 < n \cdot \partial G(\cdot)/\partial c_i, \partial G(\cdot)/\partial c_i > 0, \tag{2}$$

with *n* representing the group size. Condition (2) ensures for the monetary maximizing individual that the problem constitutes a social dilemma—as the selfish individual optimum and the

[3] It is possible, for a variety of reasons, that a decision maker by default contributes a positive amount. In the spirit of parsimony and modeling convenience, we abstract from such cases, but we shall revisit this point in the Experimental Results section.

collective optimum are in conflict. The function $u(\pi_i)$ is strictly concave, i.e., $u'(\pi_i) > 0$ and $u''(\pi_i) < 0$. One can interpret the concavity of the utility function for monetary payoffs as diminishing marginal utility or risk aversion, but one can also view it as a weight for utility derived from monetary payoffs vis-à-vis the costs *k* and *s*. For simplicity, we will speak of concavity as risk aversion.

The second term of the utility function, $k = \frac{\beta_i}{2}(m - c_i)^2$, is the cost of deviating from others' average contributions to the public good, where the average contribution of others is denoted *m;* the parameter $\beta_i \geq 0$ captures individual sensitivity to this difference. Our utility function thus incorporates other-regarding concerns in a manner similar to that of warm-glow, altruism, and reciprocity models.

Finally, s represents the "opportunity-based" specification of self-control cost, from Fudenberg and Levine (2006). We assume that the "selfish self" (in their terminology, the short-run self) is purely selfish and therefore will only maximize monetary payoff from the public good, implying a zero contribution in the linear mechanism. The potentially "pro-social self" (Fudenberg and Levine's long-run self), however, solves a maximization problem that contains all attributes in U_i and may therefore decide on a positive contribution, $c_i > 0$. Self-control cost in this framework is proportional to the difference in payoffs resulting from the payoff maximizing contribution (zero) and from the actual contribution. The cost of contributing for the pro-social (long-run) self, therefore, is $\pi_i(0) - \pi_i(c_i)$, which is a cost (non-negative) whenever $0 \neq c_i$. To account for individual differences in self-control cost, we divide this expression by $\omega_i > 0$, the individual willpower parameter. Hence, the cost of self-control for the utility maximizing agent becomes $s = [\pi_i(0) - \pi_i(c_i)]/\omega_i$.

The decision problem of the agent is straightforward. Nature exogenously determines whether or not the agent identifies conflict. The binary identification function $\varphi = \{0,1\}$ implies conflict identification when $\varphi = 1$, and no conflict identification when $\varphi = 0$. The model has a trivial solution in the case of $\varphi = 0$; the agent does not identify conflict, maximizes monetary payoff, and hence contributes nothing (see footnote 3). If the agent identifies conflict ($\varphi = 1$), however, she proceeds to the conflict stage and maximizes U_i with respect to c_i. We derive the following propositions. All proofs can be found in Appendix D.

PROPOSITION 1: *Given that the agent has identified conflict* ($\varphi = 1$) *and* ω_i *and* β_i *are sufficiently large*, *raising willpower increases contributions.*

If the agent has identified conflict, she needs to determine how much she wishes to contribute. If she is sufficiently pro-social, this will be a positive contribution. Contributing, however, is costly both in terms of money and in terms of self-control. An increase in willpower reduces marginal self-control cost of contributing and therefore raises contributions.

PROPOSITION 2: *Given that the agent has identified conflict* ($\varphi = 1$) *and* ω_i *and* β_i *are sufficiently large*, *raising willpower leads to a smaller increase in contributions if risk aversion is high.*

As willpower increases, the marginal self-control cost of contributing decreases, which increases contributions—given that the individual has identified self-control conflict, and is sufficiently pro-social to prefer a positive contribution level. The reduction in marginal self-control cost resulting from an increase in willpower, however, diminishes if risk aversion increases (concavity increases). Consequently, there is a smaller increase in contributions. We therefore expect an interaction between willpower and risk preferences on contributions. The more concave the utility function is, the smaller is the positive effect of willpower on contributions.[4]

Having outlined the model and derived the main predictions, we next present the experimental design and procedure.

5. Experimental Design and Procedure

5.1 The Basic Public Goods Game and the Strategy Vector Method

In our experiment, the public goods game builds on the following linear payoff function for individual *i*

$$\pi_i = 20 - c_i + 0.4\sum_{j=1}^{4} c_j, \tag{3}$$

[4] The effect in Proposition 2 holds when risk aversion refers to the concavity in the utility of monetary payoffs.

where c_i denotes the contribution of individual *i* to the public good. Each group consists of four randomly matched individuals, and each individual receives an endowment of 20 experimental points (the experimental currency unit). The marginal per capita return (MPCR = $\partial G(.)/\partial c_i$) from investing in the public good is 0.4, fulfilling the conditions for a social dilemma. Assuming that participants are rational and self-interested, any MPCR < 1 yields a dominant strategy to free-ride. From the perspective of social welfare, it is optimal to contribute the entire endowment because MPCR·n > 1.

The preference elicitation and the incentive mechanism in our experiment closely follow Fischbacher *et al.* (2001). More specifically, participants are asked to make two decisions: first, to make an unconditional contribution to the public good, and, then, to submit a conditional contribution schedule. The unconditional contribution is a single integer number satisfying $0 \leq c_i \leq 20$. For the conditional contribution, participants indicate how much they would contribute to the public good for any possible average contribution (rounded to integers) of the other three players in their group. For each of the 21 possible averages from 0 to 20, participants must decide on a contribution between (and including) 0 and 20. This is a variant of the strategy vector method (Selten, 1967).

To ensure incentive-compatibility, both the unconditional and the conditional contributions are potentially payoff-relevant. For one group member, randomly determined by the toss of a four-sided die, the conditional contribution is relevant; unconditional contributions are relevant for the other three group members.[5] More specifically, the three unconditional contributions from a group, and the corresponding conditional contribution (for the specific average of the three unconditional contributions), determine the sum of contributions to the public good. One can then compute individual earnings, according to equation (3).

In addition, participants are asked to guess the average unconditional contribution of the other three group members (rounded to integers). The guessing stage is implemented after the contribution stages and is not mentioned in the written instructions. As in Gächter and Renner (2010), participants are monetarily rewarded depending on the accuracy of their guesses. However, we use a slightly stronger incentive mechanism. If a participant's guess equals exactly the average unconditional contribution of the other three group members, the participant earns nine additional points from the guess; if there is a difference of one between the guess and the

[5] Each group member is assigned a number from one to four. At the end of the experimental session, and monitored by the experimenter, a randomly selected group member rolls the die.

average, the participant earns six additional points; and a difference of two results in additional three points earned. Larger differences are neither rewarded nor punished.

