

2017

# The Dynamics of Prosocial Leadership: Power and Influence in Collective Action Groups

Ashley Harrell

*University of South Carolina*

Follow this and additional works at: <http://scholarcommons.sc.edu/etd>



Part of the [Sociology Commons](#)

---

## Recommended Citation

Harrell, A. (2017). *The Dynamics of Prosocial Leadership: Power and Influence in Collective Action Groups*. (Doctoral dissertation). Retrieved from <http://scholarcommons.sc.edu/etd/4179>

This Open Access Dissertation is brought to you for free and open access by Scholar Commons. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of Scholar Commons. For more information, please contact [SCHOLARC@mailbox.sc.edu](mailto:SCHOLARC@mailbox.sc.edu).

THE DYNAMICS OF PROSOCIAL LEADERSHIP: POWER AND INFLUENCE IN  
COLLECTIVE ACTION GROUPS

by

Ashley Harrell

Bachelor of Arts  
University of South Carolina, 2008

Master of Arts  
University of South Carolina, 2010

---

Submitted in Partial Fulfillment of the Requirements

For the Degree of Doctor of Philosophy in

Sociology

College of Arts and Sciences

University of South Carolina

2017

Accepted by:

Brent Simpson, Major Professor

Barry Markovsky, Committee Member

Shane Thye, Committee Member

Robb Willer, Committee Member

Cheryl L. Addy, Vice Provost and Dean of the Graduate School

© Copyright by Ashley Harrell, 2017  
All Rights Reserved.

## DEDICATION

To Dad, who since childhood has taught me all the important things in life: that bearings make things go 'round, that one should always (loudly) test out the sound system in the car before buying it, and, just as importantly: to figure out what I want, and then go get it. Thanks, Daddy.

## ACKNOWLEDGEMENTS

I am very grateful for the financial support I received in order to complete my dissertation. Specifically, the work was supported by a SPARC Graduate Research Grant from the Office of the Vice President for Research at the University of South Carolina, a Bilinski Educational Foundation Fellowship, and grant SES-1058236 from the National Science Foundation.

I have received many helpful comments and much support over the course of this project from my committee members: Barry Markovsky, Shane Thye, and Robb Willer. Many of my graduate student colleagues provided comments, insights, and friendship. I also appreciate my family and friends who asked questions—but not too many questions—about how my work was going, and provided advice, help, and sanity checks when needed. Extra special thanks and love to my husband Josh, who supported me in so many ways as I completed this project and my degree.

Last but certainly not least, I am particularly grateful to my committee chair, Brent Simpson. Over the years I have known him, Brent has given me countless hours of his time, advice, and support. I owe so many of my accomplishments to his guidance.

## ABSTRACT

This project bridges insights from theories of collective action, power, and influence to address the conditions under which group leaders solve collective action problems. Specifically, I show how group leaders' behaviors impact the success of collective action groups as a whole via both power and influence processes. The results of a laboratory experiment support the prediction that other-regarding (prosocial) leaders increase their contributions to the group after ascending to leadership, while self-regarding (proself) leaders reduce their contributions. Further, I show that rank and file group members are influenced by their leaders' contribution behaviors; as a result, prosocial-led groups as a whole are substantially more productive than proself-led groups. Indeed, as predicted, prosocial leaders were even more effective in maintaining large group contributions than the standard peer sanctioning system, where the ability to punish others is distributed equally among all group members. Importantly, these results suggest that prosocial leaders—but not proself leaders—are an effective solution to collective action problems. Therefore, I also address whether group members tend to select prosocials for leadership positions (Study 2a), and whether they are able to identify prosocials when all group members are able to compete for the leadership position by vying to be elected to the role (Study 2b). Results from these studies suggest that people prefer prosocials for the leader position, and that while group members do compete for leadership, both prosocial and proself individuals compete at similar rates, such that

prosocials remain higher contributors than their prosocial counterparts. As a result, prosocials are particularly likely to be selected for leadership positions when group contributions are known. Study 2b also demonstrates that groups that hold democratic elections for leadership may induce more cooperative behavior in their members not only once the leader is installed, but even before leadership hierarchies emerge as a result of competition to be elected. Taken as a whole, the findings suggest that putting power and influence in the right hands solves collective action problems and promotes collective welfare. Leadership—specifically, democratically elected, prosocial leadership—promises an effective solution to collective action problems.

## TABLE OF CONTENTS

DEDICATION .....	iii
ACKNOWLEDGEMENTS.....	iv
ABSTRACT .....	v
LIST OF TABLES .....	viii
LIST OF FIGURES .....	ix
CHAPTER 1: INTRODUCTION.....	1
CHAPTER 2: BACKGROUND.....	6
CHAPTER 3: THEORY .....	25
CHAPTER 4: STUDY 1 .....	33
CHAPTER 5: LEADER ELECTION VS. APPOINTMENT .....	59
CHAPTER 6: STUDIES 2A AND 2B.....	69
CHAPTER 7: CONCLUSION.....	91
REFERENCES .....	98
APPENDIX A – SVO MEASURE, STUDY 1 AND STUDY 2B .....	111
APPENDIX B – INSTRUCTIONS, STUDY 1 .....	114
APPENDIX C – INSTRUCTIONS, STUDY 2A .....	122
APPENDIX D – INSTRUCTIONS, STUDY 2B .....	131



## LIST OF TABLES

Table 4.1 Contributions to the Public Good, Non-Punishment Rounds .....	41
Table 4.2 Change in Total Contributions, Non-Punishment Rounds to Punishment Rounds .....	44
Table 4.3 Leaders' Previous Contribution Decision Predicts Non-Leaders' Current Contribution Decision .....	47
Table 4.4 Contributions to the Public Good, Punishment Rounds .....	52
Table 4.5 Earnings Deductions Received, Punishment Rounds .....	54
Table 4.6 Earnings, Punishment Rounds .....	56
Table 6.1 Contributions to the Public Good, Non-Punishment Rounds .....	81
Table 6.2 Contributions to the Public Good, Punishment Rounds .....	86

## LIST OF FIGURES

Figure 2.1 Causal Diagram: How Leaders Impact the Rank and File via Power and Influence Processes.....	29
Figure 4.1 Group-Level Contributions in Non-Punishment and Punishment Rounds, by Condition .....	41
Figure 4.2 Change in Total Contributions, Non-Punishment and Punishment Rounds, by Condition and Leader Status.....	45
Figure 4.3 Mediation Analyses: Leaders' Previous Behavior Mediates the Relationship between Condition and Group Contributions .....	48
Figure 4.4 Average Earnings Deductions Received, by Round and Condition.....	49
Figure 4.5 Punished Group Members' Contributions vs. Leader's Contributions .....	51
Figure 4.6 Average Earnings after Punishment, by Round and Condition.....	55
Figure 6.1 Group-Level Contributions in Non-Punishment and Punishment Rounds, by Condition .....	81

## CHAPTER 1

### INTRODUCTION

How people resolve their individual interests with those of the groups they belong to is a critical question in the social sciences (Cook, Hardin, and Levi 2005; Hardin 1982; Horne 2009; Olson 1965). In a *collective action problem*, what is best for the individual is in conflict with what is best for the group, yet if all pursue their own self-interest, the group as a whole is worse off (Heckathorn 1996; Marwell and Oliver 1993; Willer 2009). As Kollock (1998) notes, collective action problems are a ubiquitous part of social life, ranging from everyday problems such as productivity in work groups to global efforts to combat climate change. Since Marwell and Ames (1979, 1980), social scientists have studied contributions to non-excludable public goods as the quintessential collective action problem, given that self-interested individuals would prefer to free-ride on public goods produced by others rather than bearing the costs of contributing themselves (Olson 1965).

Perhaps the most widely considered solution to enhance contributions to public goods has been the introduction of *sanctions*. Allowing for the punishment or reward of group members reduces the tension between individual and group goals by making free-riding, or not contributing, less attractive (Fehr and Gintis 2007; Heckathorn 1989; Macy 1993; Yamagishi 1992). Much past work has shown that people are willing to expend valuable resources to punish free riders, and this effectively increases contributions to public goods (Fehr and Gächter 2002; Henrich et al. 2006; Ostrom, Walker, and Gardner

1992; Yamagishi 1986). Real-world groups often use sanctions to enhance contributions to public goods (Horne 2009; Strimling and Eriksson 2014).

The prevailing model of sanctions in the experimental literature on collective action is diffuse or “peer-to-peer” sanctions, where each group member can expend his or her own resources to deduct earnings from each other group member (Nosenzo and Sefton 2014; Shinada and Yamagishi 2007). While diffuse punishment systems increase contributions to collective efforts, researchers have pointed to several critical disadvantages of using them, as described in the next chapter. Just as importantly, modelling sanctions in the lab as peer-to-peer treats groups facing public good dilemmas as *non-hierarchical*—that is, it ignores the role that social differentiation processes play in administering punishment and solving collective action problems in real-world groups (Grossman and Baldassarri 2012).

Perhaps for these reasons, several researchers have begun to address how designating a single *group leader* to administer costly sanctions can impact contributions to collective action groups (Baldassarri and Grossman 2011; Grossman and Baldassarri 2012; Kosfeld and Rustagi 2015; Van Vugt and De Cremer 1999). However, in each of these studies, the leader was either completely external to the group or, once the leader had been selected from among the group members, they no longer made contributions and thus had no vested interest in the collective good. Therefore, no prior studies have addressed whether and how group leaders’ own contribution behaviors change after receiving the leader position. Nor have they shown whether and how leaders’ contributions influence the contributions of rank and file members and group-level outcomes more generally. Finally, no studies have addressed how groups in which the

capacity to sanction is given to a single group member relative to the standard solution in the literature, where sanctioning capacity is equally distributed among all group members.

Here I integrate insights from the literatures on collective action, power, and influence to address how and when leadership can help solve collective action problems. In the next chapters, after describing my conceptualization of leadership, I argue that *prosocial*, or other-regarding, group leaders will lead for the benefit of their groups, whereas *proself* (self-regarding)-led groups, on the whole, will be worse off. Perhaps more importantly, I address why prosocial leaders are so effective. I develop a theory in Chapter 3 that demonstrates how, via the exercise of *power* and *influence*, prosocial (but not proself) leaders establish norms of high cooperation among members of the rank and file. In addition, I argue that the benefits of prosocial leadership will exceed those provided by the standard solution to collective action problems, where the ability to punish is distributed equally among group members.

I test my hypotheses in a laboratory study. In Study 1, participants face a standard public good dilemma without sanctions. Then, in the punishment phase of the study, I manipulate whether one person (the *leader* conditions) or all group members (the *peer punishment* condition) may punish each other group member. I compare two different leader conditions: one where a prosocial has the sole ability to punish, and one where a proself is given punishment ability. Results suggest that prosocial leaders increase their contributions to the group—above and beyond their baseline tendencies to contribute—after ascending to leadership. Proself leaders, on the other hand, significantly decrease their contributions after receiving the leader position. For both types of leaders, via an

*influence* process, leaders' contribution decisions predict rank and file group members' future contribution decisions. The result is that prosocial-led groups on the whole produce larger public goods than either proself-led groups *or* groups where all members have the ability to punish. A different pattern emerges for earnings from the public good, however. Although proself leaders and their rank and file counterparts contribute less to public goods than peer-punishment groups, both prosocial- *and* proself-led groups earn more from the public good than their peer-punishment counterparts, due to the lower use of wealth-destroying punishments in leader groups. The experimental design, procedures, and results of Study 1 are detailed in Chapter 4.

The substantial improvement in welfare of groups with prosocial leaders in particular, relative to those with peer-to-peer sanctioning systems, may explain the tendency for real-world groups to limit sanctioning capacity to designated leaders or monitors (Baldassarri and Grossman 2011; Eriksson and Strimling 2012; Ostrom 2000). However, the results of Study 1 suggest that the benefits of leadership can only accrue to groups if *prosocial* individuals ascend to leadership positions, by whatever mechanism. Thus, one important question that is not addressed in Study 1 is whether prosocials actually do tend to end up as leaders of collective action groups. Do groups tend to select their more prosocial members to lead them? I also argue that the election of leaders, rather than leader appointment, will enhance public good production above and beyond the ability of elected-leader groups to choose more effective leaders. I discuss these issues and the motivation for Studies 2a and 2b in Chapter 5.

In a vignette study (Study 2a) I demonstrate that prosocials are preferred for positions of leadership, and if given the opportunity to select the leader, group members

facing collective action problems would prefer to elect a prosocial. In an experiment (Study 2b) I show that allowing group members to select their leaders via an election (compared to, in Study 1, appointment to leadership) enhances public good production in two ways. First, supporting some past work, groups (both leaders and rank and file group members) contribute more when the leader was elected rather than appointed. Second, election to leadership promotes public good production because group members *compete* for the leadership position, by contributing more *before* the election than they do before a leader will be appointed. Importantly, although allowing groups to elect their leaders induces competition for leadership, because both prosocials *and* proselves cooperate more when competition for leadership is possible, groups are still able to identify prosocials and select them for leadership positions. The experimental design, procedures, and results of Studies 2a and 2b are detailed in Chapter 6.

Taken as a whole, the studies presented here suggest that putting power in the right hands leads to substantially greater collective welfare than does the current prevailing solution to collective action problems. In the final chapter I summarize the findings, and discuss the implications of the results. Finally, I suggest new directions for future research on leadership and collective action.

## CHAPTER 2

### BACKGROUND

In order to provide their members with benefits, groups must overcome *collective action problems* (Olson 1965). In a collective action problem, individual rationality is in conflict with collective rationality, such that people may engage in self-interested actions that do not further the benefits of the group (Heckathorn 1996; Marwell and Oliver 1993; Willer 2009). How people resolve their individual interests with those of the groups they belong to is a critical question in the social sciences (Cook, Hardin, and Levi 2005; Hardin 1982; Horne 2009; Olson 1965). This is particularly true given the ubiquity of collective action problems: from everyday interpersonal exchanges like productivity in work groups to the mobilization of large-scale social movements or efforts to combat climate change, situations often occur where what is best for the individual is at odds with what is best for the group.

Since Marwell and Ames (1979, 1980), social scientists have studied contributions to non-excludable *public goods* as the quintessential collective action problem, given that self-interested individuals would prefer to free-ride on the public goods produced by others than to bear the costs of contributing themselves (Olson 1965). In a public good dilemma, individual group members must decide whether or not to contribute their private resources to produce a public good that benefits everyone, regardless of how much each individual contributed—that is, public goods are *non-excludable*. For example, everyone can enjoy public radio, blood supplies, or clean air



regardless of whether, or how much, they have contributed to the provision of the good. Given that public goods are non-excludable, it is best for any one individual to “free ride” off of the contributions of others, by enjoying the benefits of the public good without contributing to its contribution or maintenance (Kollock 1998). But if everyone in the group does so, then all will be worse off, as the public good will not be provided. Many important problems facing groups and societies involve the production of public goods: as briefly noted above, the maintenance of public radio and television, efforts to maintain clean air and water supplies, blood donation, and a variety of other critical issues involve the production of a public good.

Given that public goods problems—and collective action problems more generally—are pervasive in daily life, much past work has considered how they can be solved. Contrary to a pure rational self-interest perspective, people do make cooperative choices in collective action situations (Kollock 1998). How do groups motivate their members to forgo self-interest and contribute to group efforts? Past work has proposed a variety of explanations for how to increase cooperative behavior in collective action situations (for reviews, see Kollock 1998; Messick and Brewer 1983). For example, allowing group members to communicate with each other (Balliet 2010; Dawes, McTavish, and Shaklee 1977; Ostrom et al. 1992; van de Kragt, Orbell; and Dawes 1983) and make pledges to cooperate (Chen and Komorita 1994; Kerr et al. 1997; Orbell, van de Kragt, and Dawes 1988) have been shown to increase cooperation. In addition, enhanced feelings of social identification with the group can play an important role in determining whether people will engage in more group-serving behaviors (De Cremer and Van Vugt 1999; Kramer and Brewer 1984; Van Vugt and Hart 2004).

But perhaps the most widely considered solution to collective action problems has been the introduction of *sanctions*, such that group members have the ability to punish free-riders, reward cooperators, or both (Yamagishi 1986, 1988a; Kollock 1998; Komorita 1987). Olson (1965) considered “selective incentives”, or private rewards and punishments to encourage contributions to collective efforts, as a necessity for groups to solve collective action problems. Since Olson, a host of past research has considered whether rewards or punishments can successfully induce individuals to put aside their self-interest and cooperate to solve large-scale collective action problems. Allowing for the punishment or reward of group members reduces the tension between individual and group goals by making free-riding, or not contributing, less attractive (Fehr and Gintis 2007; Heckathorn 1989; Macy 1993; Yamagishi 1992).

Of course, sanctions will only impact cooperation if they will actually be administered—or at least, if there exists a credible threat that they will be administered (Fehr and Gächter 2000; 2002; Fehr et al. 2002). Importantly, much past work has shown that people are indeed willing to expend valuable resources to punish free riders, and this effectively increases contributions to public goods (Fehr and Gächter 2002; Fehr and Fischbacher 2004; Henrich et al. 2006; Ostrom 1990; Ostrom et al. 1992; Yamagishi 1986, 1988).

In particular, sanctions enhance cooperation via two separate mechanisms (Fehr and Fischbacher 2004a; 2004b; Shinada and Yamagishi 2007). First, because a majority of people are willing to punish free-riders, there is a credible threat that reduces the attractiveness of free riding. That is, the threat of punishment discourages would-be defectors. Second, because would-be defectors can be punished, would-be cooperators

are assured that others will cooperate as well. Thus punishment can also encourage cooperation by reducing potential contributors' fear that *others* will not cooperate (Fehr and Fischbacher 2004a; 2004b; Shinada and Yamagishi 2007). Fehr and Gächter (2000; 2002) have conducted experiments showing that without sanctions, cooperation typically declines over time. However, when group members are given the ability to pay to sanction each other, cooperation dramatically increases. Importantly, this increase in cooperative behavior is observed even on the very first round of interaction, which suggests that just the threat of punishment itself is credible enough to enhance group-oriented behavior (Fehr and Gächter 2000; 2002).

The prevailing model of sanctions in the experimental literature on collective action—including in the work reviewed above—is *peer sanctioning*, where each group member can punish every other group member (Nosenzo and Sefton 2014; Shinada and Yamagishi 2007). As noted above, peer punishment systems do successfully increase contributions to collective efforts. However, they also have a number of critical drawbacks. First, there are often significant costs associated with contributing to the provision or maintenance of a sanctioning system (Kollock 1998; Yamagishi 1986). Indeed, sanctions themselves are a public good (Heckathorn 1989; Oliver 1980) and, as a result, pose a “second-order free rider problem,” such that group members may be motivated to enjoy the fruits of sanctions (the resultant increase in group contributions) without bearing the costs of providing them (Heckathorn 1989; Oliver 1980). The difficulty of coordinating peer punishment can also result in over-punishment when group members “pile on” a free-rider, meting out a punishment that outweighs the offense (Nosenzo and Sefton 2014). Relatedly, the costs of sanctions sometimes exceed

the benefits of increased cooperation, thus reducing overall welfare (Dreber et al. 2008; Ostrom et al. 1992; Nikiforakis 2008). Finally, the potential for retaliatory acts of counter-punishment can increase negative emotions and decrease contributions to the public good (Nikiforakis 2008).

Compared to peer sanctioning systems, institutions in which a *single* person is responsible for punishment are less subject to coordination problems, over-punishment and the associated earnings losses (Eriksson, Strimling, and Ehn 2013; Grossman and Baldassarri 2012; O’Gorman, Henrich, and Van Vugt 2009). Just as importantly, sanctioning capacity in many real-world groups is centralized in a single individual or role-occupant, rather than distributed equally among all group members (Baldassarri and Grossman 2011; Eriksson and Strimling 2012; Guala 2012; Ostrom 2000). Thus, not only is there is a mismatch between the sanctioning institutions we tend to observe in the real world, which generally limit who can sanction whom, and the standard solution—peer sanctions—that is proffered by the literature; some evidence suggests that the solution more commonly observed in real-world groups may be the more effective one.<sup>1</sup>

Perhaps for these reasons, several studies have shown that designating a group *leader*—in these studies, a single individual able to monitor group members’ behavior

---

<sup>1</sup> Of course, other types of sanctioning systems are also possible. For example, giving sole punishment ability to a group leader is one form of centralized sanctioning, but *pool punishment*, where group members invest in a formal institution that punishes according to predetermined rules, is another (e.g., Yamagishi 1986). Moreover, groups are often characterized by multiple forms or sources of sanctions. For instance, in any given group, formal rewards and sanctions may be administered by a group leader, while informal rewards and sanctions (e.g., peer pressure or disapproval) flow between rank and file members (Kitts 2006; Shinada and Yamagishi 2007).

and administer punishment—can promote contributions to collective action (Baldassarri and Grossman 2011; Grossman and Baldassarri 2012; Kosfeld and Rustagi 2015; Van Vugt and De Cremer 1999). Yet as described in more detail below, in each of these studies, the leader was either completely external to the group or, once the leader had been selected from among the group members, they no longer made contributions and thus had no vested interest in the collective good. Therefore, prior studies have not addressed whether and how group leaders' *own* contribution behaviors change after receiving the leader position; nor have they shown whether and how leaders' contributions influence the contributions of rank and file members. Finally, no studies have addressed how groups in which the capacity to sanction is given to a single group member fare relative to the standard solution in the literature, where sanctioning capacity is equally distributed among all group members. This work aims to address these issues.<sup>2</sup>

In the sections to follow, I integrate insights from the literatures on collective action, power, and influence to address how and when leadership can help solve collective action problems. After describing my conceptualization of leadership, I argue that *prosocial*, or other-regarding, group leaders will lead for the benefit of their groups, whereas *proself* (self-regarding)-led groups, on the whole, will be worse off. Perhaps more importantly, I address why prosocial leaders are so effective, showing how, via the exercise of *power* and *influence*, prosocials—but not proselfs—establish norms of high cooperation among members of the rank and file. In addition, I argue that the benefits of

---

<sup>2</sup> That is, the goal of the work presented here was to conduct a first investigation of the dynamics and effectiveness of “pure” leadership systems versus “pure” peer sanctioning systems. I therefore do not empirically address how the two systems interact. This is a critical question for future work, as discussed in the final chapter.

prosocial leadership exceed those provided by the standard solution to collective action problems, where the ability to punish is distributed equally among group members.

### *Group Leadership*

There are many different forms—and definitions—of leadership (Bass and Bass 2009). But the literature generally defines *leaders* as those who use various tools—most commonly, power and influence—to induce rank and file members to produce desired outcomes (Ahlquist and Levi 2011, Bass and Bass 2009). Effective use of these tools can facilitate group interaction and group goals, producing benefits to both leaders and their followers (Anderson and Brown 2010; Ahlquist and Levi 2011; Van Vugt, Hogan, and Kaiser 2008). Groups facing intergroup and intragroup conflict support appointing a leader, particularly when the group is performing poorly (Benard 2012; see also Samuelson et al. 1984; Van Vugt and De Cremer 1999). Yet whether leadership will ultimately help or harm group outcomes depends strongly on leader characteristics, including leadership style (Van Vugt and De Cremer 1999; 2002), and, most relevant here, how the leader is affected by the *power* inherent in the leadership role (Anderson and Brown 2010).

