

INTERGROUP CONFLICT AND INTRAGROUP DYNAMICS: HOW CONFLICT
CREATES NORMS AND HIERARCHIES

A Dissertation

Presented to the Faculty of the Graduate School
of Cornell University

In Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy

by

Stephen William Benard

August 2008

© 2008 Stephen William Benard

INTERGROUP CONFLICT AND INTRAGROUP DYNAMICS: HOW CONFLICT
CREATES NORMS AND HIERARCHIES

Stephen William Benard, Ph. D.

Cornell University 2008

Abstract: In this dissertation, I examine a fundamental sociological question: how does conflict between groups affect relations within groups? I present theoretical arguments specifying the conditions under which groups in conflict will develop stricter norms and more centralized systems of leadership than groups that do not experience conflict. I evaluate the predictions using two laboratory studies, in which small groups interact under varying levels of conflict. The results of the studies show that conflict influences the emergence of social norms, group cohesion, and the creation of hierarchies. The project has broader theoretical implications for the study of institutions, social control, and collective action, as well as practical relevance for developing conflict resolution strategies.

BIOGRAPHICAL SKETCH

Steve Benard came to Cornell University in 2001, after receiving a BA from the University of Rochester. He also received his M.A. in Sociology from Cornell in 2004. His research and teaching interests include social psychology, group conflict, collective action, social networks, gender, and inequality. In the fall, Steve will join the Sociology Department at Indiana University as an Assistant Professor.

To my parents, Michael and Christine Benard.

ACKNOWLEDGMENTS

The support of many people made completing this dissertation possible, both directly and indirectly. While it is impossible to adequately thank everyone who has contributed in some way, in this section I recognize some people who have been especially influential.

I thank my parents, Michael and Christine Benard, and my brothers, Mike and Nick for their immense support and encouragement throughout the years.

I have learned much from my teachers throughout my life, but there are two who were especially influential in shaping my career as a sociologist: Daniel Olson, from whom I took my first sociology course in the 12th grade (which included my first attempt to collect and analyze survey data...probably the less said about this, the better), and James Strauss, who did much to improve my writing and thinking skills.

I thank my undergraduate advisor, Tom Smith at the University of Rochester, who, with great patience, taught me much about theory and research, as well as the discipline as a whole. His support prepared me to study sociology at the graduate level, and was influential in shaping my research interests. I also benefited greatly from the guidance (and sense of humor) of Kamran Asdar Ali. Other faculty at Rochester who were more than gracious with their time and advice include Anthony Carter, Tom DiPiero, and Dean Harper.

At Cornell, I have greatly benefited from the guidance of my dissertation committee, consisting of Michael Macy, Shelley Correll, Doug Heckathorn, and Ed Lawler. My advisor, Michael Macy, has been a fantastic source of insight and contributed much to my research. I have also been fortunate to have the chance to work closely with Shelley Correll during her five years at Cornell, which was both a

tremendous learning experience and a great deal of fun. Doug and Ed have also each provided much-appreciated help and advice over the years.

Others who have provided helpful discussion and comments on the dissertation research include Emilio Castilla, Christine Horne, Ko Kuwabara, Brent Simpson, Robb Willer, and the Cornell Sociology lab group.

When I first arrived at Cornell, Robb Willer went out of his way both to make me feel welcome, and to help me get my research program started on the right foot. I very much appreciate both his insight and his friendship.

I would not have made it through graduate school (at least, not without my sanity intact) without the support of several close friends who have been great sources of joy during this time, including Pat Barclay, Ed Carberry, Jorge Martinez, Colin Meehan, In Paik, Esther Quintero, Sharon Palchak, and Dave Wirth.

I also thank Youngjoo Cha for all of her support and encouragement throughout the dissertation writing and job search process.

Finally, I thank the National Science Foundation (Dissertation Improvement Grant SES-0526214) and the Cornell Department of Sociology for supporting this research. Arguments in this dissertation are not necessarily those of the funding organizations.

TABLE OF CONTENTS

Biographical Sketch	iii
Dedication	iv
Acknowledgments	v
Table Of Contents	vii
List Of Tables	viii
Chapter 1: Introduction	1
Chapter 2: Theory	7
Chapter 3: Study 1	21
Chapter 4: Study 2	60
Chapter 5: Discussion	80
References	100

LIST OF TABLES

Table 3.1: Incentive Structure of the Within-Group Dilemma	29
Table 3.2: Incentive Structure of the Between-Group Dilemma	30
Table 3.3: Proportions by Condition	40
Table 3.4: Multilevel Model of Norm Enforcement	43
Table 3.5: Multilevel Model of Norm Enforcement, by Contribution Level	47
Table 3.6: Multilevel Model Coefficients of Contribution	50
Table 3.7: Multilevel Model Coefficients of Preference for a Leader	54
Table 4.1: Means or Proportions by Condition	66
Table 4.2: Multilevel Model of Norm Enforcement	69
Table 4.3: Multilevel Model of Contribution	74
Table 4.4: Multilevel Model of Preference for a Leader	76

CHAPTER 1

INTRODUCTION

Conflict between groups is an enduring aspect of the human experience. Intergroup conflict appears to have been widespread among early human societies, so much so that it may have contributed to keeping population growth well below its theoretical maximum for approximately 100,000 years (Bowles 2006). Despite radical changes in every aspect of human societies since then, intergroup conflict remains a persistent feature of contemporary life. For example, a recent review of the social psychological literature on the topic refers to intergroup conflict as the “problem of the century” (Fiske 2003: 123).

Intergroup conflict has many consequences, some more obvious than others. It is readily apparent that intergroup conflict changes relationships between groups. For example, it produces greater mutual dislike between groups as well as negative stereotyping of the outgroup (Brewer 2001). A subtler consequence is the way in which conflict changes relationships within groups. While intergroup conflict creates hostility between groups, sociologists have long argued that it also increases cohesion within groups (Coser 1956; Simmel 1955[1908]; Sumner 1906). This argument is the subject of far more theory than evidence (Stein 1976), nonetheless, some patterns are well-documented. In terms of attitudes, intergroup conflict increases positive evaluations and liking of ingroup members (Blake, Shepard, and Mouton 1961; Brewer 2001; Sherif 1966). In terms of behavior, intergroup conflict increases the extent to which people will cooperate with their group, in the form of making personally costly contributions to the group welfare (Bornstein 1992, 2003; Bornstein and Ben-Yossef 1994; Bornstein, Winter, and Goren 1996; Goren and Bornstein 2000).

If conflict causes people to contribute to the group, does it also cause them to *enforce* contribution? When threatened by outsiders, do group members become more likely to pressure their peers into making sacrifices for the group's well-being? In particular, does conflict cause groups to develop social norms requiring that their members contribute to the group? To answer these questions, this dissertation develops and test theories predicting that intergroup conflict promotes the emergence of intragroup social norms. In addition, while the research focuses on social norms, I also test whether conflict causes groups to create hierarchies by granting the capacity to enforce contribution to a leader.

I examine both norms and hierarchies because, in a sense, they are two sides of the same coin. Each is a way in which people can attempt to coordinate group behavior. Enforcing norms is a form of bottom-up social control, in which people use rewards or punishments to influence other ingroup members. Establishing hierarchies is a form of top-down social control, in which people grant leaders the ability to set and enforce behavioral standards. Coordinating group behavior in intergroup conflict is challenging, because group members have incentives to free-ride, or avoid contributing to the group (Bornstein 2003; Olson 1965). Group members may free-ride because of individual-level incentives, such as when a soldier flees a battle for fear of injury, or because of collective incentives, such as when one group refrains from attacking another for fear of provoking a sustained and mutually costly conflict.

Norms and hierarchies provide ways to reduce free-riding by increasing its cost, or by increasing the benefits of contribution. The incentive structures that motivate free-riding and enforcement will be discussed in more detail in chapter 3. By examining the effect of conflict on the emergence of these two forms of social structure, the dissertation seeks to contribute to two lines of research: the literature on

group conflict and group cohesion, and the literature on the emergence of social norms.

The present research aims to contribute to the conflict-cohesion literature by testing whether conflict increases the likelihood that people will enforce norms and establish hierarchies. As noted above, previous research has documented the positive effect of conflict on individuals' feelings towards and contribution to their groups. In contrast, the causal relationship between conflict, norms, and hierarchies is less well established. The present research thus examines under-studied dimensions of group cohesion. Their importance, however, stems not from their novelty, but because norms and hierarchies have important properties that set them apart from other forms of cohesion.

By distinguishing between contribution and enforcement of contribution, the dissertation shifts the question of how intergroup conflict affects intragroup dynamics from the individual level to the relational level. By testing whether conflict affects not only the choices individuals make, but also the rules they impose on others and the systems of authority they establish, the research offers an account of how a group develops from a collection of individuals into a distinct entity with a coherent structure. Enforcement of cooperation is thus not simply another way to contribute to the group, but a way to define the group. Once in place, social structures may outlast the individuals who initially established them, such as when norms are transmitted across generations (Sherif 1936). In contrast, the commitment and enthusiasm an individual brings to the group may be less likely to persist if that individual leaves the group.

This dissertation also seeks to contribute to the literature on norm emergence, by testing whether conflict serves as a source of social norms. Understanding the origin of social norms has been the focus of substantial research across the social

sciences (See Hechter and Dieter-Opp 2001 for an overview). The interest in social norms stems from the fact that they are an important building block of human societies, and yet enforcing social norms seems to run counter to individual's rational self-interest.

Social scientists have long called on the concept of norms to explain a wide range of human behaviors (Horne 2001b). Historically, social norms are thought to have structured human societies long before the advent of formal systems of law (deQuervain et al 2004). Today, norms still regulate much human behavior. In areas where law is weak or nonexistent, norms limit the extent to which people can exploit their neighbors through force or fraud (Ellis 1971). Even when people have recourse to the legal system, norms may still provide a more preferable and efficient means of settling disputes and keeping order (Ellickson 1994). Norms are sometimes beneficial, but other times harmful to individuals or groups (Benard 2008; Hechter and Dieter-Opp 2001).

Norms have been variously defined, which has made the term difficult to measure. In this dissertation, I define norms as rules that mandate or prohibit a particular behavior, and are supported by sanctions in the form of punishments, rewards, or both. While a consensus definition of norms does not exist in sociology, the definition I adopt is widely supported (Coleman 1990; Hechter and Opp 2001; Homans [1951] 1992; Horne 2001).¹ This definition also provides a clear metric for measuring norms. By this definition, norms exist to the extent that group members invest resources in enforcing them, and so can be measured by the frequency of enforcement. The costs of enforcing norms can include time, effort, and the cognitive, emotional, and material resources involved in both monitoring the group for deviants

¹ As Homans noted, without enforcement, norms are simply ideals (Homans [1950] 1992:124).