5.2 Elicitation of Risk Preferences

We employed the design by Holt and Laury (2002) to measure individual risk preferences. Each participant, without interacting with any other participant, is required to make ten risky choices. For each choice, participants choose between two options, labeled X and Y. Both options include a lottery with the same probabilities, but with different payoffs. Option X is the relatively safer option; its highest outcome is lower than the highest outcome from option Y, but its lowest outcome is higher than the lowest outcome from option Y. Payoffs are fixed throughout the choice sequence. However, in both options the probability of receiving the higher payoff increases by ten percentage points, from 10% in decision 1 to 100% in decision 10.[6]

As the participant moves down the sequence of choices, depending on the participant's preference for risk, the participant at some point may switch from Option X (the relatively safe choice) to Option Y (the relatively risky choice). In the case of extreme risk-loving, the participant would always choose Option Y. Switching from Y to X, or always choosing X is incompatible with consistent money-maximizing behavior. One can compute an individual's degree of risk aversion by using the point at which she switches from Option X to Option Y.[7] Upon completing this task (and the rest of the experiment), one of the ten lotteries is selected randomly and played for real. All lotteries are thus potentially payoff-relevant, and participants could in this part earn up to 3.85 euro.

5.3 Measurement of Conflict Identification and of Trait Self-Control

After risk preference elicitation, we implement a standard measure of trait self-control: the Rosenbaum Self-Control Schedule (Rosenbaum, 1980a), henceforth abbreviated RSS.[8] This measure has been validated against a battery of relevant personality measures, and against behavioral tasks associated with self-control, such as resisting pain (Rosenbaum, 1980b), coping

[6] We provide the specific numbers used for this risk elicitation procedure in Appendix C.

[7] Switching points can readily be converted into risk aversion parameters of parametric models, such as CRRA. As the choice of a model would be arbitrary, we use the switching point in our analysis.

[8] The Rosenbaum Self-Control Schedule (1980a) is included in Appendix A. We translated this scale to German; the same translation was used in Myrseth *et al.* (2015).

with stress (Rosenbaum and Smira, 1986; Rosenbaum, 1989), coping with mental disability (Rosenbaum and Palmon, 1984), coping with seasickness (Rosenbaum and Rolnick, 1983), quitting smoking (Katz and Singh, 1986), saving over spending (Romal and Kaplan, 1995), and curtailing procrastination (Milgram *et al.*, 1988).

We build on the finding from personality psychology that the tendency to apply self-control strategies represents a stable trait within the individual over time. Indeed, the tendency to apply self-control strategies remains remarkably consistent throughout life. For example, Mischel and colleagues found that a child's performance at age 4 on an instant gratification task (one cookie now, or two cookies later) predicted later in life their cognitive control (Eigsti *et al.*, 2006), ability to concentrate, self-control, interpersonal competence, SAT scores, and their drug use (Mischel *et al.* 1988; Mischel *et al.*, 1989; Shoda *et al.*, 1990; Ayduk *et al.*, 2000).

Critically, self-control strategies are relevant to the decision to indulge only when the individual has identified self-control conflict.[9] Therefore, one approach to investigating whether the problem of pro-social versus selfish behavior resembles that of self-control is to test whether self-control strategies are positively associated with pro-social behavior when the individual has felt conflicted, but less so or not at all when the individual has not. It would be appropriate, therefore, to measure experienced conflict. To capture recollection of feelings of mixed emotion, we posed in the last part of the experiment (but before administering the RSS) a question similar to that used in Aaker *et al.* (2008): *"To what extent did you experience conflict when deciding how much to contribute?"* Participants answered this question on a continuous scale ranging from 0 ("not at all") to 100 ("very much").[10]

5.4 Experimental Procedure

The computer-based experiment was conducted at the experimental laboratory MELESSA of the University of Munich in October 2009 and in March 2010, using the experimental software z-Tree (Fischbacher, 2007) and the organizational software Orsee (Greiner, 2004). A total of 144

[9] Such self-control strategies may take a variety of forms, and common examples include counteractive self-control (e.g., Trope and Fishbach, 2000; Myrseth *et al.*, 2009) and pre-commitment (e.g., Schelling, 1984).

[10] Note that the original German question clearly hinted at the normative conflict, without being too suggestive. After asking the experimental participant to recollect his or her decision about contributions, the following question was posed: "In welchem Maße fühlten Sie sich bei Ihrer Entscheidung in einem (inneren) Zwiespalt?" The term "Zwiespalt" can also be translated to English as "dilemma".

undergraduate students from all disciplines, except economics, participated in six sessions, each with 24 participants. Approximately 62% of participants were female. Sessions lasted up to 1½ hours, and the average payoff was 13.4 euro, including a show-up fee of 4 euro.[11]

Upon arrival, experimental participants were seated in separate cubicles. Each session started with instructions for the public goods game. At this stage it was made clear that there would be additional parts of the experiments, but that the instructions for those parts would only be handed out after the completion of the current part. It was also emphasized to participants that decisions in one part would be completely unrelated to those in the other parts. Participants received neutrally-framed, written instructions (see Appendix B), which were read aloud to ensure common knowledge.[12] Everybody had the opportunity to ask questions in private. The experiment continued only after all participants had completed a series of computerized exercises (where they calculated profits for different contribution levels in the public goods game), and after all participants had correctly understood the procedures. It was made very clear that feedback and profit information would only be given at the end of the experiment. This was done to reduce the potential spillover effects of earnings, from one part of the experiment to the next.

Upon completing the public goods game (part 1), participants received instructions for the risk preference elicitation (part 2) and a variant of the trust game (part 3).[13] Following part 3, participants answered the conflict experience question, the RSS, and some questions about socio-demographics and individual background. The final stage of the experiment included feedback on the decisions of group members in the public goods game, chance moves, and the individual earnings. Payments were made privately and in cash.

6. Experimental Results

We hypothesized that self-control would positively correlate with contributions to the public good for individuals who had identified a self-control conflict between better judgment to cooperate and the temptation to act in self-interest. We did not expect a significant correlation for

[11] Each experimental point earned in the public goods game is exchanged at the pre-announced rate of 1 point = 0.33 euro.