Key to my conception of leadership is that leaders possess disproportionate power over rank and file members (Anderson and Brown 2010; Bass and Bass 2009; French and Raven 1959). Since Weber (1947), sociologists and social psychologists have conceived of power as the ability of an individual to impose their will on others via their control over valuable resources (Cook and Emerson 1978; Emerson 1962; French and Raven 1959; Markovsky, Willer, and Patton 1987; Sell et al. 2004). Following Weber,

sociological work has focused on power as the capacity to reward (Molm 1988, 1989, 1990) and punish (Baldassarri and Grossman 2011; Lawler and Bacharach 1987; Molm 1988, 1989, 1990; Ostrom et al. 1992; Sell and Wilson 1999; Yamagishi 1986).

Following the literature on which this work builds (Baldassarri and Grossman 2011; Grossman and Baldassarri 2012; Kosfeld and Rustagi 2015; Van Vugt and De Cremer 1999), here I focus specifically on leaders whose power is based on their sole ability to punish other group members. This focus on leaders who can punish—versus reward—facilitates comparisons with peer sanctioning systems which, with very few exceptions, typically focus on punishment (see Shinada and Yamagishi 2007 for a review).

Because leadership typically entails disproportionate power over others (Anderson and Brown 2010; Bass and Bass 2009), it is important to know whether and how those who move into leadership positions are affected by power. Indeed, how leaders wield their power has important consequences in a variety of micro and macro organizational settings (Anderson and Brown 2010; Flynn, Gruenfeld, Molm, and Polzer 2011). One line of research holds that power corrupts (Bass and Bass 2009; Kipnis 1972, 1976). For instance, compared to their less powerful counterparts, the powerful tend to be more attentive to rewards, and construe others as a means to attain rewards (Keltner, Gruenfeld, and Anderson 2003). Similarly, power reduces individuals' ability to accurately perceive others and the relations between them (Keltner and Robinson 1996, 1997; Simpson, Markovsky and Steketee 2011). The powerful are more likely to rely on stereotypes when judging interaction partners (Fiske 1993). The powerful spend more money on themselves compared to others, a pattern that is reversed for the powerless (Rucker, DuBois, and Galinsky 2011). In addition, individuals with high power are less

likely to adopt others' perspectives (Galinsky et al. 2006) and more likely to treat others hostilely (Keltner et al. 2001).

This past work, then, suggests that power enhances self-regarding tendencies—and reduces other-regarding tendencies. From this perspective, it might seem puzzling that leadership is so ubiquitous in real-world collective action groups—and that past work has considered whether solitary punishers in the lab will promote cooperation. Rather, the “power corrupts” approach seems to suggest that, given enhanced power over others, leaders will respond by engaging in self-serving behaviors.

Indeed, as noted above, no past research on whether a single group leader can sustain contributions to collective action problems has considered how receiving the leader position impacts the leader's *own* cooperative behavior. Rather, existing work on the group leader solution to collective action problems has typically designated a leader who is *external* to the group, and thus uninvolved in the provision of—and unable to benefit from—the public good (e.g. Baldassarri and Grossman 2011; Grossman and Baldassarri 2012; Kosfeld and Rustagi 2015).<sup>3</sup> The one exception (O’Gorman et al. 2009) gave a solitary punisher the ability to administer sanctions only after group members made their contributions to the public good. Thus, each group member did not know

---

<sup>3</sup> To be clear, Grossman and Baldassarri (2012; Baldassarri and Grossman 2011) refer to leaders in their study as internal to the group, meaning the leader was selected from among group members. But, once the leader was given punishment power, she no longer made contributions and had no vested interest in the collective good.



when deciding how much to contribute if she would be able to punish others, or if she would instead face punishment from another.<sup>4</sup>

The omission of work on internal, contributing group leaders with the sole power to punish others may be partly due to the prominence of the “power corrupts” perspective described above. That is, this work suggests—and researchers might have assumed—that leaders given sole ability to punish their group members would take advantage of their position, free-riding on the efforts of the group while punishing others who choose to do so. But given that, in real-world groups, leaders are typically internal (vs. external) group members (Grossman and Baldassarri 2012), it is important to know whether and how group members will alter their efforts to the group once they have received a position of leadership within that group.

Although the work reviewed above suggests that power increases self-oriented behaviors and tendencies, some additional research has taken a substantially more nuanced view of how power impacts those who possess it (Blader and Chen 2012; Chen et al. 2001; DeCelles et al. 2012). These *person x situation* approaches argue that power accentuates baseline dispositions, such that those with other-regarding preferences will become *more* other-regarding when they acquire power, while those with more self-regarding preferences will become more self-regarding than they were before they obtained power (Chen et al. 2001). That is, power affords individuals with the freedom to express their own traits and attitudes—whether for the good or for the bad (Anderson and

---

<sup>4</sup> In addition, the solitary punisher in the O’Gorman et al. (2009) study had the punishing role for one round only; in each subsequent round, a new leader was randomly chosen. Thus at no point did leaders know they would be leaders until after they had made their contribution decision—preventing them from making contribution choices strategically based on their role (punisher or non-punisher).

Berdahl 2003; Keltner et al. 2003). More generally, this work suggests that individual-level variables will moderate whether power possession leads to positive or negative outcomes (Sturm and Antonakis 2014).

In the sections to follow, I apply this *person x situation* approach to the problem of leadership in collective action groups. Specifically, I consider whether an individual's *social value orientation* will moderate how the acquisition of a leadership position—here, having the sole ability to sanction group members—affects her behavior. Perhaps more importantly, I later address how leaders' response to the leadership position will ultimately impact the behaviors of rank and file group members, and the success of collective action groups as a whole, via both *power* and *influence* processes (Thye 2000; Thye, Willer, and Markovsky 2006; Willer, Lovaglia, and Markovsky 1997).

### *Social Value Orientation*

A number of studies in sociology, social psychology, and behavioral economics have addressed the role of different social preferences in explaining variation in cooperative behavior, including contributions to collective action (e.g. Fehr and Gintis 2007; Ones and Putterman 2007; Simpson and Willer 2008; Van Lange et al. 2007; Yamagishi et al. 2013). One of the most powerful person-level predictors of behaviors in collective action situations is an individual's *social value orientation*, defined as a relatively stable preference for how valuable outcomes are distributed between oneself and others (Balliet et al. 2009; Kollock 1998).

Although many social value orientations are theoretically possible, researchers generally focus on three “types”: *individualists*, *competitors*, and *prosocials* (Van Lange

et al. 1997). *Individualists* seek to maximize their own outcomes with less regard for the outcomes of others. *Competitors* seek to maximize the difference between their own and others' outcomes. Compared to these two "proself" value orientations, *prosocials* tend to maximize joint outcomes and to minimize differences between own and others' outcomes (Van Lange et al. 1997).

Social value orientation is predictive of a wide range of other-regarding behaviors in both the lab and in real-world settings. Compared to proselfs (both individualists and competitors), prosocials donate more to charity (Van Lange et al. 2007), are more concerned with group goals in organizational settings (Nauta et al. 2002), take more cooperative approaches to negotiations with others (De Dreu and Boles 1998), and, most relevant here, contribute more in collective action situations (Balliet et al. 2009). Here I focus on how prosocials and proselfs will respond to leadership positions, and how their behaviors as leaders impact contribution norms and the behaviors of rank and file group members.

#### *A Person x Situation Approach to Leadership and Collective Action*

As described earlier, some past work suggests that the possession of power enhances self-regarding (and reduces other-regarding) behavior. Alternatively, the *person x situation* approach to power suggests that, rather, power enhances baseline tendencies in the power holder (Chen et al. 2001). Following this line of work, in this section I consider how prosocials and proselfs will respond to leadership positions, and how their behaviors as leaders will impact contribution norms—i.e., the behaviors of their non-leader group members.

Specifically, following past work that suggests power accentuates baseline behaviors—and noting that prior work demonstrates prosocials contribute at higher rates than proselves in collective action situations—I expect that proselves given a group leader role will *reduce* their contributions to the public good, such that they will contribute *less* when they ascend to a powerful leadership position. On the other hand, prosocial leaders will contribute *more* when they become leaders. Consider a real-world example. This argument would suggest that, in a work group, a relatively prosocial leader given the ability to distribute tasks among his or her group members might choose to take on a disproportionate share of the work herself. On the other hand, a relatively prosself leader, in the same situation, would be more apt to free-ride on the efforts of his or her subordinates, delegating a substantial portion of the work to others. I outline this theoretical argument in more detail in the next chapter.

Importantly, I expect that these effects will occur *above and beyond* baseline differences in contributions between prosocials and proselves. After all, prosocials are generally more cooperative than proselves under situations of equal power (see Balliet et al. 2009 for a review). But following Chen et al. (2001), I expect that the acquisition of power will *magnify* these baseline tendencies in the power holder. That is, prosself individuals will become even more self-regarding when they receive the powerful leadership position while prosocials will become even more group-oriented.

If a leader contributes less to the group as a consequence of his or her power, it necessarily follows that the group as a whole will be worse off, since less of the public good will be provided when any one group member does not contribute his or her resources. Similarly, if a leader contributes more to the group, then the group as a whole

will benefit given that groups do better the more that any one individual contributes. More important, however, is how leaders' own behaviors shape group-level outcomes by impacting *others'* contribution behaviors. In the next section, I consider two routes through which leaders might promote (or inhibit) the welfare of the groups they lead: *influence* and *power use*.

### *Leader Impact on Group-Level Outcomes: Influence*

Besides enhanced power, leaders normally enjoy greater *influence* in groups (Anderson and Brown 2010; Lucas 2003). Whereas power entails the ability to impose one's will on others via control over resources, influence involves modifying others' behavior without the use of sanctions, e.g., via persuasion, information, or advice (Willer et al. 1997; Willer, Troyer, and Lovaglia 2005). A long tradition of research in sociology connects standing in status hierarchies to disproportionate influence over other group members (Berger et al. 1977; Ridgeway 2001); some of this work has focused specifically on the influence enjoyed by group leaders (e.g. Lucas 2003).

Several studies have addressed the role of status-based influence in collective action groups (Clark, Clark and Polborn 2006; Sell 1997; Simpson, Willer, and Ridgeway 2012; Willer 2009). These studies find that higher status group members initiate contributions to their groups and influence others to give at higher levels than they would otherwise, thus establishing cooperative norms over time. But as noted above, no prior work has considered whether group leaders—following past work, those with the sole ability to punish their group members—*influence* others to give via their own increased (or decreased) contributions to collective efforts.

Leaders are afforded a special position in group decision-making hierarchies (Van Vugt 2006); group leaders are more salient than non-leaders and command greater attention (Clark, Clark, and Polborn 2006; Ahlquist and Levi 2011). Thus leaders' behaviors, in particular, serve as a focal point that other group members consider when making their own choices (Ahlquist and Levi 2011). Charities, for instance, rely on "leadership giving," or large contributions from an initial wealthy donor, to induce others to give (Andreoni 2006). In the lab, studies of sequential public good games—where an initial group leader contributes first—show that group outcomes are enhanced when a leader's contribution signals to others that they should also contribute (Potters et al. 2007; Moxnes and Van der Heijden 2003). In line with this past work, I suggest that non-leaders will focus especially on leaders' contributions when making their own contribution decisions. That is, I expect that leaders' contributions will play a larger role in the emergence of cooperative group *norms*—or *uncooperative* group norms—via an influence process.

Social norms play a central role in enhancing cooperation in groups facing collective action problems (Fehr and Fischbacher 2004a; 2004b). Norms are rules prescribing what people should or should not do in a given situation; when violated, they are enforced by sanctions (Cialdini and Trost 1998; Hechter and Opp 2001; Horne 2009). Like sanctions, norms have a dramatic impact on behavior (Hechter and Opp 2001; Horne 2009). Norms discouraging free-riding behavior, in particular, are proposed to contribute to the provision of public goods and allow groups to solve collective action problems (Fehr and Fischbacher 2004a; 2004b; Gintis et al. 2003).

One norm that has been widely considered as having an impact on cooperation to collective action is the norm of *conditional cooperation*—that is, that one should cooperate given that others cooperate (Fehr and Fischbacher 2004a; 2004b; Fischbacher and Gächter 2010; Gintis et al. 2003). Players using the Tit-for-Tat strategy in repeated social dilemmas, for instance, are behaving in a conditionally cooperative manner (Axelrod 1984; Kollock 1998). Fehr and Fischbacher (2004a) show that a large percentage of people behave like conditional cooperators: they cooperate if others cooperate, increasing their own contributions if the average contribution of others is increasing.

Thus, a norm of conditional cooperation—combined with people’s willingness to enforce the norm by punishing norm violators—can enhance cooperation in groups facing collective action problems. On the other hand, social norms will be detrimental to cooperative behavior if *self-interested* behavior, rather than group-oriented behavior, is the norm. If norms suggest self-interested behavior is appropriate, then other-oriented behaviors will decrease and self-regarding behavior will increase (e.g., Irwin and Simpson 2013). That is, normative content can moderate whether norms will help or hurt groups solving collective action problems.

Linking leaders’ ability to shape group norms via their enhanced influence, to the earlier discussion of how proselves and prosocials respond to the acquisition of leadership, implies that there will be divergent collective outcomes depending on leaders’ social value orientation. Specifically, if leaders have disproportionate influence on the behaviors of rank and file members, and prosocial leaders contribute *more* to the group upon ascending to leadership, it follows that rank and file members led by a prosocial will also

contribute more to group goals. That is, prosocial leaders can be expected to establish cooperative norms among their group members, influencing non-leaders to contribute at higher rates as well.

On the other hand, following this same logic, members of proself-led groups can be expected to make *less* cooperative choices. If proself leaders decrease their contributions upon receiving the leader position, they may establish norms of uncooperative behavior such that rank and file group members also contribute less. Consider the example given in the previous section, where prosocial and proself leaders are expected to behave differently when delegating tasks among members of a work group. When a prosocial leader of a work group takes on a disproportionate share of the work towards a common goal, her subordinates may be influenced to match the leader's high effort level in subsequent cooperative tasks. On other hand, a more proself leader may influence her subordinates to *reduce* their efforts to the group, given that the leader has also put forth less effort. In this way, prosocials' and proselfs' responses to leadership, *and* the subsequent influence that their contributions have on the behaviors of rank and file members, will determine the extent to which groups with leaders will ultimately succeed at collective action.

#### *A Second Possible Route from Leaders' Behaviors to Group Outcomes*

Thus far I have argued that leadership can generate divergent effects, since prosocials and proselfs will respond differently to the power that accrues to leaders (see also Chen et al. 2001). Further, I expect that the effect of receiving this powerful leader position on the leader will spark divergent collective outcomes via an influence effect.



But it is also possible that these different group-level outcomes will emerge as a result of the leaders' differential exercise of his or her power (i.e., the use of punishment) on rank and file group members.

Specifically, the behavioral economics literature on sanctioning defines *strong reciprocators* or *altruistic punishers* as those who, compared to their more self-interested counterparts, contribute at high levels *and* punish others who fail to do so (Fehr and Gächter 2002; Fehr and Gintis 2007). This line of research suggests that prosocials will punish low contributions at higher rates than will proselves. Therefore, prosocial leaders may not only be more likely to contribute at higher levels after obtaining a powerful leadership position (here and in the next chapter I refer to the effect of power on the powerholder as a *first-order power effect*); they may also be more likely to punish others who fail to cooperate at high rates, thus increasing their group members' cooperation levels via their use of power (a *second-order power effect*—that is, the effect of power exercise on the less powerful).

Note, however, that some research runs counter to this strong reciprocity prediction. For instance, Ones and Putterman (2007) find evidence for individual differences in both contribution behavior and punishment strategies, noting that propensities to contribute and to punish are not perfectly correlated. More recent work has found that tendencies to cooperate and tendencies to punish others who do not cooperate are completely orthogonal, contrary to strong reciprocity approaches (Eriksson et al. 2014). In any case, although I primarily expect leaders to impact rank and file members' contributions via an influence process, I will also assess these second-order power effects as a non-competing mechanism.

To review, as noted earlier, no prior work has examined whether giving a single group member the sole ability to sanction his or her group members will promote collective action when the leader is an internal, contributing group member. As a result, we do not know from existing work whether leaders promote (or inhibit) collective action in the groups they belong to via power or influence processes. Nor do we know whether groups with leaders will fare better or worse than groups with a standard peer sanctioning system in place, i.e., a sanctioning system where all group members have *equal* power to punish others. Some past work suggests power enhances self-regarding (and reduces other-regarding) behavior; the alternative *person x situation* approach suggests that, rather, power enhances baseline tendencies (Chen et al. 2001).

Following this line of work, in the next chapter I outline the theoretical argument that prosocials and proselfs will differentially respond to leadership positions. In addition, I address how, upon ascending to leadership, leaders' behaviors impact contribution norms and the behaviors of rank and file group members via power and influence processes. Finally, I consider how leadership will fare compared to the standard solution in the literature on collective action, peer sanctions.

## CHAPTER 3

### THEORY

The theory and hypotheses I detail in this chapter build on prior work on leadership, power, and collective action. As noted earlier, no prior work has examined whether giving a group leader—following past work, a single group member given the sole ability to sanction his or her group members—will promote collective action when that group leader is also a contributing member of the group. As a result, we do not know from existing work whether, and how, leaders will alter their contribution behavior after receiving the leadership position. Nor do we know whether leaders will promote (or inhibit) collective action in the groups they belong to via power or influence processes. Finally, we do now know whether groups with leaders will fare better or worse than groups with a standard peer sanctioning system in place. This work aims to address these issues.

First I consider how leadership will impact the leaders' *own* behavior before turning to how leadership affects followers' responses. Following my conception of leadership as detailed in the previous chapter, leaders are those who typically have enhanced power over their group, including the ability to administer rewards and punishments toward their non-leader group members (Anderson and Brown 2010; Bass

and Bass 2009).<sup>5</sup> Indeed, past work on the leadership solution to collective action has granted the leader the sole power to sanction his or her group members after observing their contributions to group efforts. And since Weber (1947), sociologists and social psychologists have conceived of power as the ability of an individual to impose their will on others via their control over valuable resources, including rewards and punishments (Cook and Emerson 1978; Emerson 1962; French and Raven 1959; Markovsky, Willer, and Patton 1987; Sell et al. 2004). Once people receive a powerful leadership position, then, how will they react?

Following the person x situation approach to power discussed in the previous chapter—power *accentuates* baseline dispositions, enhancing individuals’ propensities to express their own traits and attitudes (Blader and Chen 2012; Chen et al. 2001; DeCelles et al. 2012), I assume that

*Assumption 1: If an individual receives a position of leadership, they are more apt to act according to their own preferences.*

Further, as described in the previous chapter, much past work suggests that social value orientation is one of the most powerful person-level predictors of cooperative behavior, including contributions to collective action (e.g. Fehr and Gintis 2007; Ones

---

<sup>5</sup> Of course, while leadership is correlated with power, it is not a sole determinant of power-holding. Some leaders may have limited or no power, and non-leaders may have power over leaders, for instance, when choosing to elect them to leadership or not. But because leaders typically hold power and this power is considered to be critical to groups’ success or failure, here I focus on whether leaders who do obtain power over their followers can successfully solve collective action problems. That obtaining the leadership position gives the leader enhanced power over non-leaders, then, is a key scope condition for the theory outlined in this chapter.

and Putterman 2007; Simpson and Willer 2008; Van Lange et al. 2007; Yamagishi et al. 2013). Specifically, prosocials prefer to minimize differences in valuable outcomes between themselves and others, while proselves prefer to maximize their own outcomes.

*Assumption 2a: If a prosocial individual is more apt to act according to their own preferences, then they will contribute more to collective action.*

*Assumption 2b: If a proself individual is more apt to act according to their own preferences, then they will contribute less to collective action.*

From these assumptions I derive Hypothesis 1 below, which argues that the social value orientation of the group leader will moderate leader behavior. But perhaps more importantly, I expect granting leadership to a single group member will not only affect the leader him or herself (as in Assumption 1) – I also expect that granting leadership to a single group member affects non-leader group members, via an influence process. As reviewed in the previous chapter, leaders are afforded a special position in group decision-making hierarchies. They are more salient than their non-leader group members and command greater attention among followers making their own choices (Clark, Clark, and Polborn 2006; Ahlquist and Levi 2011). In line with this past work:

*Assumption 3: If an individual receives a position of leadership, then her contribution decisions to collective action are more salient among other non-leader group members.*

Importantly, as described above, if non-leaders focus especially on leaders' contributions when making their own contribution decisions, non-leaders' behaviors will diverge based on the leader's own behaviors:

*Assumption 4a: If a group member's contribution decisions are more salient among other group members, and the group member is prosocial, others will contribute more to collective action.*

*Assumption 4b: If a group member's contribution decisions are more salient among other group members, and the group member is proself, others will contribute less to collective action.*

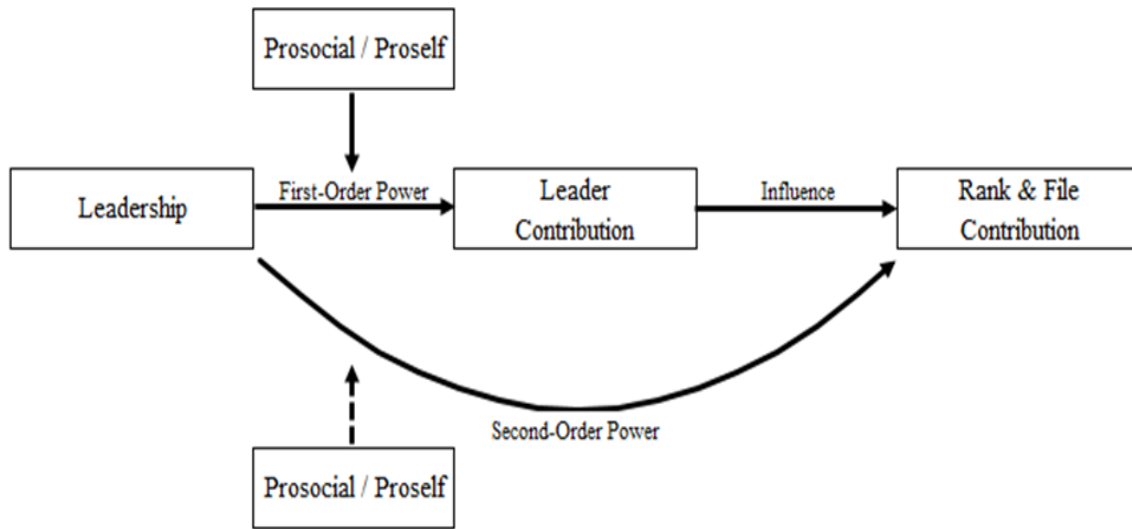
In the next section, from Assumptions 1 and 2 I derive Hypothesis 1 below, which argues that the social value orientation of the group leader will moderate leader behavior. Perhaps more importantly, I also consider how leaders' behaviors will impact non-leader group members' own behaviors—and, ultimately, group level outcomes—following Assumptions 3 and 4. I describe these hypotheses in turn.