(Hechter 1987) and administering sanctions (Oliver 1980). The costs can also include the risk of reprisals from norm-breakers who retaliate when sanctioned.

Given that norms require enforcement in order to emerge and survive, and the individual cost of enforcement typically outweighs the benefit (Oliver 1980), why would individuals enforce norms? Research has produced a number of answers to this question. For example, scholars have argued that people enforce norms in exchange for other benefits, either material or reputational (Axelrod 1987; Barclay 2006; Baron and Swistak 2001; Coleman 1990; Homans [1950] 1992; Horne 2001a, 2004; Horne and Cutlip 2002; Macy, Willer, and Centola 2005; Posner 1997). Heckathorn shows that the benefits of enforcing norms can outweigh the costs for “hypocritical cooperators” who enforce norms that they themselves violate, when groups are threatened by collective sanctions from an external agent (Heckathorn 1988, 1990). Other factors underlying the emergence of norms could include low trust of others and high risk of exploitation (Yamagishi 1986, 1988), cultural evolution (Allison 1992), individual learning (Macy 1993), emotions (Fehr and Gächter 2002; Frank 1988; McAdam 1997), preference for normative groups (Güerker, Irlenbusch, and Rockenbach 2006), and neurological rewards for norm enforcement (de Quervain et al 2004). Surprisingly, research has not yet tested whether intergroup conflict contributes to the emergence of costly norms (Fehr and Fischbacher 2003). The present research seeks to contribute to understanding how norms emerge by testing whether intergroup conflict plays a causal role. Given the frequency of conflict in history (Bowles 2006), it may have contributed to the normative structure of a broad range of human societies.

Why would conflict motivate individuals to enforce norms and establish hierarchies? Drawing on and extending research from both the conflict-cohesion and norm emergence literatures, I suggest three theoretical mechanisms that might answer

this question. First, the *signaling* explanation states that people enforce norms to a greater extent in conflict because (1) enforcing norms signals trustworthiness and group commitment and (2) conflict increases the motivation to signal these traits. Second, the *intrinsic preferences* explanation states that people (1) have an intrinsic preference for enforcing norms and (2) have an evolved tendency to do so to a greater extent in intergroup conflict. Third, the *framing* explanation states that conflict increases the extent to which people rank the group's well being above their own, and therefore increases their tendency to invest resources in enforcing norms that protect the group. In Chapter 2, I will consider each of these theoretical perspectives in more detail.

In order to test these ideas, I developed a set of laboratory experiments in which two groups interact in either a conflict or a no conflict setting. By systematically manipulating the level of conflict, I can test for the causal relationship between conflict and norm enforcement. Using laboratory experiments was necessary to establish the causal effect of conflict. As will be discussed in chapter 2, ethnographic studies have documented that conflict and normative control appear to be correlated, but the direction of causality is ambiguous. Using experimental methods will address this limitation. In chapter 3, I examine the effect of conflict on norms based on punishing non-contributors. In chapter 4, I follow up by examining the effect of conflict on norms based on rewarding contributors. In chapter 5, I review the results, reflect on their limitations and broader implications, and discuss plans for future research.

CHAPTER 2

THEORY

In chapter 1, I discussed the link between intergroup conflict and norm enforcement as a theoretical and empirical problem and laid out its implications in detail. In chapter 2, I will discuss three theories that converge to predict a relationship between intergroup conflict and norm enforcement, contribution, and the emergence of hierarchies. The theories discussed in this chapter have generally sought to explain enforcement of contribution or the effect of conflict on contribution, but not the effect of conflict on enforcement (but see Horne and Chen 2005). I show that the theories converge on the prediction that conflict increases enforcement of contribution. As the theories make converging predictions, I will not attempt to adjudicate between them, instead postponing that task for later work.²

In this chapter, I begin by discussing classical theories regarding the problem, and shortcomings of those approaches. I then address these shortcomings by discussing more contemporary theories and the evidence in support of those theories. In particular, I focus on theories based on signaling, intrinsic preferences for norm enforcement, and framing. These theories posit different mechanisms, but make similar predictions.

Classical Theories

The general argument that intergroup conflict amplifies intragroup cohesion has a long history in social science theory. More than century ago, William Graham Sumner (1960 [1906]:18) argued that “[t]he relation of comradeship and peace in the

² The studies necessary to test the basic conflict and norm enforcement hypotheses, described in detail in chapters 3 and 4, required complex study designs and large (for a laboratory experiment) sample sizes. Attempting to differentiate the between the explanations suggested would have required studies that were beyond the scope of the current project. The theories are not mutually exclusive, and the mechanisms they describe may jointly contribute to the effect of conflict on norm enforcement.

we-group and that of hostility and war towards others-groups are correlative to each other.” In his view, groups become more cohesive in conflict in order to compete with outgroups more effectively. Sumner emphasized that cohesion increases not spontaneously, but through the emergence of social structures such as laws and political institutions (1960 [1906]:18).

In a similar but more developed account, Georg Simmel argued that conflict increases cohesion through two social mechanisms. First, Simmel (1955:93) argued that conflict causes groups to become intolerant of dissent, because “[t]hey cannot afford individual deviations from the unity of the coordinating principle beyond a definitely limited degree.” For example, the Catholic Church may eject or excommunicate those it defines as heretics, in order to more efficiently vie for material or political power with its competitors (Simmel 1955:93).

In a systematic evaluation of Simmel’s argument, Coser concurred with the broad claim that conflict increases intolerance of dissent and centralization, but cautioned that the process is likely contingent on a number of factors, including preexisting levels of solidarity, the type of conflict, and the values of the group members (Coser 1956:89-104). He suggested that careful theoretical and empirical analysis is necessary to determine the conditions under which conflict will promote or prevent group cohesion, and warned researchers not to assume an overly simplistic link between conflict and cohesion.

Sumner, Simmel, and Coser all drew on functionalist theories to explain their argument that conflict increases group cohesion (Coser 1956:16-20). They assumed that conflict exists because it serves a beneficial purpose, namely, integrating and strengthening societies. While functionalist explanations once dominated sociology (Turner 1982), scholars have since identified several problems with the approach. In functionalist theories, the selection process by which functional mechanisms arise is

not identified (Elster 1989), significant social changes are unexplained (van den Berg 1998), and theories are typically descriptive rather than explanatory (Erasmus 1967).

Although the validity of the functional mechanism that Simmel and others suggested is doubtful, there are nevertheless theoretical reasons to trust their intuitions regarding the link between conflict and group cohesion. In the following section, I will review three theoretical accounts predicting that conflict increases norm enforcement, contribution, and the creation of hierarchies. The posited mechanisms include signaling to develop a reputation, an intrinsic preference for sanctioning, and a social psychological framing account. In the following sections, I discuss each of these theories in turn.

Signaling

People may enforce norms to a greater extent in conflict because (1) enforcing norms and contributing to the group provide a way to signal that one is a committed group member and (2) conflict increases a person's motivation to prove that she or he is committed to the group. A number of researchers have argued that people enforce norms because doing so carries reputational benefits, for example, signaling that one is a committed or trustworthy group member (Axelrod 1986; Barclay 2006; Gintis, Smith, and Bowles 2001; Homans 1950; Hechter 1987; Horne 2004; Posner 1997). People prefer to interact with and look favorably upon individuals who are reliable and committed to the ingroup's wellbeing (Horne and Chen 2005). Investing valuable resources (time, effort, wealth) in enforcing norms is one way to signal that one is committed to the group and its goals. This assumption is supported by findings showing that people reward group members who enforce norms, and are more likely to enter into profitable exchanges with norm enforcers than with other group members (Horne 2001; Horne and Cutlip 2002; Horne 2004).

Further empirical support of the notion that people think highly of norm enforcers comes from a series of studies by Barclay (2006). This work shows that people view peers who punish uncooperative group members as more trustworthy, respectable, and group-focused than those who do not punish. They are also more likely to entrust norm enforcers with valuable resources in situations in which the truster is vulnerable to exploitation. Importantly, these effects depend on norm enforcement appearing fair and justifiable (e.g., directed at uncooperative but not cooperative group members), supporting the notion that norm enforcement must be seen as in the service of the group in order to constitute an effective signal of group commitment.

A computational model developed by Centola, Willer, and Macy (2005) shows that signaling behavior could support norm emergence in large populations, even when a majority of the population privately disapproves of the norm. Examples of such unpopular norms include professing liking for popular but inscrutable scholars, binge drinking among college students, and informing on one's neighbors in oppressive regimes. Centola, Macy, and Willer show that these behaviors can arise and persist when people enforce norms to signal their group loyalty and evade punishment from peers. The assumptions underlying this model were supported by a series of experimental studies, which showed that people will enforce norms that they privately disapprove of when in public, but not in private (Willer, Kuwabara, and Macy 2008). In addition, a model by Gintis, Smith, and Bowles (2001) shows that a strategy of signaling desirable qualities by punishing non-contributors can be evolutionarily stable in a single-shot n-person public goods dilemma.

If people enforce norms in order to signal group loyalty, they should be more likely to do so when their group is in conflict with an outgroup. One reason for this, suggested by Horne and Chen (2005), is that people care deeply about their group

affiliation. Indeed, social identity research shows that people demonstrate bias towards ingroups even when those groups are created arbitrarily and have existed for only a short time (Abrams and Hogg 1990; Hewstone et al 2002; Tajfel 1982). Conflict with an outgroup may make group identity more salient, draw sharper distinctions between in and outgroup members, and increase the motivation to signal one's commitment to the ingroup.

In addition, Willer (2006:98-100) makes a theoretical argument that conflict could increase the status rewards for contributing to the group, because contributions are seen as more critical or more badly needed during intergroup conflict than at other times. While this argument focuses on contribution, the results may extend to norm enforcement, as this can be another form of contributing to the group's welfare.

Another possibility is that conflict with an outgroup increases the motivation to signal loyalty to the group, by providing grounds for more negative interpretations of selfish behavior. In the absence of conflict with an outgroup, individuals who do not contribute to the group may be viewed as lazy or selfish. However, in conflict, interpretations of their behavior may take on a darker cast. Rather than being viewed as simply lazy, non-contributors may be viewed as unpatriotic or treasonous (Blake, Shepard, and Mouton 1964). By raising the stakes for how non-contributors may be viewed, conflict with an outgroup increases the incentive to portray oneself as a committed group member.