[12] All instructions, written and oral, were given in German. English versions are included in this paper.

[13] Kocher *et al*. (2011) analyze the association between cooperation, trust, and risk (but not self-control) based on these data.

individuals who had not identified conflict. The RSS represents our proxy for self-control, and a dummy variable, extracted from participants' self-reports of conflict intensity, represents our proxy for identification of self-control conflict.

While the response variable for conflict intensity is continuous, there is no reason to expect a linear effect of experienced conflict on the impact of trait self-control. Rather, a threshold effect of the former on the latter seems more appropriate; individuals who identified self-control conflict would draw on their self-control strategies to promote pro-social behavior, whereas others would not. A natural, theoretically motivated threshold for our analysis, therefore, would be the lowest positive, non-zero report of experienced conflict (identification). Accordingly, our conflict dummy takes the value of zero for participants reporting no conflict ("0" on the conflict intensity question), and 1 otherwise. Our subsequent pattern of results also holds qualitatively when the threshold for conflict identification is set at the median conflict response ("21" on the conflict intensity question), which corresponds to the midpoint of the response distribution. And for both conflict definitions, our results hold when including the continuous conflict measure as a control variable.

Insert Table 1 around here

Our full sample consists of 144 subjects, but 15 provided inconsistent answers in the Holt-Laury-task. In our subsequent analyses, we have included all subjects and added a dummy to control for inconsistency.[14] The summary statistics in Table 1 reveal that the average unconditional contributions in our sample, approximately 34% of the endowment, resemble those reported in the related literature (e.g., Fischbacher *et al.*, 2001; Fischbacher and Gächter, 2010). Moreover, the RSS scores of our participants appear roughly similar to those found in other studies: the standard deviation is within the range of those found in the original samples studied by Rosenbaum (1980a, b), but the mean is slightly below the corresponding range of means (16.7 vs. a range of 23 to 27). Overall, our summary statistics are in line with previous findings.

[14] The results do not depend on the inclusion of inconsistent responses; the pattern remains the same when we include only those who provided consistent answers in the Holt-Laury task.

6.1 Conditional Contributions in the Public Goods Game

We start by examining contribution schedules. Recall that each of our participants had to indicate 21 contribution levels for all possible average contribution levels (rounded to integers) of the other group members. The elicitation of the schedule was fully incentivized. Table 2 provides test results for our two propositions on the conditional contribution data. More specifically, it presents a tobit analysis of conditional contributions as a function of RSS scores (denoted *RSS*), risk preferences based on the switching point in the choice list task (denoted *Risk*), average contributions of others (denoted *Others*), the respective interaction terms, socio-demographic controls (gender, age, and an income proxy), and the inconsistency dummy.[15] We have split the estimations based on whether an individual has identified self-control conflict (specifications (6)-(10), based on 108 individuals) or not (specifications (1)-(5), based on 36 individuals).[16] Specifications (1)-(4) and (6)-(9) all replicate a commonly found pattern: the level of others' average contributions is a strong determinant of the decision maker's own contributions (e.g., Gächter, 2007; Kocher *et al.*, 2008; Fischbacher and Gächter, 2010). It is noteworthy that this variable does not appear significant in specifications (5) and (10), a point to which we shall return.

Insert Table 2 around here

Consistent with Proposition 1, specifications (7) and (9) yield positive and significant correlations between conditional contributions and RSS for those who have identified conflict; no such positive correlation is obtained in specifications (2) and (4) for those who reported not having identified conflict. Moreover, and consistent with Proposition 2, specification (9) yields a negative coefficient on the interaction term between RSS and risk preferences for those who have identified conflict; this coefficient, however, is not significant at conventional levels. We summarize our findings for conditional contributions in Results 1 and 2—corresponding to Propositions 1 and 2, respectively.

---

[15] We present here and hereafter only tobit regressions, which account for the lower and the upper contribution limits, but our results also hold for OLS. Regression tables are available upon request.

[16] Note that we have 21 observations per individual, and we report robust standard errors to account for the dependence in the data.

RESULT 1: *Conditional contributions are positively correlated with self-control, for individuals who have experienced conflict.*

RESULT 2: *The positive correlation between conditional contributions and self-control diminishes weakly, but not significantly, as risk aversion increases, for individuals who have experienced conflict.*

Specification (10) pertains to individuals who identified self-control conflict, and it includes all aforementioned variables and the respective interactions. Empirically, these interaction terms may matter. This is especially true for others' average contributions, which may influence conditional cooperators. Indeed, the interaction between RSS and Others is positive and significant. That is, the greater is the level of others' average contributions, the stronger the positive association between self-control and conditional contributions. This means that higher contributions of others make it easier for conditional cooperators to overcome the temptation to free-ride. This result is not obtained for specification (5), which includes only those who did not identify self-control conflict.

RESULT 3: *The positive correlation between conditional contributions and self-control becomes stronger as the level of others' average contribution increases, for individuals who have experienced conflict.*

Furthermore, the three-way interaction between RSS, risk preferences, and Others is negative and significant. In other words, with a higher level of others' average contributions, there is a weaker association between self-control and conditional contributions for more risk-averse individuals. Again, we do not obtain the result from specification (5), which includes only those who did not experience self-control conflict.[17] We summarize the finding in Result 4.

[17] We have for expositional purposes decided to split the data according to conflict identification. When instead aggregating the data and including in the specifications a dummy for conflict identification, the same patterns obtain. When we interact the conflict dummy with the relevant variables, the interactions are significant and confirm the results in Tables 2 and 4. However, such specifications are more cumbersome to interpret, in particular the four-way interaction between conflict identification, RSS, Risk, and Others.

RESULT 4: *With higher average contributions of others, the strength of the positive correlation between conditional contributions and self-control diminishes with higher levels of risk aversion, for individuals who experienced conflict.*

In order to illustrate the results from specification (10), which includes those who did experience self-control conflict, we plot in Figure 1 the unit increase in conditional contribution, from a one-standard-deviation increase in RSS, at different levels of risk preferences and others' contributions. At low levels of Others, there is little difference in conditional contribution for various levels of risk preferences. Similarly, at high levels of risk aversion, there is little difference in conditional contribution for various levels of Others. However, a one-standard-deviation increase in RSS yields higher levels of conditional contributions when there are both lower levels of risk aversion and higher levels of Others.

Insert Figure 1 around here

The main effect for Others, but also the significant effects of RSS and the interaction of RSS with Risk, statistically disappear in specification (10). It appears that there is no effect of others' contributions, independent of self-control and risk preferences.