### *Hypotheses*

Assumptions 1 and 2 above suggest that upon obtaining a leadership position, people will be freer to act according to their preferences; further, prosocials' and proselfs' preferences differ. Therefore I predict that the acquisition of a powerful leadership position will accentuate prosocials' and proselfs' pre-existing tendencies to contribute to collective action. That is, although I expect that prosocials will contribute more than proselfs in general (i.e., in their baseline contributions during rounds without sanctions), I also expect these differences will be significantly enhanced among prosocials and proselfs after they become *leaders*. I refer to changes in contributions as a result of obtaining power as first-order effects of power.

As shown in the causal diagram given in Figure 2.1, I predict that first-order power effects will be moderated by whether the power holder is prosocial or proself:

*Hypothesis 1: Upon receiving the leadership position, prosocial group leaders will contribute more to collective action. Proself group leaders will contribute less.*



**Figure 2.1: Causal Diagram: How Leaders Impact the Rank and File via Power and Influence Processes**

*Note: Causal diagram of two routes through which (prosocial) leadership can solve collective action problems: power and influence. I expect prosocial/proself disposition of the leader to moderate the relationship between leadership and contribution behavior—prosocials will contribute more while proselfs will contribute less. I further expect leader contribution to have a positive impact on contributions of the rank and file. The dotted arrow predicting that prosocial/proself disposition of the leader moderates second-order power use indicates that this hypothesis is exploratory (see main text).*

In the previous chapter I have also argued that the leader's social value orientation will impact collective outcomes. Specifically, leaders may either enhance or inhibit collective action via an *influence process*—that is, leaders' behavior will be more salient to non-leaders when making their own choices, such that they will model their behavior after the leader's. Thus, if prosocials contribute at higher levels after ascending to leadership positions, I predict that rank and file members of prosocial-led groups will be influenced by their leaders' behavior and will also sacrifice more for the group. Likewise, if proselfs decrease their contributions in response to attaining leadership, the rank and file will also contribute less (influence effect, Figure 2.1).

As noted above, research on strong reciprocity suggests we should also observe second-order power effects, namely that prosocial leaders will be more likely to use punishment as a means of increasing contributions from the rank and file (second-order power effect, Figure 2.1). While my main theoretical argument centers on the influence mechanism, the experiment also allows me to address this additional (non-competing) route from leaders' social value orientation to group contributions:

*Hypothesis 2: Compared to groups led by proselfs, groups led by prosocials will contribute more to the public good.*

Thus far, the hypotheses have focused on the relative advantages of prosocial versus proself leadership. Another aim is to compare the effectiveness of groups with leaders to the standard solution in the collective action literature, in which the power to sanction is distributed equally among all group members. I begin with contributions and then turn to earnings.



If proselfs reduce their contributions when they rise to leadership positions (and influence their group members to contribute less as well), we should expect groups led by proselfs to contribute less than members of groups with peer sanctions, where all group members have equal power and influence. Likewise, if prosocials increase their contributions when they become leaders (and influence their group members to contribute more), prosocial-led groups should contribute more than members of groups with peer sanctions. Thus, I expect contributions in the peer-punishment groups to fall between the high contributions of prosocial-led groups and the low contributions of proself-led groups:

*Hypothesis 3a: Compared to peer-punishment groups, prosocial-led groups will contribute more to the public good.*

*Hypothesis 3b: Compared to peer-punishment groups, proself-led groups will contribute less to the public good.*

Finally, consider earnings. Earnings from the public good *before* group members administer sanctions a direct linear function of contributions. Therefore, earnings before accounting for the costs of sending and receiving sanctions are expected to follow the same pattern as in Hypotheses 3a and 3b: compared groups where all are able to sanction all, prosocial-led groups will earn more from the public good as a direct result of their higher contributions, while proself-led groups will earn less from the public good. More interesting is how earnings between leader and leaderless groups will differ *after* accounting for the costs of sanctions.

As noted earlier, a fundamental drawback of peer sanctioning systems is the difficulty of coordinating punishments and, relatedly, the loss of earnings from over-

punishment and retaliatory counter-punishment. I expect that, after accounting for the welfare-destroying costs of sanctioning, peer-punishment groups will earn less than prosocial-led groups, where only one member can punish. I do not offer here a prediction for how proself-led groups will compare to peer-punishment groups, as it is unclear whether proself-led groups' lower contributions (see Hypotheses 1, 2, and 3b) but comparatively low sanctioning costs will be offset by the higher sanctioning costs but relatively high contributions in peer sanctioning groups:

*Hypothesis 4: After accounting for punishment costs, compared to peer-punishment groups, prosocial-led groups will earn more from the public good.*

The next chapter outlines the experiment designed to test these hypotheses. More generally, the experiment allows me to address whether group-level outcomes differ across leader types, whether this occurs via power, influence, or both, and whether the (prosocial) leader solution is as effective, or more effective, than the standard peer sanctioning solution.

## CHAPTER 4

### STUDY 1

Here I outline the experiment designed to test the theory outlined in the previous chapter. More generally, the study described below allows me to address whether group-level outcomes (i.e., contributions to and earnings from the public good) differ across the two leader types, whether these differences occur via power, influence, or both, and whether the (prosocial) leader solution is as effective, or more effective, than the standard peer sanctioning solution.

#### *Participants and Design*

Participants were recruited from large introductory classes at the University of South Carolina. Each session was conducted in a group of four and each group was randomly assigned to one of the three conditions: *prosocial leader*, *proself leader*, or *peer punishment*. 78 groups (312 participants; 182 female) completed the study in exchange for payment.

All groups first completed a standard repeated version of a public good dilemma with no punishment ability (e.g. Sell and Wilson 1991; Willer 2009). Then, each group participated in additional rounds of the game under new rules, which varied by condition. In two conditions, only one (ostensibly randomly selected) participant could punish his or her group members. In one of these, the group member with punishment ability had been categorized as prosocial. In the other, this group member was a proself. I refer to these

conditions as the *prosocial leader* and *proself leader* conditions, respectively. In the third condition, all group members were able to punish (the *peer punishment* condition). This condition, where all can punish all, is the standard one in the literature on sanctions in public good situations (e.g. Fehr and Gächter 2002; Ostrom et al. 1992; for a review, see Shinada and Yamagishi 2007).

### *Procedure*

Upon their arrival to the lab, participants were escorted to a private room where they were unable to see or communicate with other group members. Participants completed a consent form before beginning the computerized task, which was programmed using z-Tree version 3.3.11 (Fischbacher 2007).

The task began with a series of pre-study questions, including a standard measure of social value orientation (Van Lange et al. 1997, see Appendix A). The measure presented participants with a series of nine decomposed games, each consisting of three different distributions of points for self and another (unidentified) person. The other was described as “someone you do not know and that you will not knowingly meet in the future.” As noted earlier, two of the three distributions (individualist and competitive) correspond to “proself” responses and the third is a “prosocial” response. Aside from measuring participants’ preferences for prosocial or proself outcomes, I also used scores on the social value orientation (SVO) scale to assign the leader in the prosocial and proself leader conditions, as described below. At no point during the study were participants told that their responses on the SVO scale would be used to determine who would occupy the leader role.

After everyone completed the pre-study questionnaire, the computer advanced to instructions for a standard public good game without punishment. The description of the task began by stating that while participants would be working in a group with three others, at no point would they see or meet the other group members; nor would their group members see or meet them. After assuring participants of their anonymity, they were told that the study would consist of a series of choices in an “investment decision task” and that their and their group members’ decisions in the task would affect their earnings in the study. Specifically,

*At the start of each round, you (and each of your other group members) get 20 points. You can contribute anywhere from 0 to 20 of these points to a "group fund." Any points you do not contribute to the group fund remain in your personal fund, for you to keep.*

*Anything that is contributed to the group fund will be doubled. Then, the doubled amount will be divided between all four of the members of your group, whether or not they contributed to the group fund. (Similarly, other group members' contributions to the group fund will be doubled and redistributed equally among all members of the group.) Your total earnings per round are your share of the earnings from the group fund, plus whatever you did not invest.*

After reading through the instructions and several examples, participants answered quiz questions to ensure their comprehension (see Appendix B for the full text of all Study 1

instructions). The vast majority of participants were able to answer all quiz questions correctly. Incorrect answers to the quiz questions were followed by a detailed explanation of the correct answer.

The public good task began after everyone completed the quiz. Each participant was prompted to enter a whole number from zero to twenty to indicate how much they wished to contribute to the group fund. After all participants had answered, everyone was able to see their own contribution and earnings, as well as the contributions and earnings of each other group member. Once everyone finished reviewing the results, participants proceeded to the next round, where they again decided how much of a new twenty-point endowment to contribute to the group. This process was repeated for a total of six rounds. Participants were not told in advance how many rounds they would complete.

Once the sixth round was completed, a new set of instructions explained that for the subsequent portion of the study, some of the rules of the investment decision task would change. The description of these new rules varied by condition: either *peer-punishment* or *leader*. Instructions for the two leader conditions—prosocial leader and proself leader—were identical. (To avoid demand effects, at no point in the instructions was the participant given the leader role referred to as the “leader.”) The description of the rule change for each condition was as follows [instructions for leader conditions in brackets]:

*Under the new rules, every group member [one group member (chosen at random)] will have the opportunity to deduct points from their fellow group members following each round.*

*Specifically, after each round (that is, after all members' contributions for that round are known), everyone [one person] will be able to deduct points from their group members' earnings, if they choose to do so. It costs 1 point to deduct 3 points from another group member. For example, if a [the] group member chooses to spend two of his/her points to deduct points from Participant X, Participant X's earnings will be reduced by six points.*

*Group members can spend anywhere from 0 to 10 of their own points to deduct points from each of their group members.*

*[On a later screen, we will let you and your group members know who has been randomly assigned to the role described above. Whoever is randomly selected will continue in this role for the remainder of the study.]*

As in the no-punishment phase of the study, participants read through several examples and completed quiz questions (see Appendix B). Once everyone had completed the quiz, participants in the leader conditions were shown one final screen where they were given the ID number of the person (ostensibly randomly) assigned to be the leader. Depending on condition, I used scores on the social value orientation measure, administered in the beginning of the study, to select the leader. In the *proself leader* condition, the group member with the *fewest* number of prosocial answers (i.e. a high number of proself answers) was assigned to the leader role. In the prosocial leader

condition, the group member with the *most* prosocial answers was assigned to the leader position. In case of a tie, the leader was chosen randomly from among the most prosocial (prosocial leader condition) or most proself (proself leader condition) individuals. All participants in the leader conditions were told whether or not they had been assigned to the leader position—or if not, the ID number of the participant who had (ostensibly randomly) received the leader role—before advancing to the punishment phase of the study.

The rounds with punishment began similarly to the rounds without punishment. All group members made simultaneous contribution decisions and then viewed the results, including each group member's contribution and earnings. Then, non-leaders in the leader conditions were told to wait while the leader made his or her deduction decisions. Those who were leaders were again shown each other group member's contribution and earnings. Below that information, they indicated whether (and if so, how much) they wished to deduct from each person's earnings. For the peer-punishment condition, each participant viewed an identical screen and made deduction decisions. Once those who could deduct points had done so, all group members could view their and others' outcomes after deductions—including how many deductions they had received and their final earnings for the round. In the peer-punishment condition, participants could see the ID numbers of who punished whom. This process was repeated for an additional six rounds. As in the no-punishment phase of the study, participants were not told the number of rounds in advance.

Once the punishment phase of the public good task was completed, participants were paid based on their earnings across the twelve rounds of the study (ranging from



\$10 to \$15), probed for suspicion, and debriefed. Each session lasted approximately forty-five minutes.

### *Results*

I used the nine-item social value orientation scale, completed at the beginning of the study, to categorize participants' social value orientation. Following past work, I combined individualists and competitors into one "proself" category (Simpson and Willer 2008; Van Lange et al. 1997). As is standard in the social value orientation literature (Simpson and Willer 2008; Van Lange et al. 1997), a participant was classified as a given social value type when he or she answered at least six of the nine items consistently.

Six groups were omitted from analyses (four in the prosocial leader and two in the proself leader conditions) because no group member qualified as prosocial or proself, respectively. In the leadership conditions, the experimental software selected the individual with the highest prosocial or proself score to be leader, depending on condition. In three groups, this resulted in a participant who could not be classified as either prosocial or proself becoming the leader. In the remaining three groups, a proself was assigned to lead a group originally designated a prosocial leader condition. After omitting these six groups, analyses were performed on the remaining 72 groups of four, or 288 participants (166 female). These groups were fairly evenly distributed across the prosocial leader (25 groups), proself leader (22), and peer-punishment (25) conditions. Of the 288 participants included in the study, 119 (42%) were classified as prosocial, 143 (49%) as proself, and 26 (9.2%) could not be classified.

Because the data were nested (rounds within participants within groups), most of the analyses reported below employ three-level multilevel models with random intercepts for participants and groups. Multilevel modeling allows for the simultaneous analysis of different levels of data without violating the assumptions of independence required for traditional regression (Snijders and Bosker 2012).

### *Preliminary Analyses*

Prior to evaluating the hypotheses, I performed some preliminary analyses. First, all groups completed six non-punishment rounds at the beginning of the study. The instructions and procedures for this phase of the study were identical across conditions. Therefore, there should not be significant differences in contributions between conditions in the non-punishment phase. Any differences between conditions before the manipulation had occurred would suggest a failure of random assignment. It would also suggest that any differences observed in the punishment rounds may be due not to effects of the leadership manipulation, but to pre-existing differences between groups. Therefore, I first examined contributions in the non-punishment phase of the study by condition.

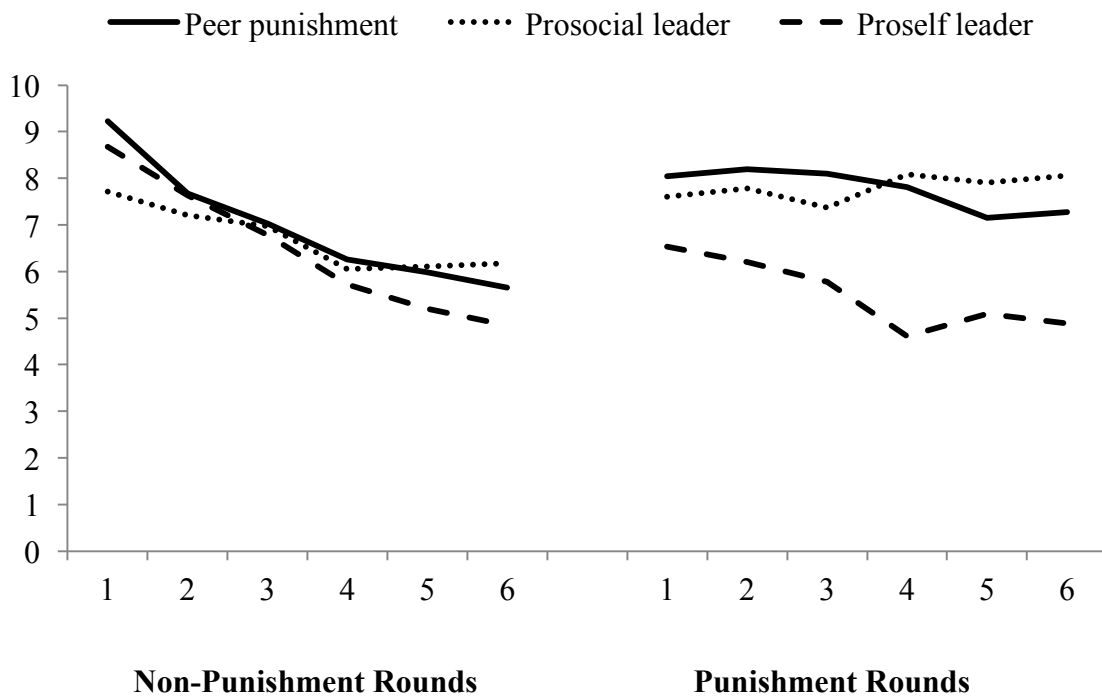
Results for the non-punishment rounds are displayed in Table 4.1. As expected, there were no significant differences in contributions between the prosocial leader, proself leader, or peer punishment conditions in the non-punishment phase ( $ps > .8$ ). There was an effect of round: within the six rounds where punishment was not possible, contributions declined significantly over time ( $p < .001$ ); this pattern was the same in all conditions (see non-punishment rounds, Figure 4.1). The degeneration of contributions over time is a standard finding in the literature (e.g. Ostrom 2000; Sell and Wilson 1991)

and is one of the key problems for groups facing collective action situations. Peer sanctioning is the standard solution proffered by the literature for avoiding this spiral of non-cooperation and maintaining high levels of contributions over time (Fehr and Gächter 2002).

**Table 4.1: Contributions to the Public Good, Non-Punishment Rounds**

	<i>B (SE)</i>
Intercept	8.78 (.83) <sup>***</sup>
Proself leader condition	-.23 (1.18)
Peer punishment condition	.27 (1.14)
Round	-.59 (.06) <sup>***</sup>
N <sub>rounds (groups)</sub> = 1728 (72)	

*Note:* \*  $p < .05$       \*\*  $p < .01$       \*\*\*  $p < .001$



**Figure 4.1: Average Group Contributions in Non-Punishment and Punishment Rounds, by Condition**

Additional preliminary analyses were conducted to determine if the leader selection process resulted in groups that differed in the distribution of gender or non-leader social value orientations. Specifically, in the leader conditions, the most prosocial or proself group member was selected to be leader. Although evidence for gender differences in social value orientation is mixed at best (Van Lange et al. 1997; Van Lange 1999), I considered whether the gender of the leader was approximately equally distributed across the prosocial and proself leader conditions. That is, if leaders were more likely to be men vs. women in one condition compared to the other, this could result in differences by condition based on the group leaders' gender, rather than the group leaders' social value orientation alone. A chi-square test, however, revealed that there was no significant difference in the gender distribution of the leaders by condition,  $\chi^2(1) = 1.95, p = .16$ .

Additionally, I considered whether, as a result of selecting the most prosocial (or proself) group member for the leadership position, the social value orientations of the *remaining* group members (i.e., the rank and file group members) were unevenly distributed across conditions. For instance, selecting the most prosocial group member to be leader might have resulted in a more heavily proself rank and file in the prosocial leader condition. Likewise, selecting the most proself group member to be leader may have resulted in a more heavily prosocial rank and file in the proself leader groups. However, a chi-square test revealed that non-leader group members' SVOs were distributed approximately equally across the prosocial and proself leader conditions,  $\chi^2(1) = 1.323, p = .25$ .

### *Change in Contributions, Non-punishment to Punishment Rounds*

As described in the previous section, and as anticipated, contributions did not differ across conditions in the non-punishment rounds. Following the non-punishment rounds, participants were told either that all participants or one randomly selected participant would be able to make deductions from the group members' earnings. How did the introduction of sanctions impact contributions to the public good?

A Wilcoxon signed-rank test revealed that group-level contributions increased significantly from the final round of the non-punishment phase to the initial round of the punishment phase in all three conditions ( $p < .03$ ). Further, a two-level (participants within groups) multilevel model with contributions in just the first round of the punishment phase as the outcome variable suggested that first-round contributions did not differ by condition ( $p > .3$ ). Taken together, these results suggest that the introduction of each type of sanctioning system (leader or peer-punishment) immediately increased contributions to the public good, and at similar levels, before the leader (or all group members, in the peer-punishment condition) had a chance to sanction.

I now turn to the first hypothesis by considering changes in individual-level contribution behavior after the introduction of sanctions—and, for the leaders, after receiving the leadership position. To do so, I examined how both condition and leader status impacted the *difference* in contributions between the six rounds where punishment was possible and the six rounds where it was not. This model allows for the assessment of whether leaders' contributions changed in the rounds in which they had sole ability to punish their group members, and also how non-leaders' contributions changed as well.

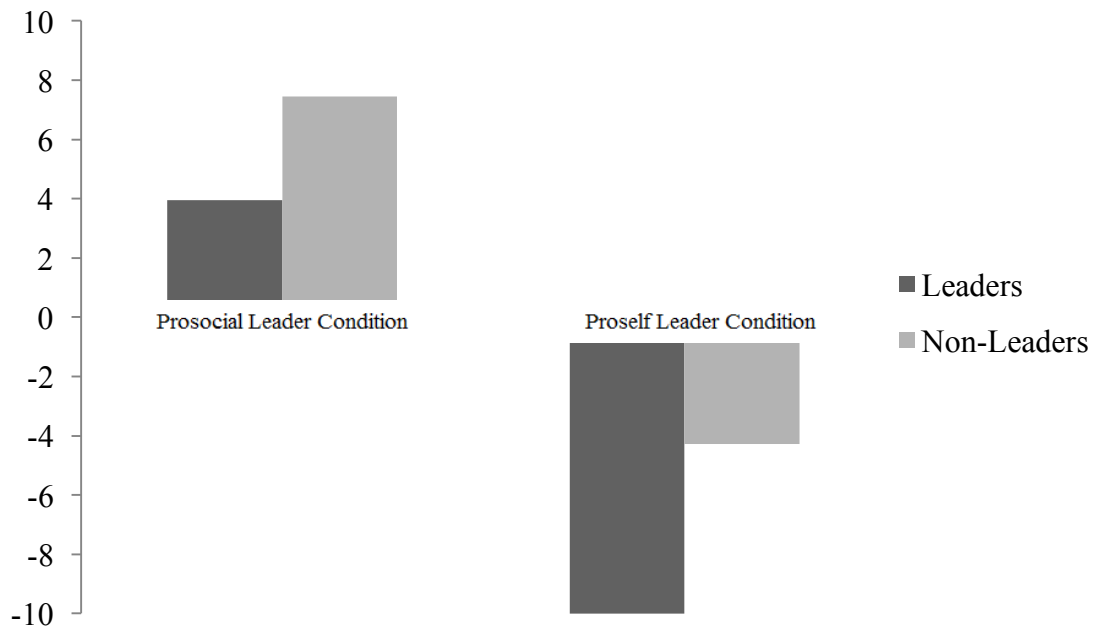
Because these models contain variables for leadership status, they are based only on the 47 groups in the two leader conditions; peer-punishment groups were omitted from these analyses and are discussed later in the chapter.

Results of a model with change in contributions as the outcome variable and random intercepts at the group level revealed a main effect of condition: those in the proself leader condition contributed less in the punishment rounds, compared to those in the prosocial leader condition ( $p = .04$ ). The main effect was qualified by a significant *condition x leader status* interaction: proself leaders, in particular, contributed less in the six punishment rounds than they did in the six non-punishment rounds ( $p = .03$ , see Table 4.2 and Figure 4.2). While proself leaders contributed an average of 10.0 points *less* than they contributed before they became leaders, prosocials contributed, on average, 4.0 points *more* when they became leaders. The findings lend support to the person x situation approach, and Hypothesis 1: although proselfs who received the leadership role decreased their contributions to the group, I observed the opposite pattern for prosocials.