Historical evidence supports this assumption, showing that people are particularly sensitive to non-contribution during conflict with an outgroup (Coben 1964; Erikson 1966; Gibson 1988). For example, Coben (1964) describes an incident at a WWII victory celebration in which a man failed to stand for the national anthem, and was subsequently shot in the back by a Navy serviceman. The crowd applauded, rewarding a particularly severe form of norm enforcement (1964:52).

Of the three measures of group solidarity discussed in chapter 1 – norm enforcement, contribution to the group, and the creation of hierarchies – the signaling theory predicts a positive effect of conflict on norm enforcement and contribution, and makes no predictions about the effect of conflict on hierarchies. Conflict may exert a stronger effect on enforcement of norms, because norms provide the clearest apparent signal of one’s group commitment. While contributing a portion of one’s resources to the group signals a certain level of commitment to the group, investing further resources in pressuring *others* to contribute signals an especially strong commitment to the group. Norm enforcement might also be a more accepted signal of group commitment than contribution, if norm enforcement is less likely to be perceived to be motivated by peer influence, and more likely to be perceived as a true indicator of one’s sympathies (Centola, Macy and Willer 2005; Willer, Kuwabara, and Macy 2008). Conflict should also have a positive, though possibly weaker, effect on contribution to the group, as it theoretically provides a less efficient way to signal group loyalty. The theory makes no predictions about creating hierarchies, as supporting hierarchies does not necessarily signal loyalty to the group.

Intrinsic Preferences

People may also enforce norms to a greater extent because humans have evolved to find punishing those who exploit them or rewarding those who help them to be intrinsically rewarding. Informal enforcement of norms (i.e., without the aid of a legal apparatus) can be difficult to explain from an individual cost-benefit perspective (but see Heckathorn 1988, 1990 for an exception). Nonetheless, informal enforcement existed for thousands of years before formal systems of law developed (de Quervain et al 2004). This has led some researchers have proposed that people have an evolved preference for both enforcement of group norms and cooperation with the group, or *strong reciprocity* (Gintis, Bowles, Boyd, and Fehr 2003). In this section, I first

discuss evidence that people find enforcing norms intrinsically gratifying at the neurological level. I then draw on theories of stability-dependent cooperation to explain why the preference for norm enforcement might be greater in conflict.

Evidence of a preference for norm enforcement has been documented by studies that monitor the brain activity of individuals engaged in social dilemma experiments. One study using PET scans found that punishing non-cooperators was associated with higher levels of activation in the caudate nucleus. This is a region of the brain associated with processing rewards in primates, and also with “reinforcers such as cocaine and nicotine,” suggesting that norm enforcement is powerfully rewarding at the neural level (de Quervain et al 2004:1256). Reciprocating cooperation in a social dilemma (which has been considered a type of norm enforcement, see Axelrod 1984) has similarly been shown to correspond to the activation of reward centers in the brain (Rilling et al 2002). Related work shows that people respond empathically to pain in others at the neural level if those others have dealt fairly with them in the past, but not if they have been unfair (Singer et al 2006), which suggests that people are less hesitant to inflict pain on people who have treated them unfairly in the past.

Further evidence for a preference for norm enforcement comes from studies showing that people prefer joining strict groups rather than more lenient ones. Groups with strong norm enforcement mechanisms have a competitive advantage in recruiting members, as well as in group performance. For example, laboratory work shows that people tend to prefer groups with norm enforcement mechanisms to groups without them, even though these groups are initially more costly for their members (Gürerk, Irlenbusch, and Rockenbach 2006). This echoes findings from the sociology of religion, indicating that strict churches that place higher costs on their members tend

to have more committed members, receive greater donations, and grow more quickly than more permissive denominations (Iannaccone 1994).

If people do have an intrinsic motivation to enforce norms, why would this motivation be greater during intergroup conflict? One possible answer is that such behavior was adaptive in humans' ancestral environment. The evolutionary environment for humans was exceptionally lethal, in large part due to intergroup conflict (Bowles 2006). Estimates place death rates due to warfare at around 15% for foraging societies. In contrast, this figure for the U.S. and Europe in the 20th century was closer to 1% (Bowles 2006: 1572). In such an environment, groups in which members responded to conflict by contributing and enforcing norms of contribution would be better able to mobilize for war, and would thus have a competitive advantage over less cohesive groups.

The theory of *stability-dependent cooperation* (Lahti and Weinstein 2005), argues that people respond with selective altruism (including altruistically punishing non-cooperators), when the group faces an external threat, but otherwise behave more selfishly. This could explain why people might respond to potentially dangerous outgroup by becoming more solidary, but not cooperate with their ingroup to the same degree in times of peace. Stability dependent cooperation could be a successful evolutionary strategy, because adopters enjoy the benefits of group solidarity when faced with a group-level external threat, but the benefits of more individually-oriented strategies otherwise. Preliminary experimental evidence supports the argument that people follow a strategy of stability dependent cooperation (Barclay and Benard 2008).

The intrinsic preference account predicts that conflict will increase norm enforcement, contribution to the group, and the creation of hierarchies. If individuals follow a strategy of stability-dependent cooperation, then enforcing norms of

contribution, contributing to the group, and creating hierarchies are all steps that can be taken to increase the group's likelihood of prevailing in the conflict.

Framing and Motivation

An additional theoretical explanation for the relationship between group conflict and group solidarity centers on the framing effects of conflict. This explanation is less well developed than the other two perspectives. It appears as a sketch of a theoretical mechanism, rather than a fully elaborated theory, in a single paper by Bornstein and Ben-Yossef (1994). However, it does offer an explanation for the conflict-norm enforcement link, and has at least some empirical support.

This perspective argues that framing a situation as a conflict between two groups causes people to value the ingroup's outcomes more than their individual outcomes (Bornstein and Ben-Yossef 1994). Once a person values the good of the group over their personal welfare, it becomes rational to sacrifice one's resources to help the group prevail in conflict. This explanation was invoked by Bornstein and colleagues to explain their finding that people tend to contribute more to the group in public goods experiments when the ingroup is in competing with another group than when it is not (Bornstein and Ben-Yossef 1994). Importantly, in their study the material incentive to contribute to the group was the same when the ingroup faced an outgroup in competition and when it did not. In both cases, the rational decision was to withhold contribution (see chapter 3 for more details). Participants contributed more when competing with an outgroup, even though the noncontribution was the rational strategy regardless of the presence of the outgroup.³ This suggests that psychological factors other than the direct material benefit of contributing play a role in determining participants' responses to conflict.

³ However, one critic of this work has argued that participants may have simply misunderstood the study, and believed that contributing was materially beneficial (Baron 2001).

The researchers found some support for their hypothesis by surveying study participants regarding their motivations. They found that, relative to groups that did not experience conflict with an outgroup, members of groups that did experience conflict reported lower motivation to maximize individual gain, greater motivation to maximize ingroup gain, and greater motivation to maximize the difference between ingroup and outgroup earnings. This suggests that conflict can shift individuals' preferences to favor the group's welfare. However, the mechanism by which conflict changes preferences is left unspecified.

Bornstein and colleagues used the motivational explanation to account for increased contribution to the group in the face of conflict with an outgroup. This account could also explain why individuals might be more likely to enforce norms and create hierarchies. Enforcing norms and creating hierarchies both make it easier to coordinate collective action, and thus maximize ingroup earnings as well as the difference between the ingroup and the outgroup.

Forms of Conflict

The three theories discussed above are united in their prediction that conflict produces norm enforcement. This claim raises a question. What kinds of conflict are sufficient to trigger the mechanisms discussed above? Framing conflict as an intergroup public goods problem, in which individuals decide whether to make costly contributions to their group, raises two possibilities. First, it may be the case that the intergroup competition aspect of the payoff structure will be sufficient to induce higher levels of solidarity, compared to a standard single-group public goods problem. This would be the case if the *potential* for conflict with an outgroup is sufficient to motivate people to enforce norms.

However, if the outgroup does not actively contribute to collective action (i.e., conflict with the ingroup), then the intergroup public goods problem closely resembles

a single-group public goods problem, in that no external costs are levied on the ingroup. This suggests a second possibility: it may be the case that actual costs must be incurred in competition with an outgroup, in order to motivate people to invest resources in enforcing norms. In this case, individuals would enforce norms to a greater extent in an intergroup public goods problem as the outgroup behaved more competitively, by investing in conflict with the ingroup.

This question can be phrased in more familiar language by asking, is it sufficient for two groups to have conflicting goals, such that they have the potential to cause harm to one another? Or is it necessary for the outgroup to actively *pursue* those goals, thereby prosecuting the conflict and posing a threat to the ingroup? While this point has not been actively debated within the literature on conflict and group cohesion, theorists have independently reached different conclusions. Muzafer Sherif's realistic group conflict theory argued that competition over incompatible goals constitutes the sufficient condition under which group conflict emerges (Sherif 1966:85). Sherif argued that competition between interdependent groups produces "glorification" of and greater loyalty to the ingroup and as well as "derogation" of and greater hostility to the outgroup (Sherif 1966:120). This theory suggests that the presence of a potentially competitive outgroup will be sufficient to motivate greater norm enforcement.

In contrast, Coser pointed out that sometimes groups become more cohesive in conflict with an outgroup, yet other at times they tend to fall apart or become more anomic. From this observation, he reasoned that conflict alone might not be enough to motivate group cohesion. Drawing on work by Robin Williams (1947), Coser argued that groups will become more cohesive in conflict to the extent that the outgroup poses a threat to the ingroup. This is most similar to the second possibility discussed above, in which individuals will enforce norms to a greater extent when the outgroup invests

resources in competing with the ingroup. In chapter 3, I will explain how I will empirically evaluate both of these possibilities.

Empirical Evidence

Several theories predict that intergroup conflict will produce norm enforcement, contribution, and the creation of hierarchies. Evidence in this regard is strongest for the group cohesion hypotheses. Some evidence for the hierarchy hypothesis exists as well. There is little systematic evidence for the effect of conflict on norm enforcement, though research suggests that such a link exists in real-world conflict.