6.2 Types of Contributors in the Public Goods Game

We followed the standard approach in classifying four types of contributors (see Fischbacher *et al.*, 2001; Fischbacher and Gächter, 2010). *Conditional cooperators* submitted a contribution schedule displaying a (weakly, with at least one strict step) monotonically increasing contribution for an increasing average contribution of the other group members.[18] *Free-riders* are characterized by zero conditional contributions for every possible average contribution of the other members. *Hump-shape contributors* (also known as *triangle contributors)* exhibit (weakly, with at least one strict step) monotonically increasing contributions up to a certain average level of others' contributions, above which their contributions schedule is (weakly, with at least one strict step) monotonically decreasing. The category referred to as *Residual* constitutes the

[18] We also included those without a weakly monotonically increasing contribution, but with a highly significant (p-value < 0.01) positive Spearman rank correlation coefficient between own and others' contributions, as in Fischbacher *et al.* (2001) and Fischbacher and Gächter (2010).

remaining participants.[19] The distribution of types based on our data, and shown in Table 3, corresponds to those found in past studies (e.g., Fischbacher *et al.*, 2001; Kocher *et al.*, 2008; Herrman and Thöni, 2009; Fischbacher and Gächter, 2010; Martinsson *et al.*, 2013).

Insert Table 3 about here

Given that free-riders by definition contribute less than do other types, and given that they happened to have about the same RSS score, and about the same risk preferences, our model would imply that they were less likely to identify a self-control conflict between keeping the money and contributing.[20] Consequently, we would predict that free-riders were less likely to have drawn on their self-control strategies to promote pro-social behavior. Indeed, consistent with this implication, free-riders reported a significantly lower average level of conflict than did other types (p-values < 0.01; two-sided Mann-Whitney-U-tests). In other words, free-riders seem to have contributed less because they were less likely to see a self-control conflict in the first place and, therefore, they were less likely to draw on their self-control strategies to promote pro-social behavior. We summarize this finding for contributor types in Result 5.[21]

RESULT 5: *Free-riders experience lower levels of conflict than do other types, but they do not exhibit different risk preferences or scores on the self-control measure (RSS).*

Finally, it is worth mentioning that the self-serving bias of conditional cooperators (i.e., the difference between perfect conditional cooperation and the actual conditional contribution of an individual) is related to trait self-control. In a regression also controlling for risk preferences, a higher level of self-control exhibits a strong and significant ($p < 0.01$) negative association with the size of the individual self-serving bias of conditional cooperators.[22]

---

[19] We elect not to label this category *Others*, as is conventional in the literature, because the label would be identical to the one that we have employed in our regression analyses. To avoid confusion, we instead refer to the residual class of contributor types as *Residual*.

[20] The RSS of free-riders is not significantly lower than that of either conditional cooperators or hump-shape contributors (all p-values > 0.4; two-sided Mann-Whitney-U-tests).

[21] The result provides ex-post evidence for the assumption in our model that no conflict identification implies low levels of contribution. See footnote 3.

[22] The regression table is available on request.

6.3 Unconditional Contributions in the Public Goods Game

Our experiment elicited conditional and unconditional contributions to the public good. While we deem the contribution schedule (conditional contributions) essential for testing our hypotheses, examining participants' unconditional contributions can provide valuable robustness checks.

An initial analysis of the data reveals the predicted association between RSS and cooperation also for participants' unconditional contributions. We compare the means of unconditional contributions by high versus low RSS scores (above vs. below the mean) and by experienced conflict versus no conflict. Among participants who reported conflict, those with high RSS scores contributed more (on average, 8.87) than did those with low RSS scores (4.62). The difference is highly significant (p-value < 0.01; Mann-Whitney-U-test). However, among participants who did not report having identified conflict, those with high RSS scores did not contribute significantly more (4.62) than did those with low RSS scores (7.13) (p-value = 0.29; two-sided Mann-Whitney-U-test).

Table 4 presents tobit regressions for unconditional contributions as a function of RSS scores, risk preferences, and the interaction between the two. As with conditional contributions, we have split the estimations based on whether individuals identified self-control conflict (specifications (14)-(16)) or not (specifications (11)-(13)). Consistent with Proposition 1, specifications (14) and (15), which exclude the interaction term, reveal that RSS is positively correlated with unconditional contributions for individuals who identified self-control conflict. However, the corresponding specifications for those who did not identify conflict, (11) and (12), yield a negative and significant correlation between RSS and unconditional contributions. That is, given that they failed to identify a self-control conflict, individuals with higher trait self-control contributed less to the public good than those with lower trait self-control. This result is consistent with the assumption that failure to identify self-control conflict implies self-interested preferences. Higher trait self-control should help the individual to better achieve her own interests, which in this case means cooperating less.

Specification (16) includes the interaction term between RSS and risk preferences, and it includes only those who have identified self-control conflict. Consistent with Proposition 2, the coefficient for RSS is positive and significant, and the coefficient on the interaction term for RSS and risk preferences is negative and significant. In other words, the positive association between RSS and unconditional contributions is weaker for more risk-averse individuals. The corresponding estimation for individuals who did not identify conflict, specification (13),

however, yields the opposite pattern. This is also consistent with the assumption of self-interested preferences among those who fail to identify self-control conflict. Although higher trait self-control helps the individual better achieve her interests, higher payoffs will matter less when the utility function is more concave.

We summarize these findings for unconditional contributions, which reinforce our conclusions from Section 4.1, in results 1' and 2', corresponding to Propositions 1 and 2, respectively.

Insert Table 4 about here

RESULT 1': *Unconditional contributions are positively correlated with self-control, for individuals who experienced conflict.*

RESULT 2': *The positive correlation between unconditional contributions and self-control diminishes with higher levels of risk aversion, for individuals who experienced conflict.*

The analogues of Results 3 and 4 are more difficult to re-test with data on unconditional contributions. Because we have elicited expectations of others' contributions immediately after asking for unconditional contributions, the data are less suited than are the conditional schedules. This is mainly because of a potential influence of unconditional contributions on expectations, for example, through the anchoring or the false consensus effect (Gächter, 2007). Accordingly, when we reproduce specification (10), substituting the expectations variable for Others, the only variables yielding statistical significance are expectations and the control variable for inconsistent responses in the choice list task.

As a final note, the coefficient on Risk alone does not explain unconditional contributions in any significant way.[23]

[23] This result is also reported in Kocher *et al*. (2011).