**Table 4.2: Change in Total Contributions, Non-Punishment Rounds to Punishment Rounds**

	<i>B (SE)</i>
Intercept	15.51 (3.57) <sup>***</sup>
Proself leader condition	-9.81 (4.71) <sup>*</sup>
Contribution, non-punishment rounds	.23 (3.46)
Is leader	1.96 (3.46)
Proself leader condition*Is leader	-11.25 (5.13) <sup>***</sup>
N <sub>participants (groups)</sub> = 188 (47)	

**Note:** +  $p < .10$       \*  $p < .05$       \*\*  $p < .01$       \*\*\*  $p < .001$



**Figure 4.2: Change in Total Contributions, Non-Punishment Rounds to Punishment Rounds, by Condition and Leader Status**

*Note:* figure displays the difference between leaders and non-leaders' total contributions in the six non-punishment rounds and their total contributions in the six punishment rounds. Positive numbers indicate that total contribution was higher in the punishment rounds than in the non-punishment rounds. Prosocial groups (i.e., leaders and non-leaders together) contributed more in the punishment rounds than proself groups; in the proself leader condition, proself leaders decreased their contributions to a greater extent than did their followers. The difference between prosocial leaders and their followers was not significant. See main text for analyses.

The significant *main* effect of condition in this model supports Hypothesis 2: it suggests that *non-leaders* in the proself condition *also* contributed less during the six punishment rounds, compared to non-leaders in the prosocial leader condition (see Figure

3). That is, proself leaders, and their followers, actually gave less than they had when there was no sanctioning system at all, while prosocial leaders and their followers gave more. The next section more closely examines the mechanisms through which contributions to groups in these two conditions diverged.

*Do Leaders Impact Contributions from the Rank and File via Influence, Power, or Both?*

Thus far I have shown that even after controlling for their baseline tendencies during non-punishment rounds, proselfs decrease their contributions further when they become leaders. Further, Figure 4.2 and the change in contributions model in Table 4.2 suggest that non-leaders' contributions fell in line with their respective leader's contributions, suggesting the predicted influence process. To formally test this prediction, I next assessed whether leaders' contribution decisions in a *previous* round predict non-leaders' contributions in the *current* round (Table 4.3, Model 1).

I also address the possibility that group members did not favor leaders in particular, but instead used the behavior of *all* group members in the previous round (i.e., both leaders and non-leaders) when making their contribution decision. Further, it is possible that group members were not responding to leaders' contribution behaviors in the previous round, but rather to leaders' use of punishment in the previous round, when they made their own contribution choices. To ensure that leaders' contributions *in particular* influenced non-leaders' contribution decisions, in a second model I included two additional predictors: average contribution of the other two rank and file members in the previous round and leaders' expenditure on punishment of the three rank and file members in the previous round.



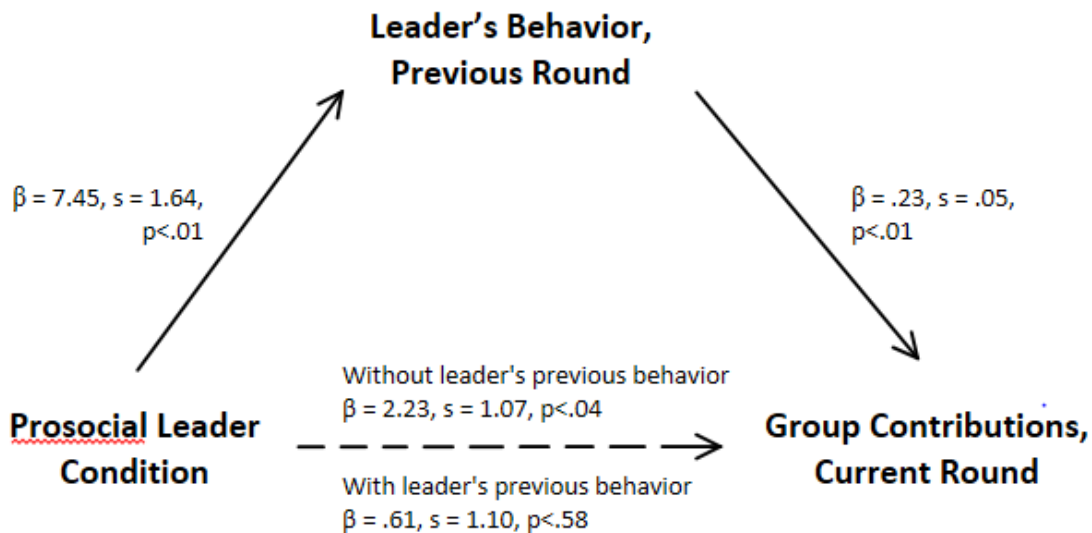
**Table 4.3: Leaders' Previous Contribution Decision Predicts Non-Leaders' Current Contribution Decision**

	Model 1 <i>B (SE)</i>	Model 2 <i>B (SE)</i>
Intercept	.85 (.76)	.02 (.70)
Proself leader condition	.52 (.66)	.31 (.55)
Round	.08 (.11)	.11 (.11)
Contribution, non-punishment rounds	.20 (.03) <sup>***</sup>	.20 (.03) <sup>***</sup>
Contribution, previous round	.41 (.03) <sup>***</sup>	.43 (.03) <sup>***</sup>
Other non-leaders' contribution, previous round		.04 (.02) <sup>*</sup>
Leader's expenditure on punishment, previous round		.10 (.05) <sup>*</sup>
Leader's contribution, previous round	.20 (.04) <sup>***</sup>	.17 (.04) <sup>***</sup>
N <sub>rounds (groups)</sub> = 705 (47)		

**Note:** +  $p < .10$       \*  $p < .05$       \*\*  $p < .01$       \*\*\*  $p < .001$

Not surprisingly, this final model (Table 4.3, Model 2) revealed a significant effect of other non-leader group members' contributions in the previous round on own contribution decision in the current round ( $p = .04$ ) and of leaders' use of punishment in the previous round on own contribution decision in the current round ( $p = .04$ ). However, as expected, the effect of the leaders' previous contribution on non-leaders' contribution in the current round was highly significant ( $p < .001$ ), suggesting that participants used the leader's own contribution behavior in particular as a cue for how much to contribute. There was no interaction with condition ( $p = .8$ ); thus, although the prosocial- and proself leaders tended to lead their groups in opposite directions, the magnitude of their influence was similar.

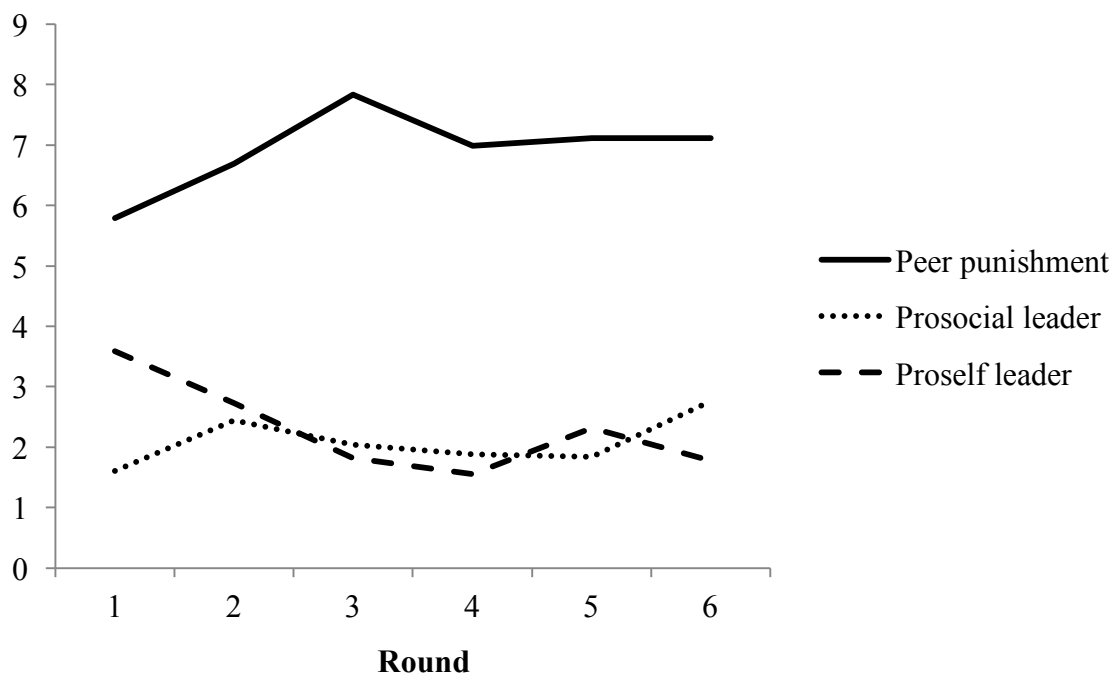
A mediation analysis revealed that the group-level contribution difference between conditions was fully mediated by the leader's influence (see Figure 4.3). Although condition significantly predicted contributions in a given round ( $p < .05$ ), and condition significantly predicted leader's contribution in the previous round ( $p < .001$ ), the effect of condition on contributions became non-significant after controlling for leader's behavior in the previous round. This suggests that the group-level differences in contributions were driven largely by the effect of the leader, specifically, the leader's *influence* on the rank and file's own contribution patterns.



**Figure 4.3 Mediation Analyses: Leaders' Previous Behavior Mediates the Relationship between Condition and Group Contributions**

*Note:* Figure gives results for mediation analyses showing that in a given round, the effect of the prosocial (vs. proself) leader condition on contributions was mediated by the leader's behavior in the previous round. The effect of condition on contributions becomes nonsignificant after controlling for leader's behavior in the previous round. All models controlled for contributions in the non-punishment rounds, round, and leader status.

The forgoing analyses show that leaders' contributions had a powerful and significant influence on non-leaders' subsequent contributions. And because prosocial leaders, but not proself leaders, contributed at high levels after ascending to the leadership position, the rank and file contributed at higher levels and prosocial-led groups as a whole were better off. Did leaders differ in their exercise of power—that is, their use of punishment? Figure 4.4 shows that prosocial and proself leaders did not differ in how much they punished ( $p > .7$ ). Moreover, the average contribution of those who were subsequently punished (and the average contribution of those who were *not* subsequently punished) did not differ by leader type ( $p = .35$ )—the average punished person contributed, on average, 4.2 points to the group in the proself leader condition and 4.9 points to the group in the prosocial leader condition.

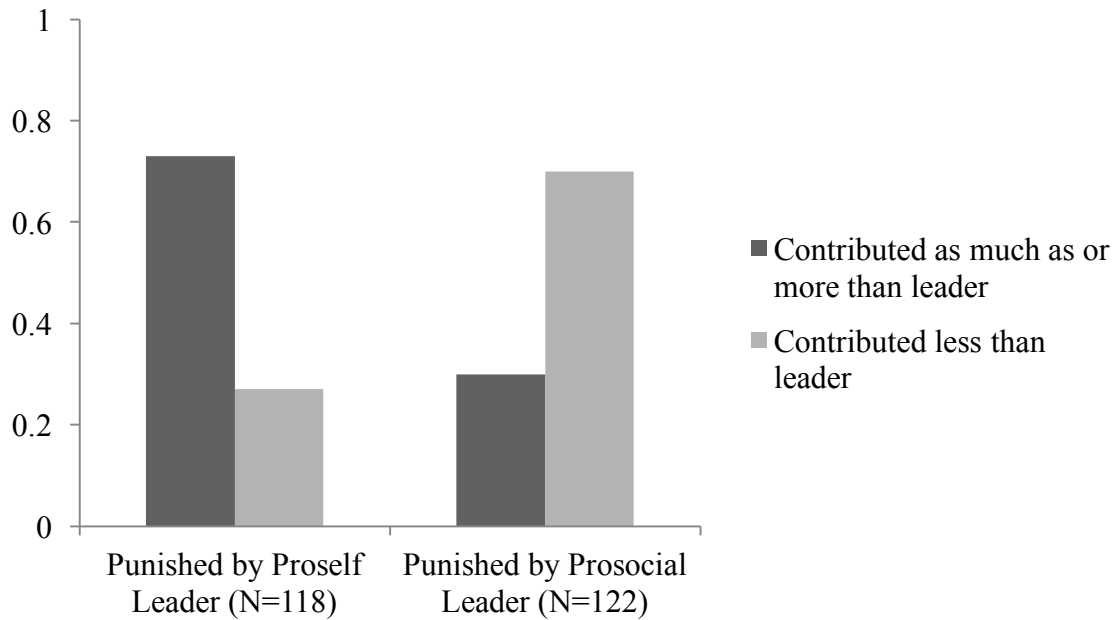


**Figure 4.4: Average Earnings Deductions Received, by Round and Condition**

*Note: Higher values indicate greater earnings loss. For the two leader conditions, the maximum a given group member could lose was thirty points (the leader could spend up to ten points to punish a given group member up to thirty points). For the peer punishment condition, the maximum was ninety points (thirty points from each of three group members).*

However, as described above, prosocial and proself leaders did not contribute to the group equally, even though they tended to punish equally. This suggests that proself leaders frequently engaged in *hypocritical punishment* of their fellow group members (Heckathorn 1990; Helbing et al. 2014)—that is, they were significantly more likely than their prosocial counterparts to punish rank and file members who contributed as much, or more than, they themselves had. Indeed, of the instances where a rank and file group member was punished by a proself leader, the target of punishment had contributed as much or more than the leader 73% of the time. The opposite pattern was observed for prosocials. For prosocial leaders, in 70% of the instances where they punished, the punishment was directed at someone who had contributed *less than* the leader (Figure 4.5).

Similarly, although the average individual punished by both a prosocial and a proself contributed about 4-5 points, as noted above, *prosocials* who punished at least one other group member had contributed, on average, 10.1 points; *proselfs* who punished, on the other hand, had contributed 3.0 points—less than the average amount (4.2 points) for which they were punishing others.



**Figure 4.5: Punished Group Members' Contributions vs. Leader's (Punisher's) Contributions**

*Note: Figure displays the 118 (122) instances where a rank and file group member was punished by a proself (prosocial) leader. Displayed here is the percentage of those instances where the punished rank and file group member had contributed as much or more than the leader (punisher) had, vs. contributed less than the leader (punisher) had.*

#### *How Does Leadership Compare to Peer Sanctioning?*

Next I turn to comparing groups in the standard peer sanctioning condition with leader groups to test Hypothesis 3: did public good provision differ by type of sanctioning system? As noted above, contributions in the first round (i.e., before sanctions took place) did not differ by condition. However, contribution patterns by condition diverged substantially across the six rounds (see punishment rounds, Figure 2). Specifically, the model in Table 4.4 shows a *condition x round* interaction: consistent with the individual-level analyses described above, groups in the proself leader condition

contributed less over the six punishment rounds, compared to prosocial-led groups ( $p < .01$ ). Thus, as already suggested and in line with Hypothesis 2, prosocial-led groups produced larger public goods than proself-led groups. As described above, this was driven by both leaders and non-leaders in the proself-led groups contributing less than leaders and non-leaders in prosocial-led groups.

**Table 4.4: Contributions to the Public Good, Punishment Rounds**

	<i>B (SE)</i>
Intercept	5.96 (.87) <sup>***</sup>
Proself leader condition	-.91 (1.26)
Peer punishment condition	.76 (1.22)
Round	.16 (.09) <sup>+</sup>
Contribution, non-punishment rounds	.19 (.02) <sup>***</sup>
Proself leader condition*Round	-.38 (.12) <sup>**</sup>
Peer punishment condition*Round	-.24 (.12) <sup>*</sup>
N <sub>rounds (groups)</sub> = 1728 (72)	

**Note:** +  $p < .10$       \*  $p < .05$       \*\*  $p < .01$       \*\*\*  $p < .001$

I also expected (Hypothesis 3a) that prosocial-led groups will be more effective than groups where sanctioning power is equally distributed to all group members. Consistent with this hypothesis, this model also shows that peer sanctioning groups contributed less over time than those led by a prosocial ( $p < .05$ ). That is, although both types of groups initially contributed at higher levels under a sanctioning system than they had under no sanctioning system at all, groups with prosocial leaders were *more* successful than the standard sanctioning system at maintaining high contributions over time. Finally, proself-led groups contributed somewhat less over time than peer-

punishment groups, but the difference was not significant ( $p = .27$ ). The findings, then, support Hypothesis 3a: prosocial-led groups, as a whole, were more successful at producing the public good than were groups using the standard solution, peer punishment. Hypothesis 3b, that proself-led groups would be worse off than peer-punishment groups, was not supported. Rather, standard peer sanctioning groups fared no better than groups led by proselfs.

### *Punishment*

That contributions in prosocial-led groups were higher than those in peer-punishment groups is noteworthy given that the risk of being sanctioned, and the amount that one could be sanctioned, was three times higher in peer-punishment groups. Each member of peer-punishment groups risked being punished by three others; in groups with leadership structures, rank and file members risked being punished only by the leader and leaders could not be punished at all. Simply based on the number of people who have the capacity to administer punishment, it is likely that significantly more resources are likely to be destroyed in groups where punishment ability is decentralized across all group members.

Indeed, Figure 4.4 above and the model in Table 4.5 show that those in peer-punishment systems were punished far more than those in either the prosocial- or proself-led groups ( $ps < .001$ ), even controlling for how much they contributed, which was itself a significant predictor of punishment ( $p < .001$ ). On the other hand, prosocial- and proself-led group members were punished to a similar extent, as already noted.

**Table 4.5: Earnings Deductions Received, Punishment Rounds**

	<i>B (SE)</i>
Intercept	5.26 (1.00) <sup>***</sup>
Proself leader condition	-.05 (1.36)
Peer punishment condition	5.17 (1.31) <sup>***</sup>
Round	-.03 (.08)
Contribution, punishment rounds	-.44 (.03) <sup>***</sup>
N <sub>rounds (groups)</sub> = 1446 (72)	

**Note:** +  $p < .10$       \*  $p < .05$       \*\*  $p < .01$       \*\*\*  $p < .001$

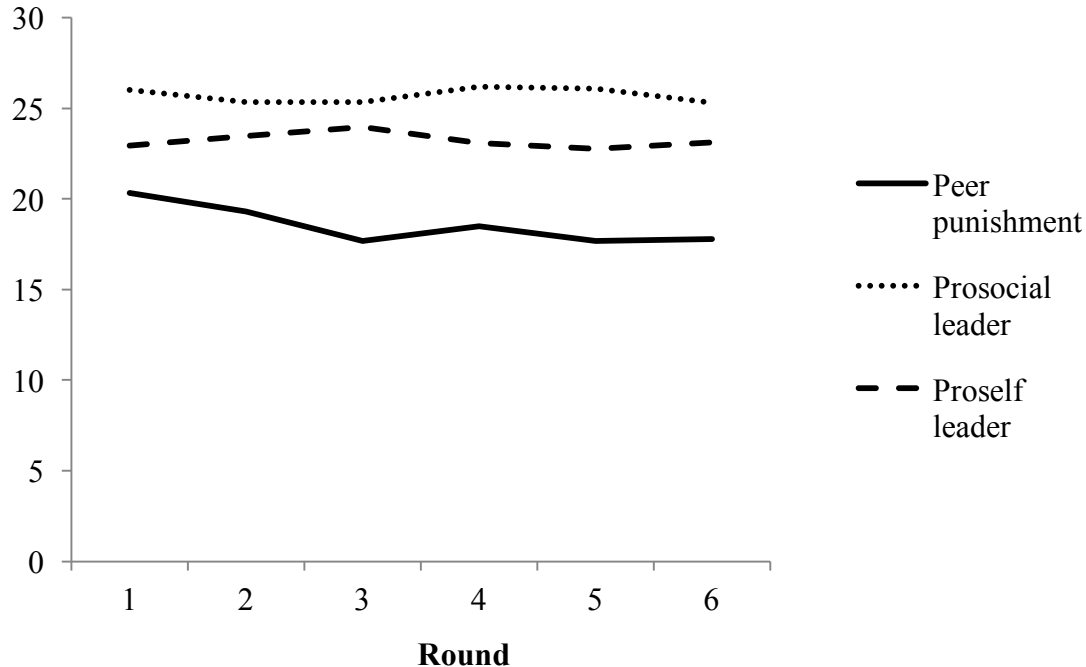
These findings on punishment are important for two reasons. First, the increased contributions observed over time in the prosocial leader condition, compared to both proself leader and peer-punishment groups, cannot be explained by a failure to punish free-riding behavior in the latter conditions. Prosocial-led groups did not produce more of the public good than their proself-led counterparts because proself leaders failed to punish—rather, the two types of leaders did not differ in their punishment behavior. And although groups with prosocial leaders also maintained high contributions over time compared to peer-punishment groups, group members led by prosocials were punished significantly *less* than members of groups using peer punishment. Second, because a single person was responsible for administering sanctions in the two leadership structure conditions, those conditions resulted in far fewer welfare-destroying sanctions, compared to the standard peer sanctioning solution. Even controlling for contributions, those in the peer-punishment condition were punished significantly more than those in groups with leaders.



### *Earnings, Punishment Rounds*

Finally I turn to differences in earnings, after accounting for punishment costs. Figure 4.6 displays group-level earnings from the public good after subtracting out costs from punishments administered to other group members and/or punishments received from other group members.

Consistent with Hypothesis 4, peer-punishment groups earned significantly less from the public good overall than groups in the prosocial leader condition ( $p < .01$ ). Indeed, although they contributed less than peer punishment groups, even proself-led groups, on the whole, earned more from the public good than peer punishment groups, due to the increased rates of punishment in the latter groups ( $p = .03$ ) (Figure 4.6 and Table 4.6).



**Figure 4.6: Average Earnings after Punishment, by Round and Condition**

**Table 4.6: Earnings, Punishment Rounds**

	<i>B (SE)</i>
Intercept	20.21 (1.49) <sup>***</sup>
Proself leader condition	3.24 (2.18)
Prosocial leader condition	5.17 (1.31) <sup>***</sup>
Round	-.48 (.15) <sup>***</sup>
N <sub>rounds (groups)</sub> = 1728 (72)	

*Note:* +  $p < .10$     \*  $p < .05$     \*\*  $p < .01$     \*\*\*  $p < .001$

On one hand, perhaps this finding should not be surprising, given that members of peer-punishment groups could be punished up to three times as much as lay-members of groups with leaders. On the other hand, the results are striking when we consider that peer sanctioning is the standard solution in the behavioral economics literature. Peer-punishment groups yielded similar contribution rates, and significantly lower earnings, than groups led by proselfs, who contributed at low levels (and influenced others to do so) and punished hypocritically.

### *Discussion*

In sum, the results from Study 1 supported the key predictions. Consistent with Hypothesis 1, prosocials who received the leader position contributed more than they had in the non-punishment rounds; proselfs who received the leader position, on the other hand, significantly decreased their contributions. This occurred above and beyond the baseline differences in their behaviors in the non-punishment rounds. Further, the results also suggest that leaders' social value orientations impact group outcomes as a whole, in support of Hypothesis 2—proself-led groups were worse off than prosocial-led groups,

which produced larger public goods. Findings revealed that the leaders' impact on group outcomes was due to leaders' influencing their group members' contributions—leaders' behaviors in particular were significantly associated with their followers' contributions in subsequent rounds. I also suggested that differences in the exercise of power among prosocial vs. proself leaders could impact rank and file contributions; however, there was no evidence that prosocials and proselfs punished at different rates. Rather, the difference between prosocial and proself leadership is driven by the impact of power on the leader's own contributions decisions, and the subsequent influence of the leaders' behavior on his or her group members.