A number of studies find that conflict increases cohesion by magnifying liking and evaluations of the ingroup. Sherif's field studies of boys in summer camp found that, in conflict, people positively stereotype ingroup members, see them as possessing more desirable qualities, and overestimate their performance on valued tasks (Sherif 1966:89-93). Similarly, experiments with corporate executives found that competition with an outgroup led to greater group cohesion, measured by participants' increased positive evaluations of their groups (Blake and Mouton 1961:421-422), as well as more centralized leadership structures (Blake, Shepard, and Mouton 1964: 19-23). There is also a long line of research in social identity theory showing that the presence of an outgroup can increase positive evaluations of and trust towards ingroup members (Brewer 2001).

Studies focusing on hierarchy have found that conflict increases the influence of leaders. This is true both in experimental work (Hamblin 1958), and in field studies of large organizations (Staw et al 1981). An informal review suggests that dictatorships tend to be preceded by conflict or another type of crisis (Hertzler 1940). However, there is less systematic evidence that conflict increases the creation of leadership hierarchies.

Evidence also suggests that intergroup conflict increases cooperation with the ingroup. A series of small group studies by Bornstein and colleagues provides evidence that conflict can increase individuals' willingness to make costly personal sacrifices for the group good (Bornstein 1992, 2003; Bornstein and Ben-Yossef 1994; Bornstein, Winter, and Goren 1996; Goren and Bornstein 2000). Bornstein and his collaborators used an experimental setting in which two teams of participants repeatedly interacted under varying levels of conflict. Bornstein and his collaborators found, across a number of studies, that participants more frequently sacrificed for the group when the two teams were in conflict than when they were not (see Bornstein 2003 for a review).

While evidence shows that conflict increases cooperation with ingroup members, there are no systematic tests of the hypothesis that conflict increases *enforcement* of cooperation. Field studies are suggestive in this regard. Reviews of the anthropological literature suggest that groups that experience regular conflict with outgroups tend to have strict norms requiring participation in conflict (Gould 2003; Black-Michaud 1975). In nineteenth century Corsica (as well as New Guinea, Montenegro, and other "honor societies"), men were forced to choose between risking their lives to support family members in revenge killings against rival families, or suffering dishonor and loss of social standing (Gould 2003:128). More recently, reports from Iraq indicate that Sunni militant groups use suicide bombings and targeted assassinations to kill and intimidate moderate Sunnis whom they perceive as disloyal, particularly those who join the national police force or the army (Wagner 2006). However, it is difficult to determine the direction of causality from these case studies: does conflict increase enforcement of norms, are strict, normative groups more likely to participate in conflict, or both?

More generally, a review of the conflict and cohesion literature in sociology, anthropology, psychology, and political science concluded that an association between conflict and group cohesion exists, but also that this association depends on a number of factors (Stein 1976). The review criticized the “casual acceptance of the [conflict and cohesion] hypothesis without any of the caveats” and “the paucity of empirical literature” regarding that hypothesis (Stein 1976:145). Thus, while research is fairly clear on a link between conflict and cohesion, the nature of this link requires much elucidation.

In summary, classical perspectives in sociology predicted a relationship between intergroup conflict and intragroup cohesion, norm enforcement, and centralization. The classical theorists drew on functionalist theories, which have a number of shortcomings. Nonetheless, several other theoretical perspectives concur with the intuitions of earlier sociologists. These include theories based on signaling, evolved preferences, and the social psychological effects of framing. Evidence that conflict increases intragroup liking and cooperation is strong, and to a lesser extent there is evidence that conflict centralizes groups. There is less evidence that conflict increases enforcement of norms promoting cooperation, though field study evidence is suggestive on this point.

In the following chapters, I attempt to test the conflict and cohesion hypothesis more systematically. I will examine contribution to the group, enforcement of contribution, and the emergence of leadership hierarchies as separate outcomes of conflict with an outgroup. I will also attempt to develop a more nuanced picture of the relationship between conflict and cohesion by determining whether the structural potential for conflict with an outgroup is sufficient to motivate group cohesion, or whether active conflict with the outgroup is necessary.

CHAPTER 3

STUDY 1: THE EFFECT OF CONFLICT ON PUNISHMENT-BASED NORM ENFORCEMENT

Chapter 2 presented several converging theories predicting that conflict among groups promotes the enforcement of norms within groups. Despite this convergence, a review of the empirical literature showed that this hypothesis has not been systematically tested. In Chapter 3, I present an experiment designed to test hypotheses derived from these theories. The theories discussed in Chapter 2 argued that conflict could increase enforcement of norms for several reasons, including:

- **Signaling:** People may enforce norms, even when they are costly, to signal that they are a trustworthy group member, and increase the likelihood that others interact with them in productive ways in the future. Signaling trustworthiness may become more important in conflict, because of increased concerns about member loyalty (Horne and Chen 2005).
- **Intrinsic Preferences:** People may experience stronger neurological rewards for punishing norm violators when their group is threatened by an outgroup than when there is no threat, which may make them more prone to disregarding the material costs of enforcing norms.
- **Framing:** Presenting a situation as a conflict between two groups may increase a person's psychological motivation to see their group do well relative to the outgroup, independent of the material incentives. Motivation to do well relative to the outgroup could motivate enforcement of norms requiring contribution to the group good (Bornstein and Ben-Yossef 1994).

Chapter 2 also pointed out that two of these theories, the intrinsic preference and framing theories, predict that conflict encourages the creation of group hierarchies.

The theories presented in Chapter 2 offer explanations of the mechanism by which conflict could increase norm enforcement and create hierarchies within the group. Chapter 2 also considered the types of conflict that could be sufficient to trigger these mechanisms. On one hand, within-group solidarity may be triggered by a set of conditions that I refer to as structural conflict, which occurs when the parties have incompatible goals and are capable of inflicting some kind of loss on one another (Oberschall 1978; Sherif 1966). On the other hand, some theorists have argued that structural conflict can produce *either* solidarity or anomie, and that a further condition is necessary for conflict to motivate solidarity-enhancing mechanisms such as norm enforcement. In particular, they argue that groups become more solidary in conflict to the extent that the outgroup poses a threat to the ingroup (Coser 1956; Williams 1947). For the purposes of this study, I define “threat to the group” to mean that the outgroup inflicts costs on the ingroup by investing resources in conflict.

I test two overarching questions. First, is structural conflict sufficient to motivate the enforcement of norms, contribution to the group and the creation of hierarchies? Second, if structural conflict is *not* sufficient, does outgroup contribution to conflict provide the necessary trigger?

These questions produce three pairs of hypotheses. First, I test the hypotheses that conflict motivates norm enforcement, under the two conditions just described:

H1a: Conflict with an outgroup will lead group members to enforce norms at greater rates.

H1b: Conflict with an outgroup will lead group members to enforce norms at greater rates, *to the extent that the outgroup behaves competitively.*

Other work has previously shown that intergroup conflict can increase cooperation within groups (Bornstein 2003). However, earlier work has not examined the distinction between structural conflict alone and structural conflict in combination

with outgroup competition. I test this distinction, as described in hypotheses H2a and H2b:

H2a: Conflict with an outgroup will lead group members to contribute to the group at greater rates.

H2b: Conflict with an outgroup will lead group members to contribute to the group at greater rates, *to the extent that the outgroup behaves competitively*.

Finally, I test the hypotheses that conflict motivates the creation of group hierarchies, under the two conditions of interest:

H3a: Conflict with an outgroup will lead group members to create leadership hierarchies at greater rates.

H3b: Conflict with an outgroup will lead group members to create leadership hierarchies at greater rates, *to the extent that the outgroup behaves competitively*.

Note that the “b” hypotheses predict an interaction between structural conflict and levels of outgroup competition, such that groups in conflict develop norms of contribution when the outgroup invests in conflict, but not when the outgroup does not invest. In the next section, I briefly describe how I will test these hypotheses, before moving on to a detailed description of Study 1.

Study Overview

In order to test these hypotheses, I used a controlled laboratory experiment in which pairs of three-person groups interacted in either a conflict condition or a no conflict condition. In the conflict condition, the two groups competed in an experimental setting commonly used to model intergroup conflict (an "intergroup prisoner's dilemma", described in detail below); in the control condition, two groups interacted in a setting that lacked conflict but was otherwise identical. The conflict x no conflict conditions served as the manipulation of structural conflict.

On each of 30 rounds, participants decided whether to make personally costly contributions to the group. These contributions created a benefit for the ingroup in both conditions. In the conflict condition, contribution to the group imposed costs on the outgroup in addition to benefiting the ingroup. I therefore used the level of the outgroup's contribution as the measure of outgroup competition. Participants also decided whether to punish, at a cost, teammates who did not contribute to the group.

After 30 rounds, participants decided whether to maintain the current democratic structure of the group, or instead to elect a leader with the ability to set and enforce contribution standards for the group. Because of time constraints, I did not ask participants to repeat the experiment after making the leader decision, although they believed that they would do so. In the following sections, I describe the methods and results of the study in detail.

METHOD

Participants

Participants consisted of 120 paid undergraduate volunteers (83 women and 37 men) who were recruited via flyers offering payment for participation in an experimental study.⁴ There were 60 participants (10 sessions, each with 2 groups of 3 persons) in each condition.⁵

Procedure

When participants arrived at the lab, a research assistant led them to a private cubicle equipped with a computer. After participants gave written consent to participate in the study, a research assistant read all participants a brief overview of the

⁴ Gender did not exert a main effect on any of the behavioral dependent measures, nor was there a significant interaction effect of gender with conflict on any of the behavioral dependent measures. For a discussion of why one would not expect gender to affect behavior in a prisoner's dilemma setting, see Simpson (2003).

⁵ Initially, 126 individuals participated, but the data for one session of six participants was lost due to a computer error.

study. Participants then completed a detailed computer tutorial that explained the structure of the study and how to complete all of the necessary tasks. In addition to the computer tutorial, all participants received a printed form featuring tables summarizing the payoffs from all possible combinations of choices. The form also included bulleted reminders of the most important aspects of the setting.

Following the tutorial, participants responded to a short quiz designed to ensure they understood the rules. A research assistant checked each participant's quiz, and discussed any incorrect responses with the participant until the research assistant was satisfied that the participant fully understood the study. The study only proceeded when the research assistant was confident that all participants understood and were comfortable with the instructions. Each session was randomly assigned to either the "conflict" condition or the "no conflict" condition. In the conflict condition, the two teams competed with one another in a task known as an intergroup prisoner's dilemma. In the no conflict condition, the two teams each interacted in separate, identical tasks, known as a three-person prisoner's dilemma. Both the prisoner's dilemma and the intergroup prisoner's dilemma are explained in the next section.