7. Discussion and Conclusion

Models that posit a self-control problem between self-interest and a better judgment to act pro-socially hold a major advantage over alternative models. They can explain commonly observed cooperation patterns independently of strategic reciprocity, and so they can account for stylized facts arising from both one-shot games and from repeated interaction. Such models can also explain behavior in dictator games, which do not provide any incentive for selfish players to induce generous behavior by others. In this paper, we have developed a self-control model of cooperation, from which we have derived and empirically tested two main predictions. Both address the hypothesis that individuals may experience a self-control conflict between acting in self-interest or in the interest of the common good.

We find support for our predictions. Self-control is positively associated both with conditional and unconditional contributions in a linear public goods game, consistent with Proposition 1. Moreover, there is a weaker association between self-control and unconditional contributions for more risk-averse individuals, consistent with Proposition 2. In addition, we find that higher levels of others' average contributions strengthen the association between self-control and conditional cooperation. Finally, and addressing the widely observed phenomenon of imperfect conditional cooperation, we observe that the self-serving bias decreases in higher levels of self-control.[24] The aforementioned results hold only for individuals who reported feeling conflicted during the allocation task, also in line with our model. Our findings thus corroborate prior evidence for the idea that the social dilemma may be understood as a problem of self-control (e.g., Martinsson *et al.*, 2014; Myrseth *et al.*, 2015; Osgood and Muraven, 2015).

We also study the distribution of contributor types. Our analysis reveals that free-riders are similar to other types, both in their levels of self-control and in their risk preferences, but differ in their reported experience of conflict; free-riders seem to have cooperated less because they were less likely to see a self-control conflict in the first place—and thus less likely to draw on self-control strategies to promote pro-social behavior. This is consistent with findings from Martinsson *et al.* (2014), who used similar experimental procedures, though without measuring risk preferences. A recent mouse-tracking study by Kieslich and Hilbig (2014), however, yields the opposite pattern. They find more conflict for free-riders than for cooperators, but they gauge

[24] Note that our rationale, in principle, also could account for a reduction of contribution levels over and above a certain level of average others' contributions, as observed for hump-shape contributors. It would require, beyond that level, a strictly convex temptation function.

conflict from movements of the mouse cursor on the computer screen rather than from a self-report measure, as we do here.[25] Future research might investigate whether results depend on the method by which conflict is measured.

Aside from addressing theoretically motivated predictions, our study also yields two interesting incidental results. The first concerns the question of altruism. Our main regression specifications (10) and (16) failed to yield positive intercepts. Hence, we find no evidence of altruistic behavior that is independent of our three theoretically motivated determinants of cooperation: self-control, contributions of others, and risk preferences. The second result refers to an analogous finding for the commonly observed main effect of the average contribution of others. Our main regression specification for conditional contributions (10) yields a non-significant coefficient. It appears that the average contribution of others does not influence cooperation independently of our three theoretically motivated determinants. Future research might explore the stability and meaning of these results, across measures and experimental paradigms.

These results notwithstanding, a note of caution is due. Our empirical strategy is based on an analysis of correlations, and one should thus be careful in inferring causality. However, our theory makes clear causal predictions, with which our pattern of correlations is consistent. It is difficult to come up with plausible, parsimonious alternative accounts of our pattern of results, obtained both for conditional and unconditional cooperation, but we do acknowledge that the question of causality merits further investigation. Future studies might, for example, manipulate the independent variables that were measured here. To this end, Martinsson *et al.* (2014) implemented a perceptual framing manipulation to influence identification of self-control conflict in a public goods game. Consistent with our results, they find that the frame hypothesized to promote conflict identification yielded a stronger correlation between cooperation and the Rosenbaum (1980a) measure of self-control than did the frame hypothesized to inhibit identification.

While we have provided evidence for the conceptualization that temptation to act in self-interest may conflict with better judgment to act in the interest of others, we do not wish to overstate the generality of our findings. We have reason to think that our conceptualization applies in situations where feelings of greed dominate those (if any) to act pro-socially. And the standard experimental protocol for the public goods game is a fitting case. Of course, as O'Donoghue and Loewenstein (2007) suggest, and Andreoni *et al.* (2011) imply, the pattern in

[25] See Myrseth and Wollbrant (2015b) for a comment.

other circumstances may reverse. Specifically, when empathetic emotion is particularly strong, individuals may feel tempted to act pro-socially—even knowing that they ought not.

References

Aaker, J., Drolet, A., and Griffin, D. 2008. "Recalling mixed emotions: How did I feel again?" *Journal of Consumer Research,* 35, 268-278.

Achtziger, A., Alós-Ferrer, C., and Wagner, A. K. 2015. "Money, depletion, and prosociality in the dictator game." *Journal of Neuroscience, Psychology, and Economics*, 8, 1-14.

Ainslie, G. 1992. "*Picoeconomics: The strategic interaction of successive motivational states within the person*." Studies in rationality and social change. Cambridge: Cambridge University Press.

Alós-Ferrer, C., and Strack, F. 2014. "From dual processes to multiple selves: Implications for economic behavior." *Journal of Economic Psychology*, 41, 1-11.

Ambrus, A., and Pathak, P. 2011. "Cooperation over finite horizons: A theory and experiments." *Journal of Public Economics,* 95, 500-512.

Anderson, S., Goeree, J., and Holt, C. 1998. "A theoretical analysis of altruism and decision error in public goods games." *Journal of Public Economics,* 70, 297-323.

Andreoni, J. 1988. "Why free ride?: Strategies and learning in public goods experiments." *Journal of Public Economics*, 37, 291-304.

Andreoni, J. 1990. "Impure altruism and donations to public goods: A theory of warm-glow giving." *Economic Journal*, 100, 464-477.

Andreoni, J. 1995. "Cooperation in public goods experiments: kindness or confusion?" *American Economic Review,* 85, 891-904.

Andreoni, J., Rao, J. M., and Trachtman, H. 2011. "Avoiding the ask: A field experiment on altruism, empathy, and charitable giving." Working Paper, UC San Diego.

Ayduk, O., Mendoza-Denton, R., Mischel, W., Downey, G., Peake, P.K., and Rodriguez, M. 2000. "Regulating the interpersonal self: Strategic self-regulation for coping with rejection sensitivity." *Journal of Personality and Social Psychology,* 79, 776–792.

Battaglini, M., Bénebou, R., and Tirole, J. 2005. "Self-control in peer groups." Journal of Economic Theory, 123, 105-134.