Perhaps more importantly, another aim of Study 1 was to examine whether leadership was as effective as what is considered the standard solution to public good dilemmas in the sociological and social-psychological literature: peer to peer sanctions. In Hypothesis 3, I predicted that peer-punishment groups would contribute less than prosocial-led groups (Hypothesis 3a) but more than proself-led groups (Hypothesis 3b), given the predictions that power would increase prosocials' (and their group members') contributions but decrease proselfs' (and their group members') contributions. Hypothesis 3a was supported, but Hypothesis 3b was not—peer-punishment group members contributed at similar levels to proself-led groups. This result suggest that the peer-punishment solution is not the most viable one for increasing public good production: rather, prosocial leadership allows groups to build larger public goods. Hypothesis 4, that peer-punishment groups would earn less than leader groups after accounting for the increased punishment costs that occur by allowing all group members

to punish, was supported—thus lending more support to the claim that the leadership solution is the better solution.

Although the arguments and findings of Study 1 advance an understanding of when and why leadership promotes collective action, Study 1 had several limitations. One important limitation was that in order to compare the effectiveness of prosocial versus proself leaders, I appointed members—either the most prosocial group member or the more proself group member—to leadership. But leaders of real-world groups are often selected by their members. Thus, one important question is whether (and how) real-world groups tend to select group-beneficial, i.e., prosocial, leaders.

Although some research is suggestive (e.g., Milinski, Semmann, and Krambeck 2002), I know of no prior work that directly answers this question. But if groups are unable or unwilling to select prosocials for leadership positions, the leader “solution” may not be a solution at all. Therefore, in Study 2a I conducted a vignette study to preliminarily address whether people facing a collective action problem, like the one in the study reported above, would have selected prosocials as leaders. In Study 2b I conducted a second experiment to consider whether allowing group members to elect their leaders enhances public good production not only once the leader is elected, but also *before* the leader is elected. That is, do group members behave more prosocially when they anticipate being elected to leadership? And if so, how does this ultimately impact group members’ ability to select prosocials as leaders? Given that prosocial leaders in particular enhance public good production, determining whether (and how) real-world groups tend to select group-beneficial, i.e., prosocial, leaders, will be key for determining whether leadership will ultimately help or harm groups.

## CHAPTER 5

### LEADER ELECTION VS. APPOINTMENT

To compare the effectiveness of prosocial versus proself leaders, in Study 1 leaders were *appointed* to their positions via random assignment by the experimental program. And, following the argument detailed above, the results demonstrate that the benefits of leadership accrue mainly to groups led by prosocials. But leaders of real-world groups are often selected by their members via democratic elections rather than appointment (Grossman and Baldassarri 2012). Thus, one important question that Study 1 does not address is whether (and how) real-world groups tend to select group-beneficial, i.e., prosocial, leaders. If not, then the results of Study 1 suggest that the leader “solution” will not be a solution at all.

Some of the past work on granting sole punishment ability to a single group member has appointed the group leader (O’Gorman et al. 2009), while others have considered elected leaders (Kosfeld and Rustagi 2015; Van Vugt and De Cremer 1999) or have directly compared appointed to elected leadership (Baldassarri and Grossman 2011; Grossman and Baldassarri 2012). Yet, as described previously, in most past work considering election (or appointment) of leaders to solve public goods problems, the leader was completely *external* to the group and thus had not, and did not, contribute to group efforts. The only exceptions (Baldassarri and Grossman 2011; Grossman and Baldassarri 2012) allowed group members to select the leader from among those in their group after two rounds of the public good dilemma, but groups did so *without* knowing

each others' (i.e., the potential leaders') contribution behaviors. Rather, group members' prior contribution behaviors were kept anonymous before the leader was chosen.<sup>6</sup>

Therefore, prior studies have placed limits on whether and how group members can select their leaders on the basis of their prosocial (or proself) actions. Do groups tend to select their more prosocial members to lead?

Although no work has directly considered this question, much past research suggests that those who engage in prosocial behaviors are more likely to be rewarded with a variety of reputational and material benefits. Indeed, *reciprocity* is considered a key explanation for why people engage in prosocial behaviors toward unrelated others, even when those others may be unable to directly reciprocate their actions (Nowak and Sigmund 2000, 2005; Molm 2010; Simpson and Willer 2008). Reciprocity can be either direct or indirect (Molm and Cook 1995; Molm et al. 2007). In direct reciprocity, other-regarding behaviors are returned by the receiver, e.g., A gives to B because B gave to A. Indirect reciprocity, on the other hand, occurs when people help those who have helped others, e.g., A gives to B because B gave to C.

Both direct and indirect reciprocity can explain why people grant power and status—including leadership positions—to others. For example, past work has demonstrated that donors to charity receive more votes, from third party observers, for a

---

<sup>6</sup> Perhaps it is not surprising, then, that these studies did not find that people tended to elect their higher contributing group members to the leader position. Rather, elected leaders were no more cooperative in the public good dilemma than were their randomly appointed counterparts (Baldassarri and Grossman 2011; Grossman and Baldassarri 2012)—likely because voters could not be certain who had behaved prosocially and who had not when casting their votes.

powerful political position (Milinski et al. 2002). And those who contribute more to collective action are awarded more status by their fellow group members (Willer 2009).

More generally, that leaders and followers mutually provide benefits to each other has been considered a key explanation for why leadership has evolved among humans and other species. Price and Van Vugt's "service for prestige" theory of leadership proposes that people provide leaders with the status and prestige of the leadership position; in return, leaders are expected to provide followers with collectively shared benefits and resources. That is, people award leadership positions to those who they perceive as worthy of the position—specifically, those who are able and willing to benefit followers (Price and Van Vugt 2014).

Therefore, the work reviewed above suggests that people will prefer to award leadership to prosocials. But as noted earlier, no work has directly tested this question. Studies 2a and 2b were designed to address whether people facing a collective action problem, as in Study 1 reported above, would select prosocials as leaders. If not, the "leader solution" may not be an effective solution after all, given that proself-led groups were worse off than those led by prosocials *and* those with no leaders at all. But following past work on reciprocity and the disproportionate rewarding of power and status to more other-regarding group members, I expected that, given the ability to choose their leaders, groups tend to select prosocials for the position:

*Hypothesis 6: Given the ability to elect their leaders, people will disproportionately vote for other-regarding, vs. self-regarding, group members.*

To test Hypothesis 6, Studies 2a and 2b allow group members to elect their leaders—unlike in Study 1, where leaders were appointed. But allowing group members

to elect their leaders may provide a host of additional benefits to groups facing collective action problems. In the sections to follow, I consider how the election—versus appointment—of leaders will further benefit groups, above and beyond the benefits given by the appointed leadership structure tested in Study 1.

Specifically, I consider whether election to leadership promotes public good production in two ways. First, past work that has compared appointed to elected leader systems has consistently shown that electing leaders enhances public good production after the leader is installed (Baldassarri and Grossman 2011; De Cremer and Van Dijk 2008; Grossman and Baldassarri 2012). But second, I consider an *additional* mechanism by which leader election (vs. appointment) can promote public good production—by allowing group members to *compete* for leadership before the election takes place. I discuss this additional mechanism in the next section.

### *Do People Compete for Leadership?*

As described above, much past research suggests that those who engage in prosocial behaviors are more likely to be rewarded with a variety of reputational and material benefits. Importantly, past work also suggests that people respond strategically to the presence or absence of these benefits (e.g., Nowak and Sigmund 2005). Cooperation in social dilemmas is higher when individuals expect to be rewarded for engaging in prosocial behavior—but when the receipt of these rewards is no longer possible, cooperation decreases dramatically (Milinski et al. 2002). Relatedly, people contribute significantly more to public goods when they anticipate receiving greater status and prestige for their cooperation (Barclay and Willer 2007; Hardy and Van Vugt



2006; Willer 2009). More generally, this work suggests that people will behave in more other-regarding ways when they can receive material or non-material rewards for doing so.

Thus far, I have suggested that people will prefer to confer power and status on their more other-regarding group members. That is, given the ability to choose their leaders, groups will tend to select their more cooperative group members for the position. But importantly, the work described above suggests that when people anticipate receiving benefits (including, as discussed above, power and status) from behaving prosocially, they increase their cooperative behavior. This strategic behavior can result in *competitive* prosociality, such that people actively compete to be more generous when they can benefit from being chosen by their group members for special positions (Barclay 2004; Barclay and Willer 2007; Hardy and Van Vugt 2006). People enjoy the social and material rewards that come with power and status—and positions of leadership generally come with both (Lovaglia et al. 2006). Thus, I argue that expected or anticipated elections may trigger motivations to be elected, and that leads people to contribute more to group efforts—even *before the leader is installed*. A key consequence of this is that the possibility of leadership *vacancies* leads to higher contributions, something that has been overlooked in prior work on the leader solution to collective action problems.

Following the work described above, in Study 2b I consider whether people compete for the leadership position, by engaging in more cooperative behaviors (i.e., contribution to the public good) when they know a leader will be democratically elected by the group rather than exogenously imposed:

*Hypothesis 7a: When they can receive the leadership position via election (vs. random assignment), group members will contribute more to the public good.*

Importantly, although reputation systems are considered a key explanation for why people engage in other-regarding behavior, people still cooperate even in anonymous, one-shot interactions—that is, even in situations where material or social benefits for cooperating are not possible (Fehr and Gintis 2007; Kollock 1998). Some past work has shown that prosocials are less sensitive to the possibility of reputation formation than are proselfs (Simpson and Willer 2008). This work demonstrates that prosocials tend to cooperate whether there is the possibility for incentives or not. On the other hand, proselfs engage in other-regarding behaviors only in the presence of incentives that encourage them to do so. From this perspective, then, it is possible that the competition for leadership effect will be moderated by social value orientation. This would suggest that proselfs *in particular* will behave more cooperatively when they can compete for the leadership role. Thus, an alternative to Hypothesis 7a (which predicts that *both* prosocials and proselfs compete for leadership) is that proselfs, in particular, will compete:

*Hypothesis 7b: When they can receive the leadership position via election (vs. random assignment), proselfs, but not prosocials, will contribute more to the public good.*

Importantly, if Hypothesis 7b is supported, prosocials and proselfs who know that a leader will later be elected may contribute at similar rates—that is, prosocials contribute at high rates regardless of whether incentives are available or not, while proselfs in particular raise their contributions in order to compete for leadership. Indeed,

contributions patterns may be so similar that, when the election takes place, group members may not be able to identify prosocials for the leadership position. Rather, they might select proselfs who have competed for the position—and those proselfs may or may not continue contributing at high rates once they have received leadership. Thus not only does Study 2a consider whether people prefer prosocials as leaders (Hypothesis 6), Study 2b does as well: can people identify prosocials and select them as leaders—even when competition for leadership is possible?

Hypotheses 7a and 7b suggest that elected leadership systems (vs. exogenously imposed leadership) may benefit groups facing public good dilemmas *even before the leader is installed*, if people compete for leadership. While prior work has overlooked this mechanism by which elected leadership may benefit public good production, some past research has considered whether elected vs. appointed leadership impacts public good production *after* the leader is installed. Aside from competition for leadership before the leader is elected, in the next section I also consider whether the election process will promote cooperation *after* the elected leader receives the position. That is, do groups produce larger public goods when they are led by an elected (vs. appointed) group member?

### *Leader Election vs. Appointment*

Past work has consistently shown that allowing group members to elect their leaders is associated with higher levels of contribution in public good dilemmas (Baldassarri and Grossman 2011; Dal Bo, Foster, and Putterman 2010; Grossman 2013; Grossman and Baldassarri 2012). The existing literature suggests that the leader

selection process—specifically, allowing group members to participate in the selection of their leader—affects *both* the followers’ and the leaders’ behaviors.

Among followers, people prefer elected rather than appointed leaders (Van Vugt and De Cremer 1999). Part of the reason is that elected leaders are perceived as more legitimate (Weber 1978; Walker and Zelditch 1993). Allowing followers to participate in the leader selection process enhances followers’ perceptions of institutional legitimacy (Baldassarri and Grossman 2011; Dal Bo, Foster, and Putterman 2010), and leaders are expected to exert their authority—and others are expected to comply—to a greater extent when leaders are legitimated (Burke, Stets, and Cerven 2007). Thus experimental work comparing appointed versus elected leadership has consistently shown that followers are more willing to cooperate with institutions that they themselves have selected—democratically chosen institutions are associated with greater cooperation rates than exogenously imposed institutions (Baldassarri and Grossman 2011; Dal Bo, Foster, and Putterman 2010; Grossman and Baldassarri 2012).

Aside from followers’ enhanced cooperation with elected leaders, election to the position may also enhance leaders’ *own* cooperative behaviors. Lab and field experiments show that elected leaders feel more responsible and accountable to their followers, leading to greater cooperative behavior (Grossman 2013). Followers of elected leaders are less accepting of norm violating behavior by their elected leaders, and elected leaders anticipate this by behaving more fairly; elected leaders also report feeling more social responsibility toward their group members (De Cremer and van Dijk 2008). And cross-national studies have shown that elected leaders provide more public goods (Lake and

Baum 2001; Bueno de Mesquita et al. 2004; Stasavage 2005) and are less corrupt (Mauro 1995) than their appointed counterparts.

Of course, election may also result in enhanced cooperation among leaders and followers because election allows groups to select more *effective* leaders—i.e., prosocials. But there is little evidence to support this claim. For instance, Baldassarri and Grossman (2011) do not find evidence that voters elect leaders who are more effective in terms of either their own cooperation or in their willingness to use sanctions. Indeed, they find no difference in appointed vs. elected leaders' prosociality or use of punishment, although they still find that people are more cooperative when leaders are elected. Rather, they find that people are more likely to vote for leaders based on status markers like wealth and gender—although they also do not find evidence that wealthier and male leaders are more effective than other types of elected leaders (Baldassari and Grossman 2011; see also Grossman and Baldassarri 2012). This work suggests that election vs. appointment to leadership enhances cooperation above and beyond people's ability to select more effective leaders when the leader can be selected—specifically, due to the mechanisms described above, including enhanced feelings of social responsibility in the leader and of legitimacy in followers.

The main goal of this study is to consider whether public good production is enhanced before leaders are chosen by the group—that is, whether people compete for leadership. But I also consider whether, following past work, election to leadership promotes cooperation once the leader has been installed. Importantly, I compare cooperation under elected vs. appointed leaders even when leader social value orientation is held constant across elected vs. appointed leaders. That is, Study 1 has already

demonstrated that prosocials are more effective leaders than proselves. But above and beyond this effect, following the work described above, I expect elections to promote cooperation after the leader has taken his or her position:

*Hypothesis 8: Compared to groups with appointed leaders, groups with elected leaders will contribute more to the public good.*

In sum, in the studies described in the next chapter, I examine whether people prefer prosocials for leadership (Hypothesis 6, Studies 2a and 2b). I also consider, in Study 2b, whether leader elections promote public good production via two mechanisms.

Specifically, does holding a group election promote contributions before the leader is installed (Hypotheses 7a and 7b)? And does it continue to increase cooperative behavior after (Hypothesis 8) the leader is installed? The next chapter describes the studies I conducted to test these hypotheses.

## CHAPTER 6

### STUDIES 2A AND 2B

In this chapter, I outline, first, the vignette study (Study 2a) and second, the experiment (Study 2b) designed to test the hypotheses given in the previous chapter. Specifically, the studies described here allow me to address whether group members who can choose their leaders will prefer prosocials for the leadership position (Hypothesis 6, Studies 2a and 2b). Study 2b also allows me to consider whether people compete for leadership positions by contributing to the public good to a greater extent when their group members will later be able to elect a leader (Hypotheses 7a and 7b); importantly, I also address whether or not this competition for leadership prevents people from being able to effectively identify prosocials when choosing leaders. Finally, Study 2b tests whether, as shown in some past work, allowing groups to elect their leaders (rather than randomly appointing them) promotes public good production once the leader is installed (Hypothesis 8).

#### *Participants and Design, Study 2a*

Study 2a was a first step in considering whether people (in this study, third-party observers to a group producing a public good) select prosocials as leaders. The study was conducted via Amazon's Mechanical Turk (Buhrmester, Kwang, and Gosling 2011; Weinberg, Freese, and McElhattan 2014). 50 Amazon Mechanical Turk (AMT) users located in the United States were asked to imagine that they were a group member

participating in a decision-making task, as described in more detail below. Participants were paid \$0.50 for completing the survey, which took about ten minutes to complete (this is a relatively high payment for AMT; see Weinberg et al. 2014). Four respondents who responded incorrectly to at least one of two basic attention check questions (described in the next section) were omitted from analyses. The final sample size was 46 participants (56.5% male).

### *Procedure, Study 2a*

Respondents in Study 2a were presented with task instructions nearly identical to those given in Study 1 (see Appendix C for the full text of the instructions for Study 2a). Specifically, they read a brief description of the original public good task adapted from the Study 1 instructions. After reading these basic instructions, they read through several examples and answered quiz questions to ensure their understanding of the task.

Next, the participants were presented with a table containing the actual round-by-round contribution decisions made in the non-punishment rounds for two randomly selected groups from Study 1. One of these groups had been assigned to the prosocial leader condition and the other had been assigned to the proself leader condition; the order in which these two groups were presented to AMT users was counterbalanced. Specifically, participants in the AMT study viewed a table displaying each of the four Study 1 participants' contributions, expressed in dollars, across each of the six rounds. The instructions on this screen also reminded AMT participants that the group members could have contributed anywhere from \$0 to \$20 in each of the six rounds, and asked



participants to review the contributions decisions and become familiar with them, but noted that participants did not need to memorize the table (see Appendix C).

Participants then read a new set of instructions that described how the rules would change for subsequent rounds—i.e., that one group member (the “deductor”) would be able to make deductions from his or her group members’ earnings. The description of the rule change, as the description of the original task, was adapted from the instructions from Study 1. They could then view the contributions table again. Above the table was a reminder that 1) the table presented group members’ contributions in the task over six rounds and 2) that the group members did not know that the rules would change until after they made the decisions given in the table. Below the table they were prompted to imagine “*that you were a member of the group, and that you could select a participant to the role of Deductor for the next phase of the study.*” They were asked to indicate which of the four group members they would prefer to be appointed to the deductor role.

After selecting their preferred group member, participants were prompted to write a few sentences explaining their choice. Next, they were asked to select the participant who contributed the most and the participant who contributed the least over the six rounds, as basic attention check questions. Respondents could view the contributions table while they answered these questions. Finally, participants repeated this process once more (“for a different group”).

### *Results, Study 2a*

Participants in the Mechanical Turk study were not given information about group members’ social value orientations. Rather, they viewed the group members’ contribution

decisions in the non-punishment rounds. However, consistent with past work (see Balliet et al. 2009 for a review), Study 1 participants' SVO significantly predicted their contributions in the non-punishment rounds ( $p < .01$ ), such that prosocials tended to contribute more than proselfs. Therefore, I considered whether Study 2a participants tended to select high contributors as leaders, given that, as in much past work (Balliet et al. 2009), Study 1 high contributors tended to have a prosocial social value orientation.

30 participants (65%) preferred the highest contributor for the leadership position in one of the scenarios and 28 participants (61%) preferred the highest contributor in the second scenario. Importantly, both of these highest contributors had been classified as prosocial. Participants who did not prefer the highest contributor selected one of the other three participants for the leader position at approximately equal rates. A chi-square goodness of fit test revealed that the preference for the highest contributor for leadership was highly significant in both scenarios ( $\chi^2(3) = 41.48, p < .001$ , Cohen's  $w = .95$  and  $\chi^2(3) = 33.13, p < .001$ , Cohen's  $w = .85$ , respectively).

That is, participants tended to select the group member who was most cooperative to be the leader, in both scenarios. And in both scenarios, the most cooperative group member had been categorized as prosocial. Indeed, responses to the open-ended question revealed that participants preferred the highest contributor for leadership because that participant, for example, "sacrificed the most for the group", "was the most generous", or "contributed the most to the pool over the course of the game...he would be the fairest of all the players and make fair choices as the deductor". Thus, Hypothesis 6 was supported: prosocials generally were preferred for positions of leadership, and if given the

opportunity to select the leader, Study 1 participants would likely have selected a prosocial.

Study 2a was a first step in considering whether people prefer prosocials to leadership. However, it was a vignette study such that third party participants (rather than contributing group members) were asked to indicate who they would prefer for leadership, if they were a member of the group. Perhaps more importantly, this study cannot tell us whether election to leadership—rather than appointment—can enhance public good production, aside from giving group members the ability to select effective (prosocial) leaders. Therefore, I conducted a follow-up study to directly compare appointed to elected leader structures—do elections to leadership promote public good production, both before and after the election takes place? Study 2b was designed to address these issues.

#### *Participants and Design, Study 2b*

Participants were recruited from large introductory classes at the University of South Carolina. Each session was conducted in a group of four and each group was randomly assigned to one of two conditions: *elected leader* or *appointed leader*. 65 groups (260 participants; 69% female) completed the study in exchange for payment.

All groups first completed a standard repeated version of a public good dilemma with no punishment ability. Then, one group member was either elected (*elected leader condition*) or randomly assigned (*appointed leader condition*) to be group leader in the punishment phase. As in Study 1, the group leader was given the sole ability to deduct

from the others' earnings following each round of the punishment phase of the public good dilemma.

### *Procedure, Study 2b*

Upon their arrival to the lab, participants were escorted to a private room where they were unable to see or communicate with other group members. Participants completed a consent form before beginning the computerized task, programmed using z-Tree version 3.3.11 (Fischbacher 2007). The task began with a series of pre-study questions, including the same social value orientation (SVO) measure used in Study 1 (Van Lange et al. 1997; see also Appendix A). As in Study 1, I used scores on the SVO scale to assign the leader in the appointed leader condition, as described below. At no point during the study were participants told that their responses on the SVO scale would be used (in the appointed leader condition) to determine who would occupy the leader role.

After everyone completed the pre-study questionnaire, the computer advanced to instructions for a public good dilemma without punishment. The description of the task was similar to those used in Study 1—instructions began by assuring participants of their anonymity. Then, participants were told that the study consisted of several parts. Part 1 would consist of a series of choices in an “investment decision task”; participants were told that their and their group members' decisions in the task would affect their earnings in the study. As in Study 1, participants read the description of the public good

(“investment decision”) task, read through several examples, answered several quiz questions, and then completed two practice rounds<sup>7</sup> to ensure their comprehension.

After participants completed the instructions for Part 1, they were told that, to save time, they would now learn about Part 2 of the study before beginning the actual task (see Appendix D for the full text of the instructions for Study 2b). Specifically,

*In Part 2, you will continue to decide how many points, from 0 to 20, you wish to contribute to the group fund, and how many to keep for yourself. However, in Part 2, one group "leader" will have the opportunity to deduct points from their fellow group members' earnings following each round.*

Unlike in Study 1, participants in Study 2b learned about the punishment phase before beginning the non-punishment phase so that competing for leadership (in the election condition) was possible. That is, participants in Study 2b were aware that later a leader would be selected or elected (depending on condition), and they thus had the opportunity to (potentially) alter their contribution patterns in the first phase as a result of this knowledge. Aside from the changes to the order in which participants were given the instructions, the instructions were similar to the punishment phase instructions for Study 1: leaders could spend anywhere from 0 to 10 of their own points to deduct anywhere from 0 to 30 points from each other group member.

---

<sup>7</sup> For the practice rounds, participants were (correctly) told that the other group members' choices were simulated by the computer, and that earnings in the practice rounds would not be counted toward their earnings during the actual study.