The Intergroup Prisoner's Dilemma

The intergroup prisoner's dilemma (abbreviated IPD) belongs to a class of models designed for studying intergroup conflict (Palfrey and Rosenthal 1983; Bornstein 1992, 2003; Bornstein and Ben-Yossef 1994; Bornstein, Winter, and Goren 1996; Goren and Bornstein 2000). It is an extension of the standard prisoner's dilemma, which is a conceptual model commonly used to study cooperation in situations in which group and individual interests conflict (see Rapoport and Chammah 1965, Axelrod 1984, or Komorita and Parks 1996 for introductions to the prisoner's dilemma literature).

In a standard prisoner's dilemma (abbreviated PD), an individual, as part of a dyad or larger group, must choose to either cooperate or defect. In the one-shot game, defection earns the higher payoff, regardless of the partner's choice. Thus defection is expected if individuals are acting egoistically. Individuals earn one of four ordered payoffs, depending on their choices and the choices of the other group members. Individuals earn the highest payoff, labeled "T" (for "temptation"), by defecting when other group members cooperate. The next highest payoff (labeled "R", for "reward") is earned by cooperating when others also cooperate. The second worst payoff (labeled "P", for "punishment"), is earned when all group members defect. The worst payoff (labeled "S", for "sucker") is earned by cooperating while others defect. The PD models a broad range of decisions in which participants choose between cooperation and defection, such as whether to recycle, donate blood, attend a political demonstration, and, in the example motivating the original PD, whether to testify against one's partner in crime in return for a lighter sentence.

In the intergroup prisoner's dilemma, cooperation takes the form of contributing to a conflict with an outgroup. The key difference between the IPD and the PD is that in the IPD, cooperating with the ingroup harms an outgroup, while in the PD no outgroup is affected by the ingroup's behavior.

The IPD was originally designed to study voting behavior in two-party systems (Palfrey and Rosenthal 1983). Voting is a costly behavior (in terms of the time and effort spent voting, as well as opportunity costs). Voting for one's own party concurrently helps that party while harming the other party in an election. The logic of the IPD applies equally well to any scenario in which individuals' actions are personally costly, help their ingroup, and harm an outgroup, including wars, labor-management disputes, and inter-departmental conflicts in corporate and academic settings. Across all of these settings, the decision of whether to contribute to one's

group poses a dilemma. Contributing to help one's group in a conflict can be costly and entail substantial personal sacrifice, but doing so can also help one's group avoid defeat. Dawes (1980) succinctly captured this tension when he wrote:

Soldiers who fight in a large battle can reasonably conclude that no matter what their comrades do, they are better off taking no chances; yet if no one takes chances, the result will be a rout and slaughter worse for all the soldiers than is taking chances. (p.170, quoted in Bornstein 2003:129)

In the IPD, as suggested by Dawes's example, individuals simultaneously participate in two nested prisoner's dilemmas. First, they participate in a *within-group* dilemma: do I contribute some of my own resources to help the group, or do I instead save my resources to further my individual interests? Second, they participate in a *between-group* dilemma: should our group behave aggressively towards the outgroup, or should we be conciliatory? Contribution to one's group (voting, attacking the enemy, joining a strike) is simultaneously cooperation in the within-group dilemma, and defection in the between group dilemma. Contributing to one's group incurs individual costs, provides benefits for the group, and imposes costs on the outgroup. Not contributing to one's group avoids incurring individual costs, withholds benefits from the group, and avoids imposing costs on the outgroup.

Formally, the (IPD) is defined by five features (Bornstein 2003). First, each session consists of 2 teams (designated A and B, by convention), each with n members. In this study, I use $n = 3$ for each team. I use triads because this is the smallest number of people per team that can be considered a group in the sociological sense (Simmel 1950). Second, each participant begins with an endowment of size e ($e > 0$), and must decide whether to contribute or keep the endowment. For simplicity, participants always decide whether to keep or contribute the *entire* endowment. Third, payoffs for members of Team A increase with the number of contributors in Team A

(m_A) and decrease with the number of contributors in Team B (m_B). An equivalent relation holds for members of Team B. Fourth, the values of all of the parameters are known by all of the participants. Fifth, contributions from participants are assumed to help their teams.

In the current study, the conflict condition is an intergroup prisoner's dilemma in which the total earnings, h , of each participant, i , per trial were determined by the following schedule, shown here for a participant on Team A.⁶

$$h_i = 5(m_A - m_B) + 15 + 10(1 - \sigma_i) \quad (1)$$

In (1), each participant's contribution to the group increments all team members' earnings (including their own) by five points, and decrements the earnings of members of the opposing team by five points. This is accomplished by multiplying the difference between the number of contributors on team A and the number of contributors on team B ($m_A - m_B$) by five points. Fifteen points are added to participants' payoffs each round to assure that their earnings are always positive. Finally, σ_i is an indicator variable that equals one if the participant contributed their ten point endowment and zero otherwise.

The intergroup prisoner's dilemma captures the incentive structure of intergroup conflict, while abstracting away many of the complexities that make group conflict difficult to study empirically. It does so by explicitly modeling the tensions between individual, group, and collective rationality characteristic of intergroup conflict. Individual-level rationality asks, "How should *a single participant* act so as to maximize her or his welfare?" Group-level rationality asks "How should *the group as a whole* act to maximize its welfare?" Collective rationality asks, "How should *everyone in the conflict* behave as a unit to maximize their welfare?"

⁶ Earnings during the study are given as abstract values ("points") rather than monetary values because some participants might value money differently than others (Homans 1961).

In the IPD, it is individually rational for participants to withhold contribution, because the amount by which contribution increases the payoff of each participant (five points) is less than each participant's starting endowment (ten points). This models the fact that the costs of participating in a particular conflict often outweigh the immediate rewards. For example, voters expend effort traveling to the polls, but usually have negligible effects on an election as individuals, a soldier joining a battle risks injury and death, but can rarely turn the tide of the battle alone, and an individual worker loses pay by striking with no guarantee of obtaining a better contract. The incentive structure at the individual level is illustrated in Table 3.1:

Table 3.1: Incentive Structure of the Within-Group Dilemma

Team B	Focal Actor (A)	Total (A)	Payoff (FA)	Label (FA)
0	0	0	25	P
0	1	1	20	S
0	0	2	35	T
0	1	3	30	R

The first column, "Team B", indicates the number of contributors in the outgroup. Here, we hold the number of contributors constant at zero. Because the function described in equation (1) is linear, this value does not affect the incentive structure of the game. The second column, "Focal Actor (A)" indicates whether a particular individual from Team A contributed. A zero indicates non-contribution and a one indicates contribution. The third column, "Total (A)" indicates the total number of actors on Team A contributing, including the focal actor. A zero indicates no members of Team A contributed, a one indicates that only the focal actor contributed, a two indicates that everyone *except* the focal actor contributed, and a three indicates that all three team members contributed. The fourth column, "Payoff (FA)" indicates the payoff to the focal actor in each situation. The final column, "Label (FA)" shows the corresponding prisoner's dilemma label for each payoff to the focal actor. This chart illustrates that individuals always earn more on a given round by not contributing

to their team. It also illustrates that the within-group game is a prisoner's dilemma, because the payoff ordering is $T > R > P > S$.

In contrast, it is rational at the group level (the between-group dilemma) for participants to join the conflict by contributing their endowments. In the IPD, this is reflected in the fact that the team, as a unit, profits most when all of its members contribute and profits least when none contribute. This models the fact that the group as a whole typically fares better when all of its members contribute than when they do not. Political parties benefit from high voter turnout among their members, armies are more likely to win when their soldiers fight aggressively, and unions have more leverage when strike votes are unanimous. The payoff ordering for the between-group dilemma is illustrated in Table 3.2:

Table 3.2: Incentive Structure of the Between-Group Dilemma

Team A	Team B	Payoff (A)	Payoff (B)	Label (A)
3	3	45	45	P
3	0	90	30	T
0	3	30	90	S
0	0	75	75	R

The first column, "Team A," indicates the number of contributors on Team A. Because I focus here on the between-group dilemma, I only consider cases in which all team members contribute or all team members withhold contribution (i.e., cases in which each group is uniformly aggressive or uniformly conciliatory). The second column, "Team B," indicates the number of contributors on Team B. The next two columns, "Payoff (A)" and "Payoff (B)" indicate the group-level payoffs (the sum of the individual payoffs within each team) for Team A and Team B, respectively. The final column, "Label", indicates the corresponding prisoner's dilemma payoff label. Recall that cooperation in the within-group game is simultaneously competition in the between group game. Therefore, the first row of Table 3.2, in which all members of both teams contribute, displays the case in which both teams defect in the between-

group dilemma. Thus, each group earns the punishment payoff of 45. In general, the ordering of the payoffs is $T > R > P > S$, indicating that the between-group dilemma is also a prisoner's dilemma.

Table 3.2 illustrates that, while it is rational at the group level for all team members to contribute, it is also *collectively rational to withhold contribution*.⁷ The collective level refers to all participants in the dilemma as a whole. The “collectively rational” option is the choice that would provide the greatest total payoff, assuming all participants were required to make the same decision. All participants in the conflict fare better as a whole if no one contributes to the conflict, because they can conserve the resources that they would otherwise devote to conflict and avoid the costs that their opponents would inflict. This is illustrated by comparing the P payoff and the R payoff in Table 3.2. The population as a whole earns 90 points when everyone contributes, and 150 points when no one contributes. Non-contribution is thus collectively rational in the IPD. In general, by accounting for conflicts between individual, group, and collective welfare, the IPD mirrors the incentive structure of many real-world conflicts.

Contribution, Norms and Hierarchies

In the current study, sessions consisted of 2 teams of 3 participants each, labeled the “Red” and “Green” teams.⁸ Each participant began every round with an endowment of ten points in a “team account”, and decided whether to invest the endowment in the team or keep the endowment. Kept endowments were added to a

⁷ In IPD research, the term “collectively rational” has been used to refer to the outcome that maximizes the total earnings of everyone in the interaction. It is the choice that would be made if the collectivity made decisions as a single actor (Bornstein 2003: 132). This differs from the more common usage of the term, which denotes the Pareto-optimal point(s) in the dilemma (e.g., Arrow 1950). Full defection is collectively rational in the IPD in both usages. In the latter usage, full cooperation is also collectively rational (Pareto optimal).