Bénabou, R., and Pycia, M. 2002. "Dynamic inconsistency and self-control: A planner-doer interpretation." *Economics Letters,* 77*,* 419-424.

Bénabou, R., and Tirole, J. 2002. "Self-confidence and personal motivation." *Quarterly Journal of Economics*, 117: 871-915.

Bénabou, R., and Tirole, J. 2004. "Willpower and personal rules." *Journal of Political Economy,*112, 848-886.

Bodner, R., and Prelec, D. 2003. "Self-signaling and diagnostic utility in everyday decision making." In: *Collected essays in psychology and economics, vol. I.,* Brocas, I., and Carrillo, J. (eds.), 105-123. Oxford: Oxford University Press.

Boyce, C., Czajkowski, M., Hanley, N., Noussair, C., Townsend, M., and Tucker, S. 2015. "The effects of emotions on preferences and choices for public goods." Working Papers 2015-08, University of St. Andrews.

Brekke, K.A., Kverndokk, S., and Nyborg, K. 2003. "An economic model of moral motivation." *Journal of Public Economics*, 87, 1967-1983.

Burks, S. V., Carpenter, J. P., Goette, L., and Rustichini, A. 2009. "Cognitive skills affect economic preferences, strategic behavior, and job attachment." *Proceedings of the National Academy of Sciences*, 106, 7745-7750.

Chaudhuri, A. 2011. "Sustaining cooperation in laboratory public goods experiments: A selective survey of the literature." *Experimental Economics,* 14, 47-83.

Cornellisen, G., Dewitte, S., and Warlop, L. 2011. "Are social value orientations expressed automatically? Decision making in the dictator game." *Personality and Social Psychology Bulletin*, 37, 1080-1090.

Croson, R. 2007. "Theories of commitment, altruism, and reciprocity: Evidence from linear public good games." *Economic Inquiry*, 45, 199-216.

Curry, O. S., Price, M. E., and Price, J.G. 2008. "Patience is a virtue: Cooperative people have lower discount rates." *Personality and Individual Differences,* 44*,* 780-785.

Dickinson, D. L., and Masclet, D. 2015. "Emotion venting and punishment in public good experiments." *Journal of Public Economics*, 122, 55-67.

Eigsti, I., Zayas, V., Mischel, W., Shoda, Y., Ayduk, O., Dadlani, M. B., Davidson, M. C., Aber, J. L., and Casey, B. J. 2006. "Predictive cognitive control from preschool to late adolescence and young adulthood." *Psychological Science,* 17*,* 478-484.

Fehr, E., and Leibbrandt, A. 2011. "A field study on cooperativeness and impatience in the tragedy of the commons." *Journal of Public Economics,* 95, 1144-1155.

Fehr, E., and Schmidt, K. 1999. "A theory of fairness, competition and cooperation." *Quarterly Journal of Economics,* 114, 817-868.

Fehr, E., and Schmidt, K.M. 2006. "The economics of fairness, reciprocity and altruism – Experimental evidence and new theories." In: *Handbook on the economics of giving,*

reciprocity and altruism, Kolm, S.-C. and Mercier Ythier, J. (eds.), 615-691. Amsterdam: Elsevier.

Figuières, C., Masclet, D., and Willinger, M. 2013. "Weak moral motivation leads to the decline of voluntary contributions." *Journal of Public Economic Theory*, 15, 745-772.

Fischbacher, U., 2007. "z-Tree: Zurich toolbox for ready-made economic experiments." *Experimental Economics,* 10, 171-178.

Fischbacher, U., and Gächter, S. 2010. "Social preferences, beliefs, and the dynamics of freeriding in public good experiments." *American Economic Review,* 100, 541-556.

Fischbacher, U., Gächter, S., and Fehr, E. 2001. "Are people conditionally cooperative? Evidence from a public goods experiment." *Economic Letters,* 71, 397-404.

Fudenberg, D., and Levine, D. 2006. "A dual self model of impulse control." *American Economic Review,* 96*,* 1449-1476.

Gächter, S. 2007. "Conditional cooperation. Behavioral regularities from the lab and the field and their policy implications." In: *Economics and psychology. A promising new cross-disciplinary field,* Frey, B. and Stutzer, A. (eds.), 19-50. CESifo Seminar Series. Cambridge: MIT Press.

Gächter, S., and Renner, E. 2010. "The effects of (incentivized) belief elicitation in public goods experiments." *Experimental Economics*, 13, 364-377.

Greiner, B., 2004. "An online recruitment system for economic experiments." In: *Forschung und wissenschaftliches Rechnen 2003*, Kremer, K. and Macho, V. (eds.), 79-93. GWDG Bericht 63, Göttingen.

Gul, F., and Pesendorfer, W. 2001. "Temptation and self-control," *Econometrica,* 69*,* 1403-1436.

Harsanyi, J. 1980. Rule utilitarianism, rights, obligations and the theory of rational behavior." *Theory and Decision*, 12, 115-133.

Hauge, K. E., Brekke, K. A., Johansson, L-O., Johansson-Stenman, O., and Svedsäter, H. 2009. "Are social preferences skin deep? Dictators under cognitive load." Working Paper in Economics 371, University of Gothenburg.

Herrmann, B., and Thöni, C. 2009. "Measuring conditional cooperation: A replication study in Russia." *Experimental Economics,* 12, 87-92.

Holt, C., and Laury, S. 2002. "Risk aversion and incentive effects." *American Economic Review,* 92, 1644-1655.

Houser, D., and Kurzban, R. 2002. "Revisiting kindness and confusion in public goods experiments." *American Economic Review,* 92, 1062-1069.

Houser, D., Montinari, N., and Piovesan, M. 2012. "Private and public decisions in social dilemmas: Evidence from children's behavior." *PLoS ONE*, 7, e41568.

Joffily, M., Masclet, D., Noussair, C., and Villeval, M.-C. 2014. "Emotions, sanctions and cooperation." *Southern Economic Journal*, 1002-1027.

Kandori, M. 2002. "The erosion and sustainability of norms and morale." Working Paper, CTRJE-F-169, University of Tokyo.

Katz, R. C., and Singh, N. 1986. "A comparison of current smokers and self-cured quitters on Rosenbaum's self-control schedule." *Addictive Behaviors,* 11, 63-65.

Kelley, H., and Stahelski, A. 1970. "Social interaction basis of cooperators' and competitors' beliefs about others." *Journal of Personality and Social Psychology*, 16, 66-91.