Next—prior to beginning the non-punishment phase—the instructions described how the leader for Part 2 would be selected. This was the only portion of the instructions that differed by condition, either appointed or elected [instructions for elected leader condition in brackets]:

*The leader for Part 2 will be randomly selected by the computer [elected by the group] at the end of Part 1.*

*[Specifically, after Part 1, you (and each of the others) will see each group members' participant numbers and their total contributions to the group fund in Part 1. Each group member will then be able to vote for one person to be the leader (you cannot vote for yourself). The winner of the vote will receive the leader role.]*

*Finally, after the leader is randomly selected [elected], participant numbers will change. Therefore, [although your behavior in Part 1 may determine whether you are elected to be the leader for Part 2,] at the beginning of Part 2 no one (including the leader) will be able to identify you based on your behavior in Part 1.<sup>8</sup>*

---

<sup>8</sup> In order to consider whether people compete for leadership by engaging in more cooperative behavior prior to an election, participants knew in advance that a punishment phase would come after the non-punishment rounds of the study. But this could also result in participants contributing at high rates in the non-punishment phase because they anticipate that, once the punishment phase begins, leaders will punish them for non-cooperative behavior in the previous (non-punishment) phase. More importantly, is it possible

As they did for Part 1, participants answered quiz questions and completed a practice round. Finally, participants read a brief review of the entire study: they reviewed the instructions for Part 1, the leader selection process (depending on condition, either appointed or elected), and Part 2, before beginning Part 1 (the non-punishment phase) of the study.

Once everyone was ready to begin, as in Study 1, each participant entered a whole number from zero to twenty to indicate how much they wished to contribute to the group fund. After all participants had answered, everyone was able to see their own contribution and earnings, as well as the contributions and earnings of each other group member. Once everyone finished reviewing the results, participants proceeded to the next round, where they again decided how much of a new twenty-point endowment to contribute to the group. This process was repeated for a total of nine rounds. Participants were not told in advance how many rounds they would complete.

Once the ninth round was completed, participants in the election condition were reminded that they would now vote for a leader for Part 2 (the punishment phase). Specifically, they were presented with a table giving each participant's total contributions to the group fund across all rounds of the non-punishment phase. Participants could select

---

that this expectation could differ by condition such that elected leaders (vs. appointed leaders) are perceived as more (or less) likely to punish others for their past behaviors. This could lead to differences between conditions in the non-punishment phase due not to competition for election, but to differences in *expectations of being punished* in the upcoming punishment phase of the study. I told participants that participant identifiers would change from the non-punishment phase to the punishment phase to eliminate this alternative explanation for increased contributions in the election condition.

one group member for whom they wished to vote. Participants could not vote for themselves, nor could they see who voted for whom. The computer program selected the leader based on the group member who had received the most votes. In case of a tie, the elected leader was selected at random from among the group members who had received the most votes.

Groups in the appointed leader condition were yoked to groups in the elected leader condition, such that there was the same number of groups with prosocial (and proself) leaders in both the elected and appointed leader conditions. Given the Study 1 result that (appointed) prosocials are more effective leaders than proselfs, this ensures that differences between these two conditions can be attributed to the election process, rather than to differences in leader SVO across conditions. In the appointed leader condition, the computer either chose a group member who could be categorized as prosocial, proself, or unclassified to be the leader, depending on the social value orientation of the elected leader in a previous group from the election condition. Ties were broken by selecting at random from among the group members who fell into the assigned social value orientation category. In both conditions, once the leader was successfully elected or appointed, participants were shown one final screen where they were told whether they had received the leader position or not.

The rounds with punishment began similarly to the rounds without punishment. All group members made simultaneous contribution decisions and then viewed the results, including each group member's contribution and earnings. Then, non-leaders in the leader conditions were told to wait while the leader made his or her deduction decisions. Those who were leaders were again shown each other group member's



contribution and earnings. Below that information, they indicated whether (and if so, how much) they wished to deduct from each person's earnings. All group members could view their and others' outcomes after deductions—including how many deductions they had received and their final earnings for the round. This process was repeated for nine rounds. As in the no-punishment phase of the study, participants were not told the number of rounds in advance.

Once the public good task was completed, participants were paid based on their earnings across the eighteen rounds of the study (ranging from \$10 to \$15) and debriefed. Each session lasted approximately forty-five minutes.

### *Results, Study 2b*

As in Study 1, I used the nine-item social value orientation scale, completed at the beginning of the study, to categorize participants' social value orientation. Following past work, I combined individualists and competitors into one "proself" category (Simpson and Willer 2008; Van Lange et al. 1997). As is standard in the social value orientation literature (Simpson and Willer 2008; Van Lange et al. 1997), a participant was classified as a given social value type when he or she answered at least six of the nine items consistently.

Five groups were omitted from analysis because they contained one or more participants who either reported confusion about the instructions (one group in the appointed leader condition), missed all five of the quiz questions designed to check understanding of the task instructions (three groups, one in the elected leader condition and two in the appointed leader condition), or a computer problem during the study

prevented them from completing both the non-punishment phase and the remainder of the study (one group in the appointed leader condition). After omitting these groups, analyses were performed on the remaining 60 groups (30 groups each in the appointed and elected leader conditions). Of the 240 participants included in the study, 124 (52%) were classified as prosocial, 84 (35%) as proself, and 32 (13%) could not be classified. As in Study 1, all data were analyzed using multilevel modelling to account for the nested structure of the data.

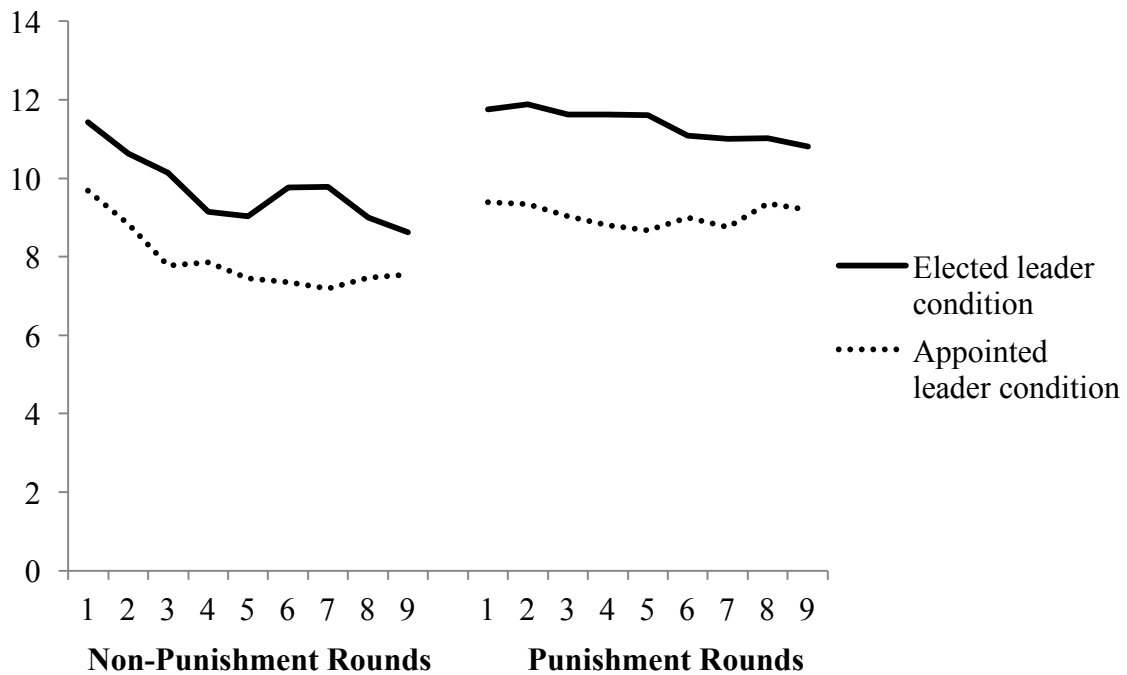
### *Contributions, Non-punishment Rounds*

Do people compete for leadership positions? In the non-punishment rounds, participants were aware that later (depending on condition) they would vote for a leader or a leader would be randomly assigned. Any significant differences between conditions in contributions in the non-punishment round, then, would be due to this knowledge of the upcoming leader selection process. If people contribute more when they anticipate election to leadership more than when they anticipate appointment to leadership, it would suggest that people compete for leadership positions.

Results suggest that people do compete for leadership, starting in the very first round. In a model considering behavior in Round 1, there was a main effect of condition: those who were told that they would later elect a leader contributed more than those who were told that a leader would later be randomly assigned ( $p = .02$ ) (see non-punishment rounds, Figure 6.1 and Table 6.1, Model 1). This finding is notably different from that of Study 1, where participants contributed similarly in Round 1 of both the non-punishment and punishment phase of the study—differences between conditions took time to emerge

as the leader responded to his or her ability to punish. In Study 2, participants immediately responded to the knowledge that a leader would later be elected (vs. randomly assigned)—and they did so by contributing more.

**Figure 6.1: Group-Level Contributions in Non-Punishment and Punishment Rounds, by Condition**



**Table 6.1: Contributions to the Public Good, Non-Punishment Rounds**

	Model 1: Round 1 <i>B (SE)</i>	Model 2: Rounds 2-9 <i>B (SE)</i>
Intercept	9.68 (.51) <sup>***</sup>	8.40 (.67) <sup>***</sup>
Elected leader condition	1.76 (.73) <sup>*</sup>	2.00 (.94) <sup>*</sup>
Round		-.18 (.07) <sup>*</sup>
	$N_{\text{participants (groups)}} = 240 (60)$	$N_{\text{rounds (groups)}} = 1920 (60)$

*Note:* +  $p < .10$  \*  $p < .05$  \*\*  $p < .01$  \*\*\*  $p < .001$

The pattern of higher contributions in the election condition held for subsequent rounds as well: in Rounds 2 through 9, the main effect of condition remained significant ( $p = .04$ ); those in the election condition were more cooperative—even though the election had not yet taken place. As is typical in public good dilemmas without punishment (Ostrom 2000; Sell and Wilson 1991), there was a main effect of round ( $p < .05$ ): contributions declined across the nine rounds. The condition  $\times$  round interaction was not significant ( $p > .7$ ), suggesting that contributions fell across the nine rounds to a similar extent in both conditions.

These findings demonstrate that participants do compete for leadership positions, which results in larger public good production than when competition for leadership is not possible. However, results also show that competition for leadership alone does not solve the main problem posed by public good dilemmas: that contributions decline over time. (e.g. Ostrom 2000; Sell and Wilson 1991). Rather, contributions declined over time to a similar extent in both conditions.

Although the knowledge that a leader would later be elected did not prevent contributions from falling, did it prevent group members from being able to identify prosocials for leadership? I also considered whether the increased contribution among those in the election condition occurred particularly among proselfs (Hypothesis 7b) versus among all group members (Hypothesis 7a) by adding terms into the models given above for participants' social value orientation. That is, were proselfs in particular more likely to compete for leadership?

In both models, as in Study 1 and much past work (see Balliet et al. 2009 for a review), social values emerged as a significant predictor of contributions: prosocials

contributed more than proselves ( $p < .001$ ) and unclassified participants ( $p < .01$ ). Proselfs and unclassified participants did not differ from each other in their public good contributions ( $p = .35$ ). After controlling for social values, the main effect of condition remained a significant predictor of contributions in round 1 and in rounds 2-9 ( $p < .01$  and  $p = .02$ , respectively). But social value orientation did not interact with condition to predict contributions in the non-punishment phase ( $ps \geq .15$ ). This suggests that people competed for leadership by contributing more when they anticipated an election, but to a similar extent, regardless of their social value orientation. Hypothesis 7a, then, was supported while Hypothesis 7b was not. People do compete for leadership, and they do so to a similar extent regardless of their social values. The result, however, is that people can still identify prosocials and elect them to leadership positions.

#### *Who Was Elected to Leadership?*

As in Study 2a, participants did not know the social value orientation of their fellow group members when choosing leaders. However, they did know how much each other group member had contributed in the non-punishment phase of the study. And they tended to vote for the highest contributor of the three others for whom they could vote (participants could not vote for themselves). 59% of votes cast were for the highest contributing group member among the three other group members from which each participant could choose. The remaining 41% of votes cast were split about equally between the lowest and middle contributor. These patterns are similar to those observed in Study 2a, as described above, and lend additional support to Hypothesis 6.

The result was that 27% of votes cast were for proselves, while 62% of votes cast were for prosocials and the remainder were cast for unclassified participants. Ultimately, a majority (67%) of elected leaders in the study were categorized as prosocial—of the 30 elected leaders, 20 were prosocial. (Given that groups in the appointed leader condition were yoked to groups in the elected leader condition, the distribution of leader SVOs in the appointed condition was identical to that in the elected leader groups.) A chi-square test revealed that prosocial SVO was associated with leadership status in the election condition, such that prosocials were overrepresented as leaders while proselves and unclassified participants were more often followers,  $\chi^2(1) = 5.38, p = .02$ .<sup>9</sup>

The results thus far suggest that participants engaged in more group-sacrificial behaviors when they knew that a leader would later be elected (vs. appointed). Further, within groups, those group members who engaged in more group-beneficial behaviors received more votes and, ultimately, were more likely to receive the leadership position. Although people compete for the leadership position, this does not appear to prevent participants from ultimately selecting the more other-regarding group members among

---

<sup>9</sup> Of course, aside from a disproportionate preference for prosocial leadership, another factor that will determine whether prosocials receive leadership positions is simply how many prosocials are present in the group. Indeed, of the 10 groups that did not elect a prosocial leader, three contained no prosocial within the group to elect. Unlike Study 1, where leader SVO was determined by condition (and thus, was balanced across conditions), in this study whether the leader is prosocial or prosself was partially determined by the SVO makeup of the group. However, results still suggest that prosocials are disproportionately preferred for leadership: of the 11 election condition groups that had exactly two prosocials and two proselves, 8 groups (73%) selected a prosocial for leadership—significantly greater than the 50% of groups that would be expected by chance,  $\chi^2(1) = 9.09, p < .01$ .

them to lead. Rather, competition for leadership appears to be one mechanism by which election, vs. appointment, of leaders can benefit the group—by promoting more other-regarding behavior ahead of the election, resulting in larger public goods. In the next section, I consider how elections can further enhance public good production—what happens *after* the elected (or appointed) leader is installed?

### *Contributions, Punishment Rounds*

Although competition for leadership enhanced public good production during the non-punishment rounds, as described earlier, it did not keep contributions high over time. How did the introduction of punishment impact contributions? As in the non-punishment phase of the study, and importantly, controlling for their (higher) contributions in the non-punishment rounds, the contributions of members in groups with *elected* leaders were significantly higher in the punishment phase, compared to groups with exogenously appointed leaders ( $p = .03$ , see punishment rounds, Figure 6.1 and Table 6.2, Model 1). This finding supports Hypothesis 8 and replicates some prior work suggesting that groups who elect their leaders produce larger public goods than groups whose leaders inherit or are appointed to the role (Baldassarri and Grossman 2011; Grossman and Baldassarri 2012). Importantly, the difference between elected and appointed leadership structures held even though the leader type (prosocial or proself) was held constant across the two conditions. This allows me to rule out that elected leadership results in enhanced public good production simply because elections allow people to select those leaders that Study 1 suggests will be more effective—prosocials.

**Table 6.2: Contributions to the Public Good, Punishment Rounds**

	Model 1: Rounds 1-9 <i>B (SE)</i>	Model 2: Rounds 1-9 <i>B (SE)</i>
Intercept	7.87 (.80) <sup>***</sup>	7.88 (.81) <sup>**</sup>
Elected leader condition	2.35 (1.10) <sup>*</sup>	2.28 (1.12) <sup>*</sup>
Round	-.04 (.07)	-.04 (.07)
Contribution, non-punishment phase	.15 (.02) <sup>**</sup>	.15 (.02) <sup>**</sup>
Is leader		.01 (.50)
Is leader*Elected leader condition		.30 (.70)
	N <sub>rounds (groups)</sub> = 2160 (60)	N <sub>rounds (groups)</sub> = 2160 (60)

*Note:* \*  $p < .05$       \*\*  $p < .001$

The difference between the elected leader and appointed leader conditions in the punishment phase was observed even after controlling for the differences in contributions in the non-punishment rounds—which also predicted contributions in the punishment rounds ( $p < .001$ ). A second model containing terms for whether the group member was the leader or not (Table 6.2, Model 2) revealed that leadership status did not impact contributions—there was no difference between leaders and non-leaders in their contributions to the group ( $p = .98$ ), nor was there an interaction between leader status and condition ( $p = .67$ ). Rather, results suggest that *all* members of groups who democratically elect their leaders—both leaders and followers alike—contribute at higher rates than those leaders and followers of groups where leaders are appointed exogenously.

Finally, unlike in the non-punishment phase, where contributions declined over time, there was no effect of round. Rather, the introduction of punishment was associated with steadily high contributions over time ( $p = .86$ ). Further, round did not interact with condition, suggesting that the introduction of either an appointed or an elected leader was



sufficient to keep contributions high across the punishment rounds ( $p = .42$ ). However, group members' contribution decisions, across all rounds, were significantly higher when the leader was elected (see punishment rounds, Figure 6.1). Taken as a whole, the results suggest *groups who elect their leaders produce larger public goods—both before and after the leader is selected.*

### *Punishment Use*

Explanations for why elected leaders are more effective than appointed leaders typically focus on perceptions of enhanced perceptions of institutional legitimacy when the leader is elected (Baldassarri and Grossman 2011). However, I also consider whether there are differences between elected and appointed leaders in how they wield their punishment power. Can increased contributions under elected leaders be explained by differences in their use of punishment?

Results revealed, first, that leaders tended to direct punishment toward those group members who contributed less to the public good ( $p < .001$ ). This was the case for both elected and appointed leaders, i.e., there was no condition  $\times$  group member contribution interaction. After controlling for group members' contributions, results revealed that leaders in the election condition punished slightly, but not significantly, more than their appointed counterparts ( $p = .14$ ). These results demonstrate that differences between elected leader and appointed leader groups was not driven by any differences in the use of punishment.

Turning to differences in punishment based on leader social value orientation, recall that Study 1 demonstrated that prosocial leaders and proself leaders did not differ

in the frequency with which they directed punishment at their fellow group members. These results held for Study 2b: prosocial leaders tended to punish somewhat more than proselves, but the difference was not significant ( $p = .07$ ). This effect did not differ by condition (i.e., there was no prosocial leader  $\times$  election interaction,  $p = .27$ ).

### *Discussion*

Leaders of real-world groups are often selected by their members (Baldassarri and Grossman 2011; Grossman and Baldassarri 2012). Thus it is important to know whether real-world groups tend to select their more other-regarding members to lead them. If groups are unable—or unwilling—to select prosocials for leadership, then the results from Study 1 suggest that leadership will be detrimental to groups' public good production. In that case, the leader “solution” will not be a solution at all.

In Study 2a, third party participants could view the behavior of group members facing a collective action problem and select from among those group members whom they would prefer for leadership. Results suggested that these participants preferred those group members who engaged in more other-regarding behaviors; if given the chance to elect a leader, they would have overwhelmingly selected prosocials. Study 2b allowed internal group members to select a leader from among those in the group. Again, participants showed a marked preference for prosocials to receive the position.

Study 2b also suggested that, supporting past work, groups with elected leaders contribute more to the public good after the leader is installed. Both leaders and their followers contributed more when the leader had been elected (vs. appointed). I matched leader type across conditions, so the observed difference between elected and appointed

leaders cannot be attributed to participants in the election condition being able to select more effective leaders. Past work has attributed this difference to enhanced feelings of social responsibility among leaders and enhanced feelings of legitimacy among followers. The results from Study 2b support this, but did not directly test these mechanisms. I return to this issue in the final chapter.

Perhaps more importantly, Study 2b also suggests that allowing group members to elect their leaders enhances public good production even *before* the leader is installed. That is, the results demonstrated that people do compete for leadership, contributing more to the public good when they are told a leader will later be elected than when they are told a leader will later be randomly assigned. Importantly, this competition for leadership does not prevent people from being able to elect prosocials. Because both prosocials and proselfs increase their contributions when leaders will be elected, groups are still successfully able to identify prosocials and install them in leadership positions. Given that prosocial leaders in particular enhance public good production, determining whether (and how) real-world groups tend to select group-beneficial, i.e., prosocial, leaders, will be key for determining whether leadership will ultimately help or harm groups. Studies 2a and 2b were designed to address whether groups select their other-regarding group members to lead them. In addition, Study 2b has demonstrated that allowing groups to select their leaders can enhance public good production both before and after elections. In the final chapter, I summarize the results of both studies, and consider their implications. Finally, I suggest new directions for future research on leadership and collective action.

## CHAPTER 7

### CONCLUSION

Here I have argued that group leadership can be an effective solution to collective action problem—given that the leader is prosocial. According to the logic outlined in previous chapters, I have suggested that appointing a single group member to leadership would impact both leader behaviors and group outcomes via *power* and *influence* processes. Specifically, following Chen et al. (2001), I argued that the acquisition of power would magnify baseline tendencies such that proself individuals become even more self-oriented (and thus, contribute even less), while prosocials become more group-oriented (and thus, contribute more), upon obtaining a leadership position.

More importantly, I expected that leaders' contribution behaviors would shape group-level contribution norms—and ultimately, group outcomes—via an *influence effect*. I also considered whether prosocial and proself leaders would produce different group-level outcomes via differences in their exercise of punishment (*second-order power effect*). Although prosocial and proself leaders punished at similar levels, proselfs tended to engage in hypocritical punishment, punishing those who contributed as much or more than they themselves had.

In sum, I have shown that leadership structures can sustain cooperation in collective action groups. But the benefits of leadership accrued mainly to groups with *prosocial* leaders: these groups both contributed more to the public good over time, and earned more, than groups with proself leaders or with the standard peer sanctioning

system. Importantly, I demonstrate that groups *do* prefer prosocials for leadership, and if given the ability to elect a group leader, they typically select their more other-regarding group members to lead.

In addition, allowing groups to choose their leaders enhances public good production above and beyond their ability, and propensity, to choose their more other-regarding group members for the position. Holding an election for group leadership promotes contributions to the public good not only after the leader is installed, but also *before* the leader is chosen—those who know a leader will later be elected contribute more than those who know a leader will later be randomly assigned. Thus competition for the leadership position also enhances cooperative behavior. Groups that hold democratic elections for leadership may induce more cooperative behavior in their members *even before leadership hierarchies emerge*. Past research has thus far overlooked this additional mechanism by which groups may produce larger public goods. Further, I also find that, as in past work, once the leader is installed groups with elected leaders contribute at higher rates.

These arguments and findings have a number of important implications for understanding collective action. As noted earlier, previous work has suggested that while standard peer sanctioning systems can increase cooperative behavior, they have important drawbacks. I found that a prosocial leader given sole sanctioning ability increases contributions significantly more than the standard solution while avoiding those problems. For instance, while peer sanctioning systems can result in over-punishment, results from my experiment show that leadership structures greatly reduce that risk.

Just as importantly, while groups with prosocial leaders fared substantially better than all other groups, even groups led by proselfs came to earn more than those with peer sanctioning systems in place. This is a remarkable finding given that peer sanctioning is the most prominent solution to collective action problems in behavioral economics (for a review, see Fehr and Gintis 2007). That members of peer sanctioning groups contributed no more—and *earned less*—than groups led by proselfs (who decreased their contributions upon becoming leaders and sanctioned fellow group members hypocritically) suggests that we should rethink the status of peer sanctioning as a favored solution to collective action problems.