⁸ Two teams are also used in the single-group PD condition to confirm that the results are due to conflict. If the conflict condition had included two teams, the no conflict condition included only 1 team, we would not know if any results found were the product of conflict itself, or the presence of a second team.

participant's total earnings for the study and were not available for further investment in the team. Thus far, the IPD as described follows the standard laboratory setting for studying intergroup conflict (Bornstein 2003). However, participants in the standard IPD setting make only one choice: whether to contribute to the conflict, or keep one's resources for oneself. In order to test the hypotheses that conflict increases norm enforcement and the creation of group hierarchies, the present study extends the IPD framework by allowing participants to enforce norms by negatively sanctioning their team members.

In addition to the team account available for investment, participants also received a ten-point endowment in a "punishment account" at the beginning of every round. When participants decided whether to invest their team account on a given round, they also decided whether to invest their punishment account in a group punishment fund. If one participant on a team did not contribute during a trial, the total amount invested in the group punishment fund on that round was subtracted from that participant's total earnings. If more than one participant did not contribute, each non-contributing participant lost the amount in the group punishment fund divided by the number of non-contributors. If a participant kept their punishment account on a given round, their ten point endowment was added to their total earnings for the study and was not available for later investment. This methodology was adapted from Yamagishi (1986, 1988).⁹

In the no conflict condition, the two teams each interacted in separate three-person prisoner's dilemmas (PD), instead of competing in a single IPD. To change the

⁹ Because individuals could not identify individual punishers, this may weaken the extent to which the signaling theory accounts for the results of the study. However, given the small ($n = 3$) group size, this may not be a serious problem. Participants may still perceive punishers as more group-motivated than punishers, and therefore behave more cooperatively (either to reward the punishers or because participants trust punishers to cooperate more) even when they do not know which of the two teammates punished. When both teammates punish, this is not a problem at all. Further work is needed to test the signaling account directly.

IPD used for the conflict condition into a PD suitable for the no conflict condition, I set m_B in equation (1) to a constant, so that payoffs are only determined by the behavior of one's team, and one can neither harm nor be harmed by the outgroup. Based on previous studies, I set m_B to 1. Other research has explored the effects of varying the constant to which m_B is set and found it does not affect the results (Bornstein and Ben-Yossef 1994).

Each session included 30 rounds, although participants were not informed of when the study would end as a way to limit “endgame” effects (e.g., Axelrod 1984).¹⁰ Participants interacted in a practice round before the rounds began to count toward their earnings. After all participants made contribution and sanctioning decisions, their computer screens displayed the results of each trial. Participants saw their gross earnings that round, the magnitude of any losses from punishment that round, their net earnings after punishment that round, and their total net earnings so far. Participants also learned how many participants on each team had contributed and punished in that round and in the session thus far.

I also extended the standard IPD setting to allow participants to create group hierarchies. After completing all of the rounds, participants were informed that they would next play another set of rounds. First, they were asked to vote on whether the group should continue with the current format, or change formats and elect a leader.¹¹ Participants were informed that the leader would be able to set rules for how much each participant must invest each round, and punish participants who did not invest

¹⁰ Reducing endgame effects may increase the incentive to cooperate, depending on the strategies of other players (Axelrod 1984). This does not change any of the theoretical predictions tested in the study.

¹¹ Participants did not actually play another set of rounds after making the leader decision. This step was taken for two reasons. Theoretically, this study focuses on explaining the emergence of norms and leaders, not leaders subsequent behavior. On a practical level, the study already lasted as long as seemed reasonable, given participants' compensation. Examining the effect of conflict on leaders' behaviors in particular would be an interesting avenue for future work.

this amounted. This aspect of the study was adapted from prior research on preferences for leaders in social dilemmas (Messick et al. 1983; Samuelson et al. 1984; Rutte and Wilke 1985; Samuelson and Messick 1986).

Following the behavioral component of the study, participants responded to a series of survey questions. These included demographic questions such as gender and year in school, as well as questions similar to those used by Bornstein and Ben-Yossef (1994) to assess the framing hypothesis. Participants responded on a 100-point scale, anchored at “Not True” and “Very True,” to the following statements: *I wanted to earn as many points as possible for myself, I wanted my group to earn as many points as possible, I wanted everyone in the game to earn as many points as possible, and I wanted my group earn more than the other group.* These questions were included to be used as a preliminary evaluation of the framing hypothesis. If conflict produces norm enforcement by shifting preferences from individual outcomes to group outcomes, responses to these measures should mediate the effect of conflict on norm enforcement. Following the survey questions, participants were then debriefed, paid, and thanked for their participation.

DATA AND MODEL

Analysis of the experimental data focuses on three dependent measures: norm enforcement, contribution to the group, and preference for a leader. These dependent measures are analyzed with three primary independent variables: the presence of conflict, the level of contribution from the outgroup, the level of contribution from the ingroup, and the level of sanctioning from the ingroup. The following paragraphs discuss the dependent and independent measures in more detail.

Dependent Measures

Norm enforcement: The primary focus of this study is to determine whether conflict increases the enforcement of norms requiring sacrifice for the group. The

measure of norm enforcement was a binary variable valued at 1 if the participant invested in the punishment fund that round, and 0 otherwise. These data were collected for every participant on every round of the study, resulting in 3,600 observations.

Contribution: The measure of contribution to the group was a binary variable valued at 1 if participants contributed their endowment to the group on a given round, and 0 otherwise. These data were collected for every participant on every round of the study, resulting in 3,600 observations.

Leader Preference: The decision to elect a leader is a binary variable valued at one if the participants voted to elect a leader, and zero otherwise. This question was asked once after participants finished the IPD or PD tasks, resulting in 120 total observations.

Independent Variables

Conflict Condition: The experimental manipulation of conflict is the key independent variable. This will allow us to test whether, as hypothesized, norm enforcement, contribution to the group, and preferences for a leader are greater in the conflict condition.

Outgroup Contribution: In addition to the hypotheses that structural conflict is sufficient to promote the emergence of norms and hierarchies, the study tests whether the effect of conflict is conditional on competition from the outgroup. To test the hypotheses that conflict causes the emergence of norms and hierarchies to the extent that the outgroup inflicts costs on the ingroup, I created a variable measuring the number of contributors in the outgroup on every round, ranging from zero to three. Outgroup competition is defined as the level of outgroup contribution because, in the conflict condition, this variable indicates the extent to which outgroup members invested resources in harming the ingroup. In the control condition, outgroup

contribution does not directly affect the payoffs of the ingroup, but might affect their behavior if ingroup members view them as symbolically competitive (e.g., because they don't want an outgroup to be more successful at mobilizing than the ingroup).

Because participants' behavior can logically only be affected by outgroup behaviors in the past (that is, behaviors that the participants have had a chance to observe), I created a one-round lag variable. This variable indicates, for each participant on every round (except the first round, for which no such information exists), the number of outgroup contributors on the *previous* round. This variable allows for a test of whether the level of outgroup competition on the previous round affected participants' behaviors on the current round.

I also created a variable for the interaction of the conflict condition with the lag variable. The interaction of conflict and the lagged outgroup contribution variable allows for a test of the hypotheses that participants become more likely to enforce norms in conflict to the extent that the outgroup behaves competitively. A positive interaction effect indicates support for this hypothesis.

Ingroup Contribution and Ingroup Norm Enforcement: While the project does not make predictions regarding the effects of ingroup behaviors (e.g., contributions to the group and norm enforcement by fellow group members), it is important to account for the effects of these variables. In particular, outgroup behaviors may be correlated with ingroup behaviors, and so it is necessary to control for the effect of ingroup behaviors to ensure we are accurately estimating the effects of outgroup contribution. To do so, I created variables for each participant measuring the number of team members contributing to the group on each round, and the number of team members investing in the punishment fund that round. I then created one-round lag versions of each variable, and variables measuring the interaction these variables with conflict.

Because the behavioral independent variables are lagged for one round, the first round of the study is excluded from analysis. This leaves a total of 3,480 participant-rounds for analysis. The behavioral independent variables are also standardized for ease of interpretation.

Data Structure

The experiment produces a data set consisting of 3,480 participant-rounds, in which rounds of the study are nested within participants, participants are nested within teams, and teams are nested within sessions. As a result, individual observations are not independent, thereby violating an assumption of OLS regression models. Prior studies of group behavior within a social dilemma framework have compensated for non-interdependence by analyzing group averages, for example, treating the mean level of contribution in a session as a single observation.

This approach is insufficient for testing the present hypotheses, which focus on the interaction between conflict and the contribution and sanctioning behavior of ingroup and outgroup members. Because the behavior of ingroup and outgroup members varies across rounds, it is necessary to use round-level data.

Model

To adjust for non-interdependence of observations while also taking advantage of the round-level data, I use a four-level hierarchical linear model, in which rounds are nested within participants, participants are nested within teams and teams are nested within sessions. At level one, norm enforcement on each round (or contribution, depending on the hypothesis being tested) is modeled as an outcome of the behavior of others (outgroup contribution, ingroup contribution, and ingroup sanctioning) and the interaction of these behaviors with conflict.¹² At level two, differences in levels of norm enforcement across individuals are modeled as random

¹² A logit link function is used for binary dependent measures.

errors that are correlated within individuals (i.e., across rounds for each individual). At level three, differences in norm enforcement across teams modeled as random errors that are correlated within teams. At level four, differences in norm enforcement across sessions are modeled as a function of condition (conflict or no conflict) as well as random error correlated within sessions. Conflict is modeled as a level-4 predictor because sessions, rather than individuals or teams, are randomly assigned to the conflict or no-conflict condition. I use a random intercepts model, in which baseline levels of norm enforcement are assumed to vary across individuals, teams, and sessions, and the effects of the predictor variables are assumed to remain constant.

I estimate the model using STATA's GLLAMM program (Rabe-Hesketh and Skrondall 2005). GLLAMM uses maximum likelihood estimation with adaptive quadrature. Adaptive quadrature is a form of numeric integration that, while computationally demanding, is well suited for multilevel analyses, including those using binary outcome measures. This approach produces unbiased parameter estimates and standard errors and allows for a test of the hypotheses while accurately accounting for the structure of the data.