Keser, C., and van Winden, F. 2000. "Conditional cooperation and voluntary contributions to public goods." *Scandinavian Journal of Economics,* 102, 23-39.

Kieslich, P. J., and Hilbig, B. E. 2014. Cognitive conflict in social dilemmas: An analysis of response dynamics. *Judgment and Decision Making*, 9, 510-522.

Klumpp, T. 2012. Finitely repeated voluntary provision of a public good. Journal of Public Economic Theory, 14, 547-572 .

Kocher, M. G., Cherry, T., Kroll, S., Netzer, R. J., and Sutter, M. 2008. "Conditional cooperation on three continents." *Economics Letters,* 101, 175-178.

Kocher, M. G., Martinsson, P., Matzat, D., and Wollbrant, C. 2011. "The role of beliefs, trust and risk preferences in contributions to a public good." Working Papers in Economics 482, University of Gothenburg.

Kreps, D., Milgrom, P., Roberts, J., and Wilson, R. 1982. "Rational cooperation in the finitely repeated prisoners' dilemma." *Journal of Economic Theory*, 27, 245-252.

Laffont, J.-J. 1975. "Macroeconomic constraints, economic efficiency and ethics: An introduction to Kantian economics." *Economica*, 42, 430-437.

Ledyard, O. 1995. "Public goods: Some experimental results." In: *Handbook of experimental economics*, Kagel, J. and Roth, A. (eds.), Chapter 2. Princeton: Princeton University Press.

Lohse, J., Goeschl, T., and Diederich, J. 2014. "Giving is a question of time: Response times and contributions to a real world public good." Discussion Paper Series No. 566, Department of Economics, University of Heidelberg.

Lotito, G., Migheli, M., and Ortona, G. 2013. "Is cooperation instinctive? Evidence from the response times in a public goods game." *Journal of Bioeconomics*, 15, 123-133.

Martinsson, P., Myrseth, K. O. R., and Wollbrant, C. 2012. "Reconciling pro-social vs. selfish behavior: On the role of self-control." *Judgment and Decision Making*, 7, 304-315.

Martinsson, P., Myrseth, K. O. R., and Wollbrant, C. 2014. "Social dilemmas: When self-control benefits cooperation." *Journal of Economic Psychology*, 45, 213-236.

Martinsson, P., Pham-Khanh, N., and Villegas-Palacio, C. 2013. "Conditional Cooperation and Disclosure in Developing Countries." *Journal of Economic Psychology*, 34. 148-155.

Milgram, N., Sroloff, B., and Rosenbaum, M. 1988. "The procrastination of every day life." *Journal of Research in Personality*, 22, 197-212.

Mischel, W., and Metzner, R. 1962. "Preference for delayed reward as a function of age, intelligence, and length of delay interval." *Journal of Abnormal and Social Psychology*, 64, 425-431.

Mischel, W., Shoda, Y., and Peake, P. K. 1988. "The nature of adolescent competencies predicted by preschool delay of gratification." *Journal of Personality and Social Psychology,* 54*,* 687-696.

Mischel,W., Shoda, Y., and Rodriguez, M. L. 1989. "Delay of gratification in children." *Science*, 244, 933-938.

Myrseth, K. O. R., and Fishbach, A. 2009. "Self-control: A function of knowing when and how to exercise restraint." *Current Directions in Psychological Science,* 18, 247-252.

Myrseth, K. O. R., and Wollbrant, C. 2013. "A theory of self-control and naïveté: The blights of willpower and blessings of temptation." *Journal of Economic Psychology*, 34, 8-19.

Myrseth, K. O. R., and Wollbrant, C. 2015a. "Intuitive cooperation refuted: Commentary on Rand et al. (2012) and Rand et al. (2014)." University of Gothenburg, Working Papers in Economics No. 617.

Myrseth, K. O. R., and Wollbrant, C. 2015b. "Less cognitive conflict does not imply choice of the default option: Commentary on Kieslich and Hilbig (2014)." *Judgment and Decision Making*, 10, 277-279.

Myrseth, K. O. R., Fishbach, A., and Trope, Y. 2009. "Counteractive self-control: When making temptation available makes temptation less tempting." *Psychological Science*, 20, 159-163.

Myrseth, K. O. R., Riener, G., and Wollbrant, C. 2015. "Tangible temptation in the social dilemma: Cash, cooperation and self-control." *Journal of Neuroscience, Psychology, and Economics*, 8, 61-77.

Neugebauer T., Perote, J., Schmidt, U., and Loos, M. 2009. "Selfish-biased conditional cooperation: On the decline of contributions in repeated public goods experiments." *Journal of Economic Psychology*, 30, 52-60.

Nielsen, U. H., Tyran, J-R., and Wengström, E. 2014. "Second thoughts on free riding." *Economics Letters*, 122, 13-139.

Nyborg, K. 2000. Homo economicus and homo politicus: Interpretation and aggregation of environmental values *Journal of Economic Behavior and Organization*, 42, 305-322.

O'Donoghue, T., and Loewenstein, G. 2007. "The heat of the moment: Modeling interactions between affect and deliberation." http//:www.arts.cornell.edu/econ/edo1/heat.pdf

Palfrey, T., and Prisbrey, J. 1997. "Anomalous behavior in public goods experiments: how much and why?" *American Economic Review*, 87, 829-846.

Piovesan, M., and Wengström, E. 2009. "Fast or fair? A study of response times." *Economics Letters*, 105, 193-196.

Rand, D. G., Greene, J. D., and Nowak M.A. 2012. "Spontaneous giving and calculated greed." *Nature*, 489, 427-430.

Rand, D. G., Peysakhovich, A., Kraft-Todd, G. T., Newman, G. E., Wurzbacher, O., Nowak, M. A., and Greene, J. D. 2014. "Social heuristics shape intuitive cooperation." *Nature Communications*, 5, 3677.

Romal, J. B., and Kaplan, B. J. 1995. "Differences in self-control among spenders and savers." *Psychology: A Journal of Human Behavior,* 32, 8-17.

Rosenbaum, M. 1980a. "A schedule for assessing self-control behaviors: Preliminary findings." *Behavior Therapy,* 11*,* 109-121.

Rosenbaum, M. 1980b. "Individual differences in self-control behaviors and tolerance of painful stimulation." *Journal of Abnormal Psychology,* 89, 581-590.