#### *Limitations and Directions for Future Research*

Although the arguments and findings here advance an understanding of when and why leadership promotes collective action, like any study, the studies presented here leave open questions that should be addressed in future research. For instance, in both studies I considered a limited conception of power, rooted in the ability of leaders to punish rank and file group members. But this is only one way leaders can exercise power and future research might consider others, e.g., leaders who can also administer (or withhold) rewards.

Study 1 reported above is a first investigation into the dynamics and relative effectiveness of two very different solutions to collective action problems: the standard solution, which distributes the power to punish equally among peers, and a system that centralizes punishment in the hands of a single leader internal to the group. Study 2 considered whether leadership systems were more effective when leaders were elected

rather than appointed, given that elected leaders can typically compete for election.

Therefore, as a first step, Study 1 treated peer punishment and leadership as mutually exclusive systems, and Study 2 considered only leadership systems.

But real world groups are often characterized by combinations of formal or “top down” sanctions and informal norms among peers, systems that can work together or in opposition (Homans 1961; Kitts 2006; Nee and Ingram 1998; Shibutani 1978). An important next step would be to extend the work here to address the joint effects of these two systems. For instance, it seems likely that the greater personal sacrifice of prosocial leaders and the tendency for them to punish rank and file members who contributed substantially less than they themselves had would have led to the emergence of informal norms among rank and file members that complement the “formal” sanctions administered by the leader. On the other hand, the tendency for proself leaders to hypocritically punish rank and file members who contributed more than the leader himself had contributed would be more apt to generate oppositional norms, characterized by disapproval or punishment of peers who contributed “too much” to the collective good. Likewise, group members who elected their leaders may be more likely to develop norms for how the leader is expected to behave, perhaps to a greater extent than leaders who were appointed. This is an important question for future research.

Relatedly, in the groups investigated here, leaders faced no risk of being sanctioned. Nor did they risk being removed from the leadership position once they had obtained it, regardless of their contribution or punishment behaviors. But rank and file members of real world groups may direct sanctions not only at each other but also at leaders. For instance, group members often can – and sometimes do – act together to oust

ineffectual leaders. The potential for loss of leadership positions, or the prospect that informal norms can emerge in opposition to a leader's goals, may moderate the effects of power, encouraging more effective and group-oriented leadership even from proselves. How such factors alter the dynamics of leadership and collective action could be addressed in straightforward extensions of the theory and experiment presented above.

Study 2a and 2b suggest that group members tend to select prosocials to lead them. Importantly, however, participants in Study 2a had direct evidence of group members' prosociality. That is, it was relatively straightforward for them to distinguish who was prosocial (those who were sacrificing for the group) and select on that basis. Indeed, even in Study 2b, competition for leadership did not prevent participants from being able to identify and select their more prosocial group members to leadership. But such direct evidence of group members' prosociality is not always available. And, whether evidence of others' prosociality is available or not, leaders of real-world groups might be selected based on other qualities which may—or may not—be related to prosociality, including willingness to punish low contributors (Van Vugt and De Cremer 1999), informal status (Willer 2009), social capital (Glowacki and von Rueden 2011), greater access to resources (Przepiorka and Diekmann 2013), or gender (Van Vugt and Spisak 2008). It is therefore possible that self-regarding types will at least sometimes end up as leaders. More work must consider how people select their leaders under situations where a variety of information about potential leaders—including some “noise” that may or may not be related to prosociality—is available.

Further, more work must consider, if proselves *are* chosen for leadership, what mechanisms may encourage them to act in group-serving ways. A variety of factors



might encourage proself leaders to act in ways that enhance group goals (or, for that matter, lead prosocial leaders to behave more self-interestedly). Indeed, in real-world groups, a mix of prosocial and proself motives may be operative, such that both types of leaders would behave more similarly than they did in Study 1. For instance, prior work finds that proselfs are nearly as cooperative as prosocials when cooperation results in reputational benefits (Simpson and Willer 2008). Future work should consider how proselfs' greater reputational concerns and, as discussed above, the prospect of losing leadership positions, might attenuate the person x situation effects observed above.

### *Conclusion*

I have asked whether a leader, given sole responsibility for monitoring and sanctioning others, can maintain cooperation in groups facing collective action problems. On the one hand, leaders prevent some of the key problems associated with peer sanctions, such as over-punishment and earnings losses. Yet some research on power suggests that the enhanced power that comes along with leadership positions might lead to self-serving actions, rather than group-oriented behavior. From this perspective, it might seem puzzling that leadership is so ubiquitous in real-world collective action groups.

But I proposed and found evidence for a person x situation approach to leadership that helps shed light on this puzzle. Social value orientation moderates the impact of leadership on both leaders' behaviors and group-level outcomes. Other-regarding (prosocial) leaders contribute more to the group after ascending to leadership, while self-regarding (proself) leaders decrease their contributions even further. Further, I found that

leaders' behaviors have large downstream effects on contributions from the rank and file. As a result, prosocial-led groups as a whole were substantially more productive than proself-led groups, pointing to the importance of putting power and influence in the right hands.

Given that prosocials are more effective leaders than proselfs—and lead their groups to produce larger public goods—do people actually tend to select prosocials to lead them? If groups are unable or unwilling to select prosocials for leadership, then the leader “solution” may not be a solution at all. However, I found that if people are given the choice to select a group leader, they typically prefer their more other-regarding group members. In addition, selecting leaders via an election process, rather than appointment, enhances public good production both before *and* after the elected leader is installed. That is, these results suggest that allowing group members to choose their leaders will result in an efficient and effective sanctioning system.

Finally, I also found that prosocial leaders were more effective in maintaining contributions than the standard peer sanctioning system, where the power to punish others is distributed equally among all group members. More surprisingly, although proself leaders decreased their contributions substantially after receiving leadership, and influenced the rank and file to contribute less as well, peer sanctioning groups did not outperform them in public good production. In fact, after accounting for the detrimental costs of punishments, peer sanctioning groups fared worse than groups with self-regarding leaders. These results cast doubt on the viability of standard solutions that rely on peer sanctions, suggesting instead that leadership structures promise more effective solutions to collective action problems.

## REFERENCES

- Ahlquist, John S. and Margaret Levi. 2011. "Leadership: What it Means, What it Does, and What We Want to Know About It." *Annual Review of Political Science* 14:1-24.
- Anderson, Cameron and Courtney E. Brown. 2010. "The Functions and Dysfunctions of Hierarchy." *Research in Organizational Behavior* 30:55-89.
- Andreoni, James. 2006. "Leadership Giving in Charitable Fund-raising." *Journal of Public Economic Theory* 8: 1-22.
- Axelrod, Robert. 1984. *The Evolution of Cooperation*. New York: Basic Books.
- Baldassarri, Delia and Guy Grossman. 2011. "Centralized Sanctioning and Legitimate Authority Promote Cooperation in Humans." *Proceedings of the National Academy of Sciences* 108:11023-11027.
- Balliet, Daniel, Craig Parks, and Jeff Joireman. 2009. "Social Value Orientations and Cooperation in Social Dilemmas: A Meta-Analysis." *Group Processes and Intergroup Relations* 12:533-547.
- Barclay, Pat. 2004. "Trustworthiness and Competitive Altruism Can Also Solve the 'Tragedy of the Commons.'" *Evolution and Human Behavior* 25:209-220.
- Barclay, Pat and Robb Willer. 2007. "Partner Choice Creates Competitive Altruism in Humans." *Proceedings of the Royal Society of London Series B* 274:749-753.
- Bass, Bernard M. and Ruth Bass. 2009. *The Bass Handbook of Leadership: Theory, Research, and Managerial Applications*. Free Press, New York.

- Benard, Stephen. 2012. "Cohesion from Conflict: Does Intergroup Conflict Motivate Intragroup Norm Enforcement and Support for Centralized Leadership?" *Social Psychology Quarterly* 75: 107-130.
- Berger, Joseph, M. Hamit Fisek, Robert Z. Norman, and Morris Zelditch, Jr. 1977. *Status Characteristics and Social Interaction: An Expectation States Approach*. New York: Elsevier.
- Blader, Steven L. and Ya-Ru Chen. 2012. "Differentiating the Effects of Status and Power: A Justice Perspective." *Journal of Personality and Social Psychology* 102:994-1014.
- Buhrmester, Michael, Tracy Kwang, and Samuel D. Gosling. 2011. "Amazon's Mechanical Turk: A New Source of Inexpensive, Yet High-Quality, Data?" *Perspectives on Psychological Science* 6: 3-5.
- Burke, Peter J., Jan E. Stets, and Christine Cerven. "Gender, Legitimation, and Identity Verification in Groups." *Social Psychology Quarterly* 70:27-42.
- Chen, S., A. Y. Lee-Chai, and J. A. Bargh. 2001. "Relationship Orientation as a Moderator of the Effects of Social Power." *Journal of Personality and Social Psychology* 80:173-187.
- Clark, C. Robert, Samuel Clark, and Mattias K. Polborn. 2006. "Coordination and Status Influence." *Rationality and Society* 18:367-391.
- Cook, Karen S. and Richard M. Emerson. 1978. "Power, Equity, and Commitment in Exchange Networks." *American Sociological Review* 43:721-739.
- Cook, Karen S., Russell Hardin, and Margaret Levi. 2005. *Cooperation Without Trust?* New York: Russell Sage Foundation.

- Dal Bó, Pedro, Andrew Foster, and Louis Putterman. 2010. "Institutions and Behavior: Experimental Evidence on the Effects of Democracy." *American Economic Review* 100: 2205-2229.
- DeCelles, Katherine A., Scott D. DeRue, Joshua D. Margolis, Tara L. Ceranic. 2012. "Does Power Corrupt or Enable? When and Why Power Facilitates Self-Interested Behavior." *Journal of Applied Psychology* 97:681-689.
- De Cremer, David and Eric van Dijk. 2008. "Leader-Follower Effects in Resource Dilemmas: The Roles of Leadership Selection and Social Responsibility." *Group Processes and Intergroup Relations* 3:355-369.
- Dreber, Anna, David G. Rand, Drew Fudenberg, and Martin A. Nowak. 2008. "Winners Don't Punish." *Nature* 452:348-351.
- Emerson, Richard M. 1962. "Power-Dependence Relations." *American Sociological Review* 27:31-41.
- Eriksson, Kimmo, Daniel Cownden, Micael Ehn, and Pontus Strimling. 2014. "Altruistic and Antisocial Punishers are One and the Same." *Review of Behavioral Economics* 1:209-221.
- Eriksson, Kimmo and Pontus Strimling. 2012. "The Hard Problem of Cooperation." *PLoS ONE* 7: e40325
- Eriksson, Kimmo, Pontus Strimling, and Michael Ehn. 2013. "Ubiquity and Efficiency of Restrictions on Informal Punishment Rights." *Journal of Evolutionary Psychology* 11:17-34.
- Fehr, E. and S. Gächter. 2002. "Altruistic Punishment in Humans." *Nature* 415:137-140.

- Fehr, E. and H. Gintis. 2007. "Human Motivation and Social Cooperation: Experimental and Analytical Foundations." *Annual Review of Sociology* 33:43-64.
- Fischbacher, Urs. 2007. "z-Tree: Zurich Toolbox for Ready-made Economic Experiments." *Experimental Economics* 10:171-178.
- Fiske, Susan T. 1993. "Controlling other people: The impact of power on stereotyping." *American Psychologist* 48:621-628.
- French, J. and B. H. Raven. 1959. "The Bases of Social Power." In D. Cartwright (Ed.), *Studies in Social Power*, pp.150-167. Ann Arbor, MI: Institute for Social Research.
- Galinsky, A. D., J. C. Magee, M. E. Inesi, and D. H. Gruenfeld. 2006. "Power and Perspectives Not Taken." *Psychological Science* 17:1068-1074.
- Grossman, Guy and Delia Baldassarri. 2012. "The Impact of Elections on Cooperation: Evidence from a Lab-in-the-Field Experiment in Uganda." *American Journal of Political Science* 56:964-985.
- Guala, Francesco. 2012. "Reciprocity: Weak or Strong? What Punishment Experiments Do (and Do Not) Demonstrate." *Behavioral and Brain Sciences* 35: 1-59.
- Hardin, R. 1982. *Collective Action*. Johns Hopkins University Press.
- Hardy, Charlie L. and Mark Van Vugt. 2006. "Nice Guys Finish First: The Competitive Altruism Hypothesis." *Personality and Social Psychology Bulletin* 32:1402-1413.
- Hechter, Michael and Karl Dieter-Opp. 2001. *Social Norms*. New York: Russell Sage Foundation.
- Heckathorn, Douglas. 1989. "Collective Action and the Second-Order Free-Rider Problem." *Rationality and Society* 1:78-100.

- Heckathorn, Douglas. 1990. "Collective Sanctions and Compliance Norms—A Formal Theory of Group-Mediated Social Control." *American Sociological Review* 55:366-384.
- Heckathorn, Douglas. 1996. "The Dynamics and Dilemmas of Collective Action." *American Sociological Review* 61:250-277.
- Helbing, Dirk, Wenjian Yu, Karl-Dieter Opp, and Heiko Rauhut. 2014. "Conditions for the Emergence of Shared Norms in Populations with Incompatible Preferences." *PLoS ONE* 9:e104207.
- Henrich, J., R. McElreath, A. Barr, J. Ensminger, C. Barrett, A. Bolyanatz, J. C. Cardenas, M. Gurven, E. Gwako, N. Henrich, C. Lesorogol, F. Marlowe, D. Tracer, and J. Ziker. 2006. "Costly Punishment Across Human Societies." *Science* 312: 1767-1770.
- Horne, Christine. 2009. *The Rewards of Punishment: A Relational Theory of Norm Enforcement*. Stanford University Press: Stanford, CA.
- Keltner, D., L. Capps, A. M. Kring, R. C. Young, and E. A. Heerey. 2001. "Just Teasing: A Conceptual Analysis and Empirical Review." *Psychological Bulletin* 127:229-248.
- Keltner, D., D. Gruenfeld, and C. Anderson. 2003. "Power, Approach, and Inhibition." *Psychological Review* 110:265-284.
- Keltner, D. and R. J. Robinson. 1996. "Extremism, power, and the imagined basis of social conflict." *Current Directions in Psychological Science* 5: 101–105.
- Keltner, D. and R. J. Robinson. 1997. "Defending the status quo: Power and bias in social conflict." *Personality and Social Psychology Bulletin* 23:1066–1077.

- Kipnis, D. 1972. "Does Power Corrupt?" *Journal of Personality and Social Psychology* 24:33-41.
- Kipnis, D. 1976. *The Powerholders*. Chicago, IL: University of Chicago Press.
- Kitts, James. 2006. "Collective Action, Rival Incentives, and the Emergence of Antisocial Norms." *American Sociological Review* 71:235-259.
- Kollock, P. 1998. "Social Dilemmas: The Anatomy of Cooperation." *Annual Review of Sociology* 24:183-214.
- Kosfeld, M. and Rustagi, D. 2015. "Leader Punishment and Cooperation in Groups: Experimental Field Evidence from Commons Management in Ethiopia." *American Economic Review* 105:747-783.
- Lawler, Edward J. and Samuel B. Bacharach. 1987. "Comparison of Dependence and Punitive Forms of Power." *Social Forces* 66:446-462.
- Lovaglia, Michael J., Jeffrey W. Lucas, Christabel L. Rogalin and Abigail Darwin. 2006. "Power, Status, and Leadership in Diverse Organizations: From Basic Research to Program Development." *Advances in Group Processes* 23:183-206.
- Lucas, Jeffrey W. 2003. "Status Processes and the Institutionalization of Women as Leaders." *American Sociological Review* 68:464-480.
- Macy, Michael. 1993. "Backward-Looking Social Control." *American Sociological Review* 58:819-837.
- Markovsky, Barry, David Willer, and Travis Patton. 1987. "Power Relations in Exchange Networks." *American Sociological Review* 53:220-236.



- Marwell, Gerald and Ruth E. Ames. 1979. "Experiments on the Provision of Public Goods I: Resources, Interest, Group Size, and the Free-Rider Problem." *American Journal of Sociology* 84:1335-1360.
- Marwell, Gerald and Ruth E. Ames. 1980. "Experiments on the Provision of Public Goods II: Provision Points, Stakes, Experience, and the Free-Rider Problem." *American Journal of Sociology* 85:926-937.
- Marwell, Gerald and Pamela Oliver. 1993. *The Critical Mass in Collective Action: A Microsocial Theory*. Cambridge, England: Cambridge University Press.
- Milinski, M., D. Semmann, and H. Krambeck. 2002. "Donors to Charity Gain in Both Indirect Reciprocity and Political Reputation." *Proceedings of the Royal Society B* 269:881-883.
- Molm, Linda D. 1988. "The Structure and Use of Power: A Comparison of Reward and Punishment Power." *Social Psychology Quarterly* 51:108-122.
- Molm, Linda D. 1989. "Punishment Power: A Balancing Process in Power-Dependence Relations." *American Journal of Sociology* 94:1392-1418.
- Molm, Linda D. 1990. "Structure, Action, and Outcomes: The Dynamics of Power in Social Exchange." *American Sociological Review* 55:427-447.
- Molm, Linda. 2010. "The Structure of Reciprocity." *Social Psychology Quarterly* 73:119-131.
- Moxnes, E. and E. Van der Heijden. 2003. "The Effect of Leadership in a Public Bad Experiment." *Journal of Conflict Resolution* 47: 773-795.

- Nee, Victor and Paul Ingram. 1998. "Embeddedness and Beyond: Institutions, Exchange, and Social Structure." in *The New Institutionalism in Sociology*. Mary C. Brinton and Victor Nee (eds). Russell Sage Foundation, New York.
- Nikiforakis, Nikos. 2008. "Punishment and Counter-Punishment in Public Good Games: Can we Really Govern Ourselves?" *Journal of Public Economics* 92:91-112.
- Nosenzo, Daniele and Martin Sefton. 2014. "Promoting Cooperation: The Distribution of Reward and Punishment Power." in *Social Dilemmas: New Perspectives on Reward and Punishment*. P.A.M. Van Lange, B. Rockenbach, and T. Yamagishi (eds). Oxford University Press.
- Nowak, Martin A. and Karl Sigmund. 2005. "Evolution of Indirect Reciprocity." *Nature* 437:1291-1298.
- O'Gorman, Rick, Joseph Henrich, and Mark Van Vugt. 2009. "Constraining Free Riding in Public Goods Games: Designated Solitary Punishers Can Sustain Human Cooperation." *Proceedings of the Royal Society B* 276:323-329.
- Oliver, Pamela. 1980. "Rewards and Punishments as Selective Incentives for Collective Action: Theoretical Investigations." *American Journal of Sociology* 85:1356-1375.
- Olson, Mancur 1965. *The Logic of Collective Action: Public Goods and the Theory of Groups*. Cambridge, MA: Harvard Univ. Press.
- Ones, Umut and Louis Putterman. 2007. "The Ecology of Collective Action: A Public Goods and Sanctions Experiment with Controlled Group Formation." *Journal of Economic Behavior and Organization* 62: 495-521.

- Ostrom, Elinor. 2000. "Collective Action and the Evolution of Social Norms." *Journal of Economic Perspectives* 14:137-158.
- Ostrom, Elinor, J. Walker, and R. Gardner. 1992. "Covenants with and without a sword: Self-Governance is Possible." *American Political Science Review* 86:404-417.
- Potters, Jan, Martin Sefton, and Lise Versterlund. 2007. "Leading-by-example and Signaling in Voluntary Contribution Games: an Experimental Study." *Economic Theory* 33: 169-182.
- Price, Michael E. and Mark Van Vugt. 2014. "The Evolution of Leader-Follower Reciprocity: The Theory of Service-for-Prestige." *Frontiers in Human Neuroscience* 8:363.
- Przepiorka, Wojtek and Andreas Diekmann. 2013. "Individual Heterogeneity and Costly Punishment: A Volunteer's Dilemma." *Proceedings of the Royal Society B*: 280.20130247.
- Ridgeway, Cecilia L. 2001. "Inequality, Status, and the Construction of Status Beliefs." In J. Turner (ed.), *Handbook of Sociological Theory*, pp. 323-342. New York: Kluwer/Plenum.
- Samuelson, C. D., D. M. Messick, C. Rutte, and H. Wilke. 1984. "Individual and Structural Solutions to Resource Dilemmas in Two Cultures." *Journal of Personality and Social Psychology* 47: 94-104.
- Sell, Jane. 1997. "Gender, Strategies, and Contributions to Public Goods." *Social Psychology Quarterly* 60:252-265.

- Sell, Jane, Michael J. Lovaglia, Elizabeth A. Mannix, Charles D. Samuelson, and Rick K. Wilson. 2004. "Investigating Conflict, Power, and Status Within and Among Groups." *Small Group Research* 35:44-72.
- Sell, Jane and Rick K. Wilson. 1991. "Levels of information and contributions to public goods." *Social Forces* 70:107-124.
- Sell, Jane and Rick K. Wilson. 1999. "The Maintenance of Cooperation: Expectations of Future Interaction and the Trigger of Group Punishment." *Social Forces* 77:1551-1571.
- Shinada, Mizuho and Toshio Yamagishi. 2007. "Bringing Back Leviathan into Social Dilemmas." Pp. 93-123 in *New Issues and Paradigms in Social Dilemma Research*, edited by Anders Biel, et al. New York: Springer.
- Simpson, Brent, Barry Markovsky, and Mike Steketee. 2011. "Network Knowledge and the Use of Power." *Social Networks* 33:172-176.
- Simpson, Brent and Robb Willer. 2008. "Altruism and Indirect Reciprocity: The Interaction of Person and Situation in Prosocial Behavior." *Social Psychology Quarterly* 71:37-52.
- Simpson, Brent, Robb Willer, and Cecelia Ridgeway. 2012. "Status Hierarchies and the Organization of Collective Action." *Sociological Theory* 30:149-166.
- Snijders, Tom A. B. and Roel J. Bosker. 2012. *Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling*. 2<sup>nd</sup> ed. London: Sage Publishers.
- Strimling, Pontus and Kimmo Eriksson. 2014. "Regulating the Regulation: Norms about Punishment." In Van Lange P, Rockenbach B, & Yamagishi T (eds.) *Reward and Punishment in Social Dilemmas*, Oxford University Press, pp. 52-69.