To illustrate the structure of the model, a simplified level one equation for norm enforcement, in which lagged outgroup contribution, ingroup contribution, and ingroup norm enforcement are represented with the term *Behavior* for clarity of presentation, is shown below.

$$Y_{ijkl} = \beta_{0_jkl} + \beta_{1_jkl}(Behavior_{ijkl}) + \beta_{2_jkl}(Behavior_{ijkl} * Conflict) + r_{ijkl} \quad (2)$$

In this model, Y_{ijkl} is the propensity to enforce norms on round i by participant j on team k in session l . The intercept β_{0_jkl} is the average propensity to enforce norms of individual j on team k in session l . The coefficient β_{1_jkl} is the effect of others' behavior on the propensity to enforce norms of individual j on team k in session l . The coefficient β_{2_jkl} is the effect of the interaction of others' behavior and conflict on

propensity of individual j to enforce norms on team k in session l . Finally, r_{ijkl} is the unmeasured individual-level error for participant j on team k in session l . The level two equations for this model are below.

$$\beta_{0jkl} = \gamma_{00kl} + u_{0jkl} \quad (3)$$

$$\beta_{1jkl} = \gamma_{10kl} \quad (4)$$

$$\beta_{2jkl} = \gamma_{20kl} \quad (5)$$

In equation (3), the average propensity β_{0jkl} to enforce norms of person j on team k in session l is modeled as γ_{00kl} , the average propensity to enforce norms on team k in session l plus u_{0jkl} , the unmeasured error for person j on team k in session l . Because β_{1jkl} the effect of others' behavior on individual propensities to enforce norms, is not theoretically predicted to vary across individuals, it is modeled in equation (4) as γ_{10kl} , the effect of others' behavior on propensity to enforce norms on team k in session l . For the same reason, β_{2jkl} , the interaction effect of conflict with others' behavior is modeled in equation (5) as γ_{20kl} . The level three equations are similar to the level two equations:

$$\gamma_{00kl} = \alpha_{000l} + \sigma_{00kl} \quad (6)$$

$$\gamma_{10kl} = \alpha_{100l} \quad (7)$$

$$\gamma_{20kl} = \alpha_{200l} \quad (8)$$

At level three, γ_{00kl} , the average propensity to enforce norms on team k in session l , is modeled as the average propensity to enforce norms in session l (α_{000l}), plus σ_{00kl} , the unmeasured error for team k in session l . The effect of others' behavior on norm enforcement on team k in session l is not expected to systematically vary across teams, and is therefore modeled as α_{100l} , the effect of others' behavior on norm enforcement in session l . The interaction effect of others' behavior and conflict on norm enforcement is also not expected to vary across teams, and is modeled as α_{200l} , the interaction effect of others' behavior and conflict on norm enforcement in session l .

The level four equations include the conflict measure, as well as the session-level error term for the intercept:

$$\alpha_{000l} = \pi_{0000} + \pi_{0001}(\text{Conflict}) + \varepsilon_{000l} \quad (9)$$

$$\alpha_{100l} = \pi_{1000} \quad (10)$$

$$\alpha_{200l} = \pi_{2000} \quad (11)$$

At level four, α_{000l} is modeled as the sum of the average propensity to enforce norms π_{0000} , the effect of conflict on norm enforcement π_{0001} , and the session-level error term ε_{000l} . Additionally, α_{100l} is modeled as π_{1000} , the effect of others' behavior on propensity to enforce norms, and α_{200l} is modeled as π_{2000} , the interaction effect of conflict and others' behavior on norm enforcement.

This simplified system of equations illustrates the structure of the model. In the model analyzed in the paper, I include separate measures of outgroup contribution, ingroup contribution, ingroup sanctioning, and time (round of the study), as well as the interaction of these variables with conflict.

RESULTS

Bivariate Relations

To begin evaluating the findings, I first compare the proportions of participants choosing to enforce norms, contribute to the group, and elect a leader, shown below in Table 3.3.

Table 3.3: Proportions of Participants Choosing Norm Enforcement, Contribution, and to Elect a Leader, by Conflict Condition

	No Conflict	Conflict	Total
Norm Enforcement	0.194 (0.24)	0.173 (0.228)	0.18 (0.23)
Contribution	0.55 (0.34)	0.642 (0.275)	0.59 (0.31)
Leader	0.40 (0.49)	0.42 (0.497)	0.41 (0.49)

Note: SDs in parentheses N = 120 (60 per condition).

Table 3.3 shows that, across conditions, participants contributed on approximately 0.60 of all rounds, enforced norms on approximately 0.18 of all rounds, and chose to elect leaders on approximately 0.41 of all rounds (reading down the final row, labeled “Total”). For the norm enforcement measure, there was a nonsignificant tendency to enforce norms less frequently in the conflict condition. Reading across row 2 (labeled “norm enforcement”), one can see that participants invested in the punishment account on about 0.17 of all rounds in the conflict condition, and on about 0.19 of all rounds in the no-conflict condition. For the contribution and leader variables, the results trend in the direction of the hypotheses, but are not statistically significant. Reading across row 3 (labeled “contribution”), one can see that participants contributed to the group about 0.55 of all rounds in the no conflict condition, and about 0.64 of all rounds in the conflict condition. Participants also voted establish a group hierarchy 0.40 of all rounds in the no-conflict condition, and 0.42 of all rounds in the conflict condition (as shown in row 4, “Leader”). None of the differences across conditions were statistically significant, as measured by t-tests of the session-level means.

The descriptive statistics suggest that, in the aggregate, structural conflict by itself may not be sufficient to motivate enforcement of cooperation. This analysis, however, does not tell us whether responses to conflict might be conditioned on the behavior of ingroup and outgroup members. Are participants in conflict more likely to enforce norms if the outgroup actively engages in competition? We can answer this question by looking at the interaction of conflict and outgroup behavior in the multivariate models considered in the following section.

Multivariate Models

Next, I present an analysis of the experimental data that systematically considers the effects of conflict, outgroup contributions, ingroup contributions, and

ingroup norm enforcement, using the multilevel model described above. I also control for the effects of round and the interaction of round and conflict, to account for the fact that participants' behavior might change over time (Rapoport and Chammah 1965; Bornstein, Winter, and Goren 1996).

Norm Enforcement

As the primary goal of this dissertation is to determine the effect of conflict on norm enforcement, I examine the norm enforcement results first. Table 3.4 presents the coefficients from a series of models estimating the effects of conflict and the standardized behavioral variables on norm enforcement. The first two models focus on the main effect of conflict: model 1 includes the conflict variable only, model 2 adds the effects of round and the interaction of round and conflict. Models 3-5 separately add the main effects and interactions with conflict of ingroup contribution, ingroup norm enforcement and outgroup contribution, respectively. Models 6-8 include the behavioral variables and their interactions with conflict in pairs: model 6 includes ingroup contribution and ingroup norm enforcement, model 7 includes ingroup contribution and outgroup contribution, and model 8 includes ingroup norm enforcement and outgroup contribution. Model 9 includes all of the variables.

I first evaluate hypothesis H1a, the prediction that participants would be more likely to enforce norms of contribution to the group in the conflict condition, relative to the no conflict condition. This hypothesis can be evaluated by examining the main effect of conflict, in row 1 of model 1 (labeled "Conflict" of Table 3.4). As the table shows, the main effect of conflict is negative and not statistically significant. This result fails to support H1a, the hypothesis that structural conflict will be sufficient to motivate increased norm enforcement. This conclusion holds when we consider other model specifications. Looking across row 1, we can see that the effect of conflict is

Table 3.4: Multilevel Model of Norm Enforcement

Fixed Effects	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9
Conflict	-0.18 (0.54)	-0.06 (0.59)	-0.04 (0.63)	-0.19 (0.57)	-0.26 (0.61)	-0.11 (0.59)	-0.17 (0.65)	-0.35 (0.58)	-0.24 (0.6)
Ingroup Contribution			-0.22** (0.11)			-0.21+ (0.12)	-0.20+ (0.11)		-0.19 (0.12)
Ingroup Norm Enf.				-0.04 (0.1)		0.01 (0.1)		-0.04 (0.1)	0.01 (0.11)
Outgroup Contribution					-0.21 (0.13)		-0.18 (0.13)	-0.21 (0.13)	-0.18 (0.13)
Round		0.05** (0.01)	-0.05** (0.01)	-0.05** (0.01)	-0.05** (0.01)	-0.05** (0.01)	-0.06** (0.01)	-0.06** (0.01)	-0.06** (0.01)
Conflict X Ingroup C			-0.05 (0.15)			-0.14 (0.16)	-0.06 (0.15)		-0.14 (0.16)
Conflict X Ingroup NE				0.25+ (0.14)		0.28+ (0.15)		0.25+ (0.14)	0.29* (0.15)
Conflict X Outgroup C					0.38+ (0.18)		0.31+ (0.18)	0.38* (0.18)	0.30 (0.18)
Conflict X Round		-0.01 (0.01)	-0.01 (0.01)	0.00 (0.01)	0.00 (0.01)	0.00 (0.01)	0.00 (0.02)	0.01 (0.01)	0.01 (0.02)
Constant	-2.06** (0.38)	-1.36** (0.41)	-1.37** (0.45)	-1.41** (0.4)	-1.36** (0.43)	-1.37** (0.41)	-1.31** (0.45)	-1.34** (0.41)	-1.30** (0.42)
Random Effects									
Intercept-1 (u_{0ikl})		2.63** (0.58)	2.98** (0.67)	2.98** (0.67)	2.94** (0.66)	3.08** (0.69)	2.97** (0.67)	2.94** (0.66)	3.06** (0.69)
Intercept-2 (σ_{00kl})	0.90 (0.65)	0.98 (0.73)	0.98 (0.79)	0.98 (0.79)	0.78 (0.71)	0.63 (0.66)	0.91 (0.76)	0.56 (0.64)	0.57 (0.67)
Intercept-3 (ε_{000l})	0.45 (0.50)	0.49 (0.59)	0.62 (0.66)	0.62 (0.66)	0.59 (0.60)	0.52 (0.54)	0.70 (0.67)	0.52 (0.53)	0.60 (0.57)

Note: Standard errors in parentheses. N = 3,480 observations for all models. The model includes 29 level one units, 120 level two units, and 40 level three units, and 20 level 4 units.

+ Significant at 10% level; * significant at 5% level; ** significant at 1% level

negative and not statistically significant in all of the models. This suggests that structural conflict between groups is not sufficient to promote the emergence of norms.