Rosenbaum, M. 1989. "Self-control under stress: The role of learned resourcefulness." *Advances in Behavior Research and Therapy,* 11, 249-258.

Rosenbaum, M., and Palmon, N. 1984. Helplessness and resourcefulness in coping with epilepsy." *Journal of Consulting and Clinical Psychology,* 52, 244-253.

Rosenbaum, M., and Rolnick, A. 1983. "Self-control behaviors and coping with seasickness." *Cognitive Therapy and Research,* 7, 93–97.

Rosenbaum, M., and Smira, B.-A. K. 1986. "Cognitive and personality factors in the delay of gratification of hemodialysis patients." *Journal of Personality and Social Psychology,* 51, 357-364.

Schelling, T. 1984. "Self-command in practice, in policy, and in a theory of choice." *American Economic Review,* 74*,* 1-11.

Schulz, J. F., Fischbacher, U., Thöni, C., and Utikal, V. 2014. "Affect and fairness: Dictator games under cognitive load." *Journal of Economic Psychology,* 41, 77-87.

Selten, R. 1967. “Die Strategiemethode zur Erforschung des eingeschränkt rationalen Verhaltens im Rahmen eines Oligopolexperiments.” In: *Beiträge zur experimentellen Wirtschaftsforschung,* Sauermann, H. (ed.), Vol. I, 136-168. Tübingen: Mohr.

Shoda, Y., Mischel, W., and Peake, P.K. 1990. “Predicting adolescent cognitive and social competence from preschool delay of gratification: Identifying diagnostic conditions.” *Developmental Psychology,* 26, 978–986.

Sugden, R. 1984. “On the economics of philanthropy.” *Economic Journal*, 92, 341-350.

Trope, Y., and Fishbach, A. 2000. “Counteractive self-control in overcoming temptation.” *Journal of Personality and Social Psychology,* 79*, 493-506.*

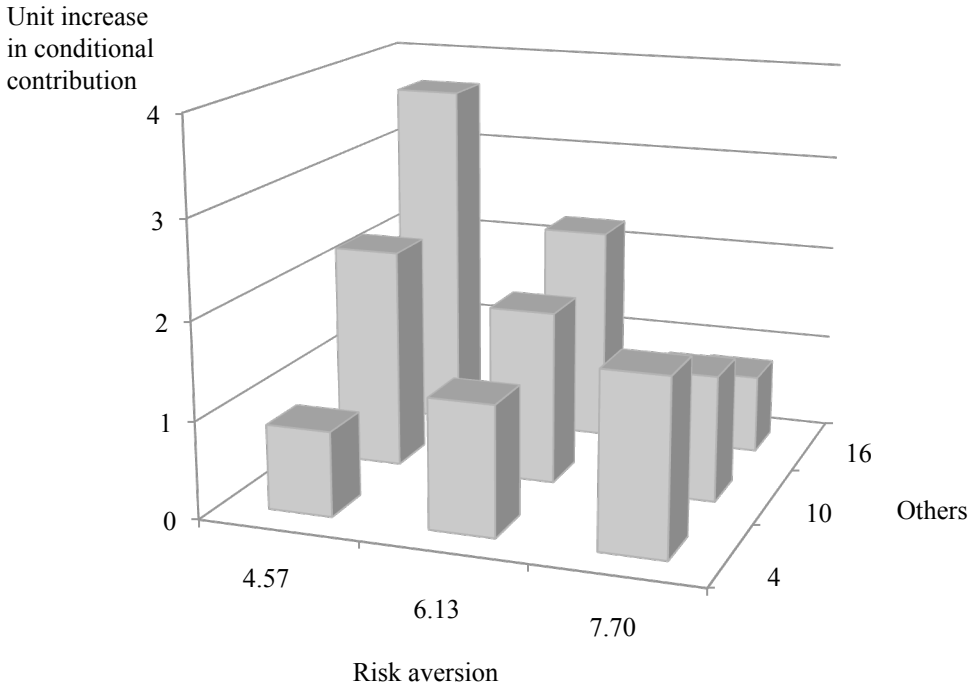
Verkoeijen, P. P. J. L., and Bouwmeester, S. (2014). Does intuition cause cooperation? *PLoS ONE*, 9, e96654.

Volk, S., Thöni, C., and Ruigrok, W. 2012. “Temporal stability and psychological foundations of cooperation preferences.” *Journal of Economic Behavior and Organization*, 81, 664-676.

Zelmer, J. 2003. “Linear public goods games: A meta-analysis.” *Experimental Economics*, 6, 299-310.

Figures and Tables

Figure 1. Unit increase in conditional contribution due to a one-standard-deviation increase in RSS, evaluated at different values of Risk and Others.



Note: The marginal effect of RSS is evaluated using specification (10) in Table 2. The change in conditional contributions due to a change in RSS can be approximately written as: ΔConditional contribution = (-0.144+ 0.031Risk + 0.031Others - 0.004RiskOthers)ΔRSS. The values chosen for each variable are the mean, one standard deviation above the mean and one standard deviation below the mean (N=144).

Table 1. Summary statistics

Variable	Description	Number of observations	Mean	Std. Dev.	Min	Max
Unconditional contribution	Unconditional contribution to the public good.	144	6.75	5.93	0	20
Conditional contribution*	Conditional contribution to the public good.	3024	6.02	6.29	0	20
Conflict intensity**	A continuous variable, ranging from 0 = "Not at all" to 100 = "Very much", in response to the question "*To what extent did you experience conflict when deciding how much to contribute?* "	144	33.14	32.06	0	100
Conflict	A dummy variable equal to zero if the participant responded 0 to the conflict intensity question and equal to one if the participant indicated a positive number.	144	0.75	0.43	0	1
Risk	Risk index derived from the experiment (switching point).	144	6.13	1.56	2	10
RSS	The Rosenbaum Self-Control Schedule score.	144	16.66	22.44	-46	76
High RSS	A dummy variable equal to one if the participant has a RSS score larger than the mean (17) and zero otherwise	144	0.51	0.50	0	1
Inconsistent	A dummy variable equal to one if the participant answered inconsistently in the risk experiment	144	0.10	0.31	0	1
Others*	A vector of integer numbers between 0 and 20 indicating all possible average contributions of the other three group members in the conditional contribution task	3024	10.00	6.06	0	20

Note: * denotes a variable constructed using the strategy method. ** denotes a response variable not used in the analysis, but transformed to a dummy; overall 36 out of 144 respondents reported zero, indicating "Not at all" as a response.