- Thye, Shane R. 2000. "A Status Value Theory of Power in Exchange Networks." *American Sociological Review* 65:407-432.
- Thye, Shane R., David Willer, and Barry Markovsky. 2006. "From Status to Power: New Models at the Intersection of Two Theories." *Social Forces* 84:1471-1495.
- Van Lange, Paul A. M. 1999. "The Pursuit of Joint Outcomes and Equality in Outcomes: An Integrative Model of Social Value Orientation." *Journal of Personality and Social Psychology* 77: 337-349.
- Van Lange, Paul A. M., Rene Bekkers, Theo N. M. Schuyt, and Mark Van Vugt. 2007. "From Games to Giving: Social Value Orientation Predicts Donations to Noble Causes." *Basic and Applied Social Psychology* 29:375-384.
- Van Lange, Paul A. M., W. Otten, E. M. N. DeBruin, and J. A. Joireman. 1997. "Development of Prosocial, Individualistic, and Competitive Orientations: Theory and Preliminary Evidence." *Journal of Personality and Social Psychology* 73:733-746.
- Van Vugt, Mark. 2006. "Evolutionary Origins of Leadership and Followership." *Personality and Social Psychology Review*: 354-371.
- Van Vugt, Mark and David De Cremer. 1999. "Leadership in Social Dilemmas: The Effects of Group Identification on Collective Actions to Provide Public Goods." *Journal of Personality and Social Psychology* 76:587-599.
- Van Vugt, Mark and Brian R. Spisak. 2008. "Sex Differences in the Emergence of Leadership During Competitions Within and Between Groups." *Psychological Science* 19: 854-858.

- Van Vugt, Mark, R. Hogan, and R. B. Kaiser. 2008. "Leadership, Followership, and Evolution: Some Lessons From the Past." *American Psychologist* 63: 182-196.
- Walker, H. and M. Zelditch. 1993. "Power, Legitimacy, and the Stability of Authority: A Theoretical Research Program." Pp 364-381 in *Theoretical Research Programs: Studies in the Growth of Theory*, edited by J. Berger and M. Zelditch. Stanford University Press.
- Weber, Max. 1947. *The Theory of Social and Economic Organization*. A. H. Henderson and T. Parsons, Eds. & Trans. New York: Free Press.
- Weber, Max. 1978. *Economy and Society*. E. Fischoff, Trans. G. Roth and C. Wittich, Eds. Berkeley, CA: University of California Press.
- Weinberg, Jill, Jeremy Freese, and David McElhattan. 2014. "Comparing Data Characteristics and Results of an Online Factorial Survey between a Population-Based and a Crowdsource-Recruited Sample." *Sociological Science* 1:292-310.
- Willer, David, Michael J. Lovaglia, and Barry Markovsky. 1997. "Power and Influence: A Theoretical Bridge." *Social Forces* 76:571-603.
- Willer, Robb. 2009. "Groups Reward Individual Sacrifice: The Status Solution to the Collective Action Problem." *American Sociological Review* 74:23-43.
- Willer, Robb, Lisa Troyer, and Michael J. Lovaglia. 2005. "Influence Over Observers of Structural Power: An Experimental Investigation." *The Sociological Quarterly* 46:263-277.
- Yamagishi, Toshio. 1986. "The Provision of a Sanctioning System as a Public Good." *Journal of Personality and Social Psychology* 51:110-116.

Yamagishi, Toshio. 1992. "Group Size and the Provision of a Sanctioning System in a Social Dilemma." In *A Social Psychological Approach to Social Dilemmas*, edited by W. B. G. Liebrand, D. M. Messick, and H. A. M. Wilke. New York: Pergamon Press.

Yamagishi, Toshio, Nobuhiro Mifune, Yang Li, Mizuho Shinada, Hirofumi Hashimoto, Yutaka Horita, Arisa Miura, Keigo Inukai, Shigehito Tanida, Toko Kiyonari, Haruto Takagishi, and Dora Simunovic. 2013. "Is Behavioral Pro-sociality Game-specific? Pro-social Preference and Expectations of Pro-sociality." *Organizational Behavior and Human Decision Processes* 120:260-271.

## APPENDIX A: SVO MEASURE, STUDY 1 AND STUDY 2B

In order to classify participants' social value orientations (SVOs) in Studies 1 and 2b, and to assign leaders in Study 1, I used the most common measure of SVO—the nine-item triple dominance measure (Van Lange, Otten, DeBruin, and Joireman 1997).

Participants in Study 1 and 2b completed this measure, reproduced below (see also Van Lange et al. 1997), before beginning the public good task instructions:

In this set of questions, we ask you to imagine that you have been randomly paired with another person, whom we will refer to simply as the “other.” Other is someone you do not know and that you will not knowingly meet in the future. Both you and Other will be making choices by circling either the letter A, B, or C. Your own choices will produce points for yourself and Other. Likewise, Other's choice will produce points for him/her and for you. Every point has value: The more points you receive, the better for you, and the more points Other receives, the better for him/her.

Here's an example of how this task works.

	A	B	C
You Get	500	500	550
Other Gets	100	500	300



In this example, if you chose A you would receive 500 points and Other would receive 100 points; if you chose B, you would receive 500 points and Other 500; and if you chose C, you would receive 550 points and Other 300. So, you see that your choice influences both the number of points you receive and the number of points the other receives.

Before you begin making choices, keep in mind that there are no right or wrong answers – choose the option that you, for whatever reason, prefer most. Also, remember that the points have value: The more of them you accumulate, the better for you. Likewise, from the Other’s point of view, the more points s/he accumulates, the better for him/her.

For each of the three choice situations below, choose A, B or C, depending on which column you prefer most.

1

	A	B	C
You Get	480	540	480
Other Gets	80	280	480

2

	A	B	C
You Get	560	500	500
Other Gets	300	500	100

3

	A	B	C
You Get	520	520	580
Other Gets	520	120	320

4

	A	B	C
You Get	500	560	490

Other Gets	100	300	490
------------	-----	-----	-----

5

	A	B	C
You Get	560	500	490
Other Gets	300	500	90

6

	A	B	C
You Get	500	500	570
Other Gets	500	100	300

7

	A	B	C
You Get	510	560	520
Other Gets	510	300	110

8

	A	B	C
You Get	550	500	500
Other Gets	300	100	500

9

	A	B	C
You Get	480	490	540
Other Gets	100	490	300

As is standard in the social value orientation literature (Simpson and Willer 2008; Van Lange et al. 1997), a participant was classified as a given social value type (*prosocial, individualistic, or competitive*) when he or she answered at least six of nine items consistently and were “unclassified” if not. Prosocial choices are: 1c 2b 3a 4c 5b 6a 7a 8c 9b; individualistic choices are: 1b 2a 3c 4b 5a 6c 7b 8a 9c; and competitive choices are: 1a 2c 3b 4a 5c 6b 7c 8b 9a (Van Lange et al. 1997). Following past work, I combined individualists and competitors into one “proself” category (Simpson and Willer 2008; Van Lange et al. 1997).

## APPENDIX B: INSTRUCTIONS, STUDY 1

Reproduced here are the materials for Study 1, including 1) the instructions, examples, and quiz questions for the non-punishment rounds of the public good dilemma (described to participants as an “investment decision task”) and 2) the instructions, examples, and quiz questions for the punishment rounds, which participants viewed after the nine rounds of the non-punishment phase were completed. All study materials were programmed and presented using z-Tree version 3.3.11 (Fischbacher 2007).

### *[Instructions, Non-punishment rounds]*

In today's study, you will take part in a series of "investment decisions." Your earnings in today's study will partly depend on how well you understand the instructions. So, please read all instructions and examples carefully.

The basic directions are as follows: you will be completing this study in a group of four. Your group members are three other participants currently in the lab. You will not meet your group members at any time, nor will you learn any identifying information about your group members. Likewise, your group members will not learn any identifying information about you.

The study consists of several periods (or rounds). At the start of each round, you (and each of your other group members) get 20 points. You can contribute

anywhere from 0 to 20 of these points to a "group fund." Any points you do not contribute to the group fund remain in your personal fund, for you to keep.

Anything that is contributed to the group fund will be doubled. Then, the doubled amount will be divided between all four of the members of your group, whether or not they contributed to the group fund. (Similarly, other group members' contributions to the group fund will be doubled and redistributed equally among all members of the group.) Your total earnings per round are your share of the earnings from the group fund, plus whatever you did not invest. The same goes for other group members.

Points earned over each round will be translated into dollars at the end of the study. **So, the more points you earn, the more money you will receive.**

Make sure you have carefully read and understand the instructions. Next, you will read over a few examples. If you have any questions about the instructions, you may slightly open your door and a research assistant will be with you in a moment.

If you understand the instructions, click "Continue."

*[Example screens, Non-punishment rounds]*

*[Screen 1]* Let's go over an example.

Imagine that each group member invests all 20 of their points. Thus, there are now 80 points in the group fund (20 points from each of 4 group members).

Since anything in the group fund is doubled, the 80 points become 160 points. Then, the group fund is divided by four (for each of the four group members). So, each group member receives 40 points.

Since everyone invested all of their points to the group fund, each player finishes the period with 40 points (40 earned from the group fund + 0 kept in the personal fund).

If you have any questions about this example, please slightly open your door and a research assistant will be with you in a moment.

If you understand the example, click "Continue."

*[Screen 2]* Here's another example.

Imagine that each group member invests none of their 20 points. Now, there are zero points in the group fund.

Since there are no points in the group fund to double, and no points to divide by all the group members, everyone earns 20 points from this round: 0 points from the group fund + 20 points kept in the personal fund.

If you have any questions, please slightly open your door and a research assistant will be with you in a moment.

If you understand the example, click "Continue."

*[Screen 3]* Imagine that three group members invest all 20 of their points, and one group member invests none of their points. Thus, there are 60 points in the group fund (20 points x 3 group members).

The 60 points in the group fund gets doubled to 120 points. Then, the 120 points gets divided into four so that every group member gets 30 points.

The three group members who invested their points end the round with 30 points each (30 points from the group fund + 0 points kept in the personal fund). The group member who invested nothing ends the round with 50 points (30 earned from the group fund + 20 kept in the personal fund).

If you have any questions, please slightly open your door and a research assistant will be with you in a moment.

If you understand the example, click "Continue."

*[Next participants completed several open-ended quiz questions. Each question was displayed on a separate screen; the first two paragraphs of the material below were presented before each question. After answering a given question, participants were informed whether or not their answer was correct. Incorrect answers were followed by a detailed explanation of the correct answer. Correct answers in **bold**]*

Next, to make sure you understand everything, you'll complete a few practice questions.

Imagine that your three group members each contribute 8 of their 20 points to the group fund. You contribute 0 points to the group fund. Please answer the following question about this scenario:

What is the total number of points in the group fund (BEFORE it is doubled)?

**24**

What is the total number of points in the group fund (AFTER it is doubled)?

**48**

How much does each group member receive from the GROUP fund (that is, after the amount in the group fund is divided amongst each of the four group members)?

**12**

How many points do your fellow group members each receive TOTAL for this round (that is, the amount earned from the group fund plus any kept in their private fund)?

**24**

How many points do YOU receive TOTAL for this round (that is, the amount earned from the group fund plus any kept in your private fund)?

**32**

*[Instructions, peer punishment condition, punishment rounds; Instructions for the leader conditions in brackets]*

For the next part of the study, you will continue to make investment decisions. However, some of the rules will change. Therefore, please carefully read the instructions so that you understand how the rules will change.

Under the new rules, every group member [one group member (chosen at random)] will have the opportunity to deduct points from their fellow group members following each round.



Specifically, after each round (that is, after all members' contributions for that round are known), everyone [one person] will be able to deduct points from their group members' earnings, if they choose to do so. It costs 1 point to deduct 3 points from another group member. For example, if a [the] group member chooses to spend two of his/her points to deduct points from Participant X, Participant X's earnings will be reduced by six points.

Group members can spend anywhere from 0 to 10 of their own points to deduct points from each of their group members. That is, they may choose not to deduct from a group member's earnings. Or, they may choose to spend some of their own points (up to 10) to deduct points (up to 30) from another group member.

[A few other things: on a later screen, we will let you and your group members know who has been randomly assigned to the role described above. Whoever is randomly selected will continue in this role for the remainder of the study.]

Please make sure that you have read the instructions carefully and understand the new rules. If you have any questions, please slightly open your door and a research assistant will be with you in a moment. Otherwise, click "Continue."

*[Quiz questions, peer punishment condition, changes for the leader conditions in brackets, correct answer in **bold**]*

Next, you'll complete a few practice questions.

[Imagine you have been randomly assigned to make deductions.] Imagine that you choose to spend 9, 5, and 0 points to deduct from your three group members' earnings. What is the total cost to you?

**14**

[Imagine that you have been randomly assigned to make deductions.] Imagine that you spent 3 points to deduct from one of your group members' earnings. By how many points will that group member's earnings be reduced?

**9**

[Imagine that you were not randomly assigned to make deductions.] Imagine that one of your group members [The person who was assigned to make deductions] spent 4 points to deduct from your earnings. By how many points will your earnings be reduced?

**12**

## APPENDIX C: INSTRUCTIONS, STUDY 2A

Reproduced here are the instructions for Study 2a. These include a description of the non-punishment rounds of the public good dilemma (described to participants as a “group decision making task”) and, second, the instructions for the punishment phase of the study, where a single group member (the “deductor”) could make deductions from the others’ earnings. The entire study was presented to Amazon’s Mechanical Turk users (see Buhrmester, Kwang, and Gosling 2011; Weinberg, Freese, and McElhattan 2014) using Qualtrics Online Survey Software.

### *[Instructions, Non-punishment rounds]*

Imagine that you were a group member participating in a task called the Group Decision Making Task. The task consists of six rounds. At the start of each round, each group member gets 20 dollars. Each can contribute anywhere from 0 to 20 dollars to a "group fund."

Any amount a group member does not contribute to the group fund remains in his/her “personal fund”, for that group member to keep.

Any money that is contributed to the “group fund” is doubled. Then, the doubled amount will be divided between all members of the group, whether or not they contributed to the group fund.

Each group members' total earnings per round are their share of the earnings from the group fund, plus whatever they kept in the personal fund.

Make sure you have carefully read and understand the instructions. Next, we will go over a few examples, then, you'll answer a few quiz questions to ensure your understanding of the task.

*[Examples, Non-punishment rounds]*

Let's go over some examples for a group of four.

Imagine each group member invests all 20 of their dollars. There are now 80 dollars in the group fund ( $\$20 \times 4$  group members).

Since anything in the group fund is doubled, the \$80 becomes \$160. Then, the group fund is divided by four (for each of the four group members). So, each group member receives \$40. Each player finishes the period with \$40 (\$40 from the group fund + \$0 kept in the personal fund)

Now, imagine that each group member invests none of their money (\$0 in the group fund).

Since there is nothing in the group fund to double and divide by all the group members, everyone earns \$20: \$0 from the group fund + \$20 kept in the personal fund.

Make sure you have carefully read and understand the examples. Later, you'll answer a few quiz questions to ensure your understanding of the task.

Finally, imagine that three group members invest all 20 of their dollars, and one group member invests nothing. Thus, there are 60 dollars in the group fund ( $\$20 \times 3$  group members).

The \$60 in the group fund gets doubled to \$120. Then, the \$120 gets divided by four so that every group member gets \$30.

The three group members who invested their points end the round with \$30 each (\$30 from the group fund + \$0 kept in the personal fund). The group member who invested nothing ends the round with \$50 (\$30 from the group fund + \$20 kept in the personal fund).

Make sure you have carefully read and understand the examples before clicking Continue. Next, you'll answer a few quiz questions to ensure your understanding of the task.

*[Quiz questions, Non-punishment rounds, correct answers in **bold**]*

Each group member can contribute anywhere from 0 to 50 dollars to the group fund.

True

**False**

Money contributed to the group fund is tripled.

True

**False**

Money in the group fund is divided evenly amongst all group members, regardless of whether or not they contributed to the group fund.

**True**

False

*[Participants in the Mechanical Turk study were then shown the contributions of each group member, per round, for one of two randomly selected groups from Study 1 (the order in which these two groups were displayed was counterbalanced)]:*

Now please imagine that the group members participating in this task contributed as follows.

Round:	1	2	3	4	5	6
Participant 1 contributed:	\$20	\$10	\$10	\$15	\$16	\$15
Participant 2 contributed:	\$5	\$20	\$20	\$20	\$5	\$1
Participant 3 contributed:	\$0	\$0	\$0	\$0	\$0	\$0
Participant 4 contributed:	\$10	\$8	\$5	\$8	\$2	\$8

Remember, group members can contribute from \$0 to \$20 per round. Briefly look over the participants' contributions, to get acquainted with them. We'll ask you some questions about the table later, but you don't need to memorize anything.

*[Instructions, punishment rounds]:*

Now imagine that the group members were told the task would change: after each round (that is, after all members' contributions for that round are known), one group member will have the opportunity to deduct money from his or her fellow group members' earnings after each round, if he/she chooses to do so. (We will call this person the Deductor.)

It costs \$1 to deduct \$3 from another group member. For example, if the Deductor chooses to spend \$2 to reduce Participant X's earnings, Participant X's earnings will be reduced by \$6.

The Deductor may choose not to deduct from a group member's earnings. Or, the Deductor may choose to spend some of his/her own money (up to \$10) to deduct points (up to \$30) from another group member.

Whoever is selected to deduct points will continue in this role for several additional rounds.

Make sure that you have read the instructions carefully and understand the new rules.

*[Finally participants were asked to choose one group member for the Deductor role:]*

Recall that the group members participating in this task contributed as follows over six rounds. Note that the group members did not know that the rules would change until after they made the decisions given below.

Round:	1	2	3	4	5	6
Participant 1 contributed:	\$20	\$10	\$10	\$15	\$16	\$15
Participant 2 contributed:	\$5	\$20	\$20	\$20	\$5	\$1
Participant 3 contributed:	\$0	\$0	\$0	\$0	\$0	\$0
Participant 4 contributed:	\$10	\$8	\$5	\$8	\$2	\$8



Imagine that you were a member of the group, and that you could select a participant to the role of Deductor for the next phase of the study. Which participant would you choose to be the Deductor?

-Participant 1

-Participant 2

-Participant 3

-Participant 4

*[Finally, participants were asked several questions about the Deductor they chose. The same contribution information presented earlier was also displayed to participants while they completed these questions]*

You chose Participant [Choice] to be the Deductor. Please explain, in a few sentences, why you chose Participant [Choice] for the Deductor role. *[open-ended]*

The participant that I chose to be the Deductor would be a fair Deductor.

-Strongly Agree

-Agree

-Neither Agree nor Disagree

-Disagree

-Strongly Disagree

The participant that I chose to be the Deductor would be a moral Deductor.

- Strongly Agree
- Agree
- Neither Agree nor Disagree
- Disagree
- Strongly Disagree

The participant that I chose to be the Deductor would be a rational Deductor.

- Strongly Agree
- Agree
- Neither Agree nor Disagree
- Disagree
- Strongly Disagree

I predict that the participant that I chose to be the Deductor would make:

- 1 (No deductions at all)
- 2
- 3
- 4
- 5
- 6
- 7 (Many deductions)

*[Attention check questions. The same contributions information presented earlier was displayed to participants while they completed these questions:]*

Which participant contributed the most (across all six rounds)?

- Participant 1
- Participant 2
- Participant 3
- Participant 4

Which participant contributed the least (across all six rounds)?

- Participant 1
- Participant 2
- Participant 3
- Participant 4

Now we want you to answer some questions about a different group.

*[Information for the second group is shown below. Participants completed the same set of questions given above for the second group.]*

Round:	1	2	3	4	5	6
Participant 1 contributed:	\$10	\$10	\$10	\$10	\$11	\$12
Participant 2 contributed:	\$6	\$8	\$10	\$12	\$11	\$0
Participant 3 contributed:	\$20	\$20	\$15	\$15	\$20	\$15
Participant 4 contributed:	\$0	\$4	\$3	\$7	\$0	\$4

## APPENDIX D: INSTRUCTIONS, STUDY 2B

Reproduced here are the materials for the punishment rounds (referred to as “Part 2”) of Study 2b. The instructions, examples, and quiz questions for the non-punishment rounds (“Part 1”) were presented before the material below; see Appendix B for these materials as they were virtually identical to those used in Study 1. Study 2b was programmed and presented using z-Tree version 3.3.11 (Fischbacher 2007).

### *[Instructions, punishment rounds]*

Before we begin Part 1, to save time, we will go ahead and tell you what Part 2 of the study will be like. Again, your earnings today will partly depend on how well you understand the instructions. So, please read all instructions carefully.

In Part 2, you will continue to decide how many points, from 0 to 20, you wish to contribute to the group fund, and how many to keep for yourself. **However, in Part 2, one group "leader" will have the opportunity to deduct points from their fellow group members' earnings following each round.**

Specifically, after each round-- that is, after all members' contributions for that round are known-- the leader will be able to deduct points from their group members' earnings, if he/she chooses. It costs 1 point to deduct 3 points from another group member. So, for example, if the leader chooses to spend 2 of

his/her points to deduct earnings from Participant X, Participant X's earnings will be reduced by 6 points.

Leaders can spend anywhere from 0 to 10 of their own points to deduct points from each of their other group members. That is, they may choose not to spend any points. Or, they may choose to spend up to 10 of their own points, to deduct up to 30 points, from each of their group members.

**No one can deduct points from the leader, and whoever receives the leader position will have that position for the remainder of the study.**

Make sure you have carefully read and understand the instructions for Part 2. Then, click "Continue".

*[Example screens, punishment rounds]*

*[Screen 1]* To explain to you a bit more about the leader role for Part 2, consider this example.

Imagine you contributed 5 points to the group fund. Further imagine that Participant X contributed 0 points to the group, Y contributed 5 points, and Z contributed 10 points. (As a result, you would earn 25 points for that round. X would earn 30 points, Y earns 25 points, and Z earns 20 points.)

If **you were the leader** for Part 2, **you** could then decide how much to deduct from each group member. For instance, you could reduce X's earnings by 0 points, 30 points, or anywhere in between, if you chose. You could also reduce Y and Z's earnings, if you choose. **You cannot deduct earnings from yourself, nor can any other group member make deductions from you, if you are the leader.** You would be able to do this for all of Part 2, which is the last Part of the study.

*[Screen 2]* Now imagine instead that **Participant Z was the leader** for Part 2. Z could decide how much to deduct from each group member. For instance, he/she could reduce X's earnings by 0 points, 30 points, or anywhere in between. He/she could also reduce your earnings, if they chose. **Z could not deduct earnings from him or herself, nor can you, or any other group member, make deductions from the leader, Z.** Z would be the leader for all of Part 2, which is the last Part of the study.

*[Instructions, appointed leader condition, elected leader condition in brackets]*

The leader for Part 2 will be **randomly selected by the computer [elected by the group] at the end of Part 1.**

[Specifically, after Part 1, **you (and each of the others) will see each group members' participant numbers and their total contributions to the group fund in Part 1.** Each group member will then be able to vote for one person to be

the leader (you cannot vote for yourself). The winner of the vote will have the leader role.]

Finally, after the leader is randomly selected [elected], participant numbers will change. Therefore, [**although your behavior in Part 1 may determine whether you are elected to be the leader for Part 2,**] at the beginning of Part 2 no one (including the leader) will be able to identify you based on your behavior in Part 1.

Next, you will complete a few quiz questions to ensure your understanding of the instructions. Click Continue.

*[Manipulation check question, correct answer varied by condition]*

How will the leader for Part 2 be chosen? The leader will be:

- Elected by the group
- Randomly chosen

*[Instructions summary, appointed leader condition, elected leader condition in brackets]*

To sum up the task today, you will first complete a series of Part 1 rounds, where no one will make deductions.

The computer will then randomly select [You and the others will then elect] a group leader for Part 2.

Participant IDs will change [after the leader is elected, but] before Part 2 begins.

Then in Part 2, the leader can deduct points from others.

Click Continue when you are ready to begin.