I now turn to assessing hypothesis H1b. Do norms emerge in conflict to the extent that the outgroup behaves competitively? We can evaluate this prediction by examining row 8 of models 4, 6, 8, and 9, in Table 3.4 (labeled “Conflict X Outgroup C”), which includes the coefficients for the interaction of conflict with outgroup contributions. Positive values of this coefficient indicate that participants increasingly enforce norms in the conflict condition to the extent that members of the outgroup contribute to the conflict. Table 3.4 indicates that hypothesis H1b is supported by the data. Study participants tend to punish noncontributors at greater rates in the conflict condition, but only when outgroup members contribute to the conflict. This can be seen by the positive interaction effect of conflict and outgroup competition. This effect is marginally significant at the 10% level across all the models 4, 6, and 8. The effect dips slightly past the 10% significance level in model 9.

To understand the substantive meaning of the coefficients, may be easiest to consider the effect of the variables on the likelihood that a given individual chooses to punish a non-cooperative group member. In model 4, for example, given the negative main effect of outgroup competition ($b = -0.21$) and the positive interaction effect of conflict and outgroup competition ($b = 0.38$), we can see that the net change in norm enforcement from a one standard deviation increase in the level of outgroup competition in the conflict condition has a net coefficient of 0.19. If we exponentiate this difference to calculate the odds ratio, we find that a one standard deviation increase in the level of outgroup competition in the conflict condition increases the likelihood of an individual investing in punishment to enforce a norm by a factor of 1.19, or about 19%.

The results thus show that the positive effect of conflict on norm enforcement is conditioned on the behavior of the outgroup. When the outgroup behaves competitively, individuals in the conflict condition become more likely to punish uncooperative group members. Individuals in the no-conflict condition are unaffected by the outgroup's behavior, as shown by the non-statistically significant main effect of outgroup contribution.

Three other statistically significant findings were not predicted theoretically. First, there is a negative effect of ingroup contribution (row 2), that is statistically significant in some of the models. This finding is unsurprising: the more an individual's teammates contributed to the team's welfare, the less likely that individual was to punish them for failing to contribute to the group. Second, there is a significant and negative main effect of round (row 5), such that participants tend to contribute less over time. This is an extremely common finding in prisoner's dilemma research (e.g. Rapoport and Chammah 1965, Bornstein, Winter, and Goren 1997), and is usually interpreted as a learning effect.¹³ Third, there is a positive interaction effect of ingroup norm enforcement and the conflict manipulation that ranges from marginally to highly significant. This means that participants were more likely to punish non-contributors if their teammates did so as well, but only in the conflict condition. While not predicted, this finding does not contradict the theory, and could have important implications. In particular, it suggests that conflict could make individuals more sensitive to normative influences from peers. This finding will be discussed in more detail in the discussion section.

¹³ It is interpreted as a learning effect because participants profit more on a single round when they do not contribute to the group. It is argued that the rationality of non-contribution in the stage game becomes clearer to participants with increased experience, thereby making them less likely to contribute.

High and Low Contributors

Recent research suggests that high contributors will be the most active punishers in a public goods setting (Gürerk et al 2006), perhaps because they are more invested in the success of the group or attaining the public good. In contrast, low contributors have been shown to be less active enforcers of contribution norms. I conduct an exploratory analysis to see if high contributors will also be more responsive to conflict than low contributors. If high contributors are more invested in the group welfare than low contributors, they may be more likely to respond to threats from outside of the group by enforcing norms of contribution. I test this possibility by dividing the sample into high and low contributors using a median split of the proportion of times an individual contributed to the group (median = 0.6167), and then running the full model (model 9 in Table 3.4) separately for high and low contributors.¹⁴ The results are shown in Table 3.5 below. I first discuss the results for high contributors, shown in column 1.

As before, there is no significant effect of the conflict manipulation, though the non-significant effect increased in magnitude. As with the complete sample, the high contributors sample offers no support for hypothesis 1a. In contrast, support for hypothesis 1b is stronger among high contributors compared to the total sample. In the conflict condition, high contributors respond to outgroup contributions by becoming much more likely to punish non-contributors. This is shown by the positive, highly significant interaction of conflict and outgroup contribution on likelihood of investing in the punishment fund. The coefficient ($b = 0.67$), is approximately twice the

¹⁴ Using a median split has advantages over using contribution as a continuous predictor. The median split approach is “fully interactive”, meaning it is equivalent to examining the interaction effect of being a high or low contributor with every other variable in the model. This allows us to easily compare differences between high and low contributors on each predictor variable. Using contribution as a continuous predictor would result in a single model with 5 additional 2-way interactions and 4 additional 3-way interactions, resulting in a much more complex interpretation task.

Table 3.5: Multilevel Model of Norm Enforcement, by Contribution Level

Fixed Effects	High Contributors	Low Contributors
Conflict	-1.404 (1.053)	0.384 (0.456)
Ingroup Contribution	-0.659** (0.174)	0.098 (0.160)
Ingroup Norm Enf.	-0.178 (0.157)	0.209 (0.146)
Outgroup Contribution	-0.325+ (0.181)	0.001 (0.173)
Round	-0.069** (0.015)	-0.051** (0.016)
Conflict X Ingroup C	0.135 (0.243)	-0.276 (0.213)
Conflict X Ingroup NE	0.568** (0.220)	-0.03 (0.201)
Conflict X Outgroup C	0.670* (0.272)	0.038 (0.244)
Conflict X Round	0.041+ (0.022)	-0.02 (0.022)
Constant	-0.59 (0.740)	-1.673** (0.314)
Random Effects		
Intercept-1 (u_{0jkl})	6.05* (3.028)	1.165 (3.090)
Intercept-2 (σ_{00kl})	1.413* (0.625)	0.000 (0.000)
Intercept-3 (ε_{000l})	0.927 (0.586)	0.020 (0.111)

Note: Standard errors in parentheses. N = 1,740 observations for all models. The model includes 29 level one units, 120 level two units, and 40 level three units, and 20 level 4 units.

+ Significant at 10% level; * significant at 5% level; ** significant at 1% level

magnitude of the equivalent coefficient for the full sample. Substantively, this means that, after accounting for the marginally significant negative main effect of outgroup contribution, a 1-standard deviation increase in the level of outgroup contribution in the conflict condition increases the likelihood of investing in the punishment account by a factor of 1.41, or about 41%. High contributors are much more responsive to conflict with an outgroup than other group members when deciding whether to enforce norms.

There are some additional differences when comparing the results for high contributors to the results for the full sample. The negative main effect of ingroup contribution is now highly significant and has tripled in magnitude. This suggests that high contributors are not only more sensitive to outgroup behavior, but are also more sensitive to ingroup behavior as well.

The positive, significant interaction between conflict and ingroup norm enforcement has nearly doubled (to $b = 0.568$), suggesting that the effect of norm enforcement may be contagious for high contributors in conflict: when their peers enforce norms, they become more likely to enforce norms as well. This is interesting, because from a rational choice perspective, it pays to allow others to shoulder the burden of enforcing norms. Thus, we might expect individuals to use costly punishment less as their peers punish more. Instead, we find the opposite pattern. It may be the case that punishment becomes more legitimate as others use it, though this conjecture requires empirical testing to be verified.

One final difference is that, for high cooperators, there is a positive, significant interaction of conflict and round. This qualifies the negative and significant main effect of round, so that norm enforcement decreases more slowly over time for high contributors than it does for the sample as a whole. This suggests that high

contributors are more persistent in their norm enforcement efforts than other individuals.

The results for high contributors contrast with those for low contributors (shown in column 2). For low contributors, the only significant finding is the negative effect of round. In addition to the lack of statistical significance, the effects of the independent variables are also substantively small. For example, the coefficient for the interaction of conflict and outgroup contribution is 0.038, or approximately 1/20th of the magnitude of the corresponding coefficient for the high contributors. This suggests that the findings for norm enforcement are largely driven by the behavior of high contributors. The implications of this finding will be discussed in greater detail in the discussion section.

Framing

Finally, I tested the mediational effects of the framing hypothesis questions (results not shown). I found that responses to the group versus individual preference questions did not mediate the significant effects of conflict on norm enforcement. While further study is needed, this suggests that the framing mechanism may not be the best explanation for this outcome.

Contribution to the Group

While the present study focuses on the emergence of norms and hierarchies, it also allows us to examine the effect of conflict on contribution to the group. Earlier studies have found that people are more willing to make costly contributions to their groups when they are in conflict with an outgroup (Bornstein 2003). However, prior work does not examine contribution to the group in conflict when a system of norm enforcement is also available. Research has found that the presence of a sanctioning system can actually undermine altruism and cooperation (Fehr and Rockenbach 2003), so it may be the case that the effects of conflict on contribution will be dampened

Table 3.6: Multilevel Model Coefficients of Contribution

Fixed Effects	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9
Conflict	0.66 (0.68)	0.47 (0.74)	0.40 (0.67)	0.53 (0.71)	0.58 (0.74)	0.44 (0.67)	0.50 (0.66)	0.66 (0.71)	0.56 (0.66)
Ingroup Contribution			0.39** (0.10)			0.23* (0.10)	0.39** (0.10)		0.23* (0.10)
Ingroup Norm Enf.				0.57** (0.09)		0.51** (0.09)		0.57** (0.09)	0.52** (0.09)
Outgroup Contribution					0.06 (0.11)		0.01 (0.11)	0.06 (0.11)	0.02 (0.11)
Round		-0.07** (0.01)	-0.06** (0.01)	-0.06** (0.01)	-0.07** (0.01)	-0.06** (0.01)	-0.06** (0.01)	-0.06** (0.01)	-0.06** (0.01)
Conflict X Ingroup C			-0.08 (0.12)			0.04 (0.13)	-0.10 (0.13)		0.02 (0.13)
Conflict X Ingroup NE				-0.30* (0.12)		-0.31* (0.13)		-0.30* (0.12)	-0.30* (0.13)
Conflict X Outgroup C					-0.28+ (0.14)		-0.20 (0.14)	-0.28+ (0.14)	-0.22 (0.15)
Conflict X Round		0.01 (0.01)	0.01 (0.01)	0.01 (0.01)	0.01 (0.01)	0.01 (0.01)	0.01 (0.01)	0.00 (0.01)	0.01 (0.01)
Constant	0.55 (0.48)	1.71** (0.53)	1.49** (0.48)	1.50** (0.51)	1.61** (0.52)	1.43** (0.47)	1.49** (0.47)	1.48** (0.50)	1.42** (0.47)
Random Effects									
Intercept-1 (u_{0ikl})	1.89** (0.40)	2.12** (0.45)	2.50** (0.51)	2.26** (0.47)	2.10** (0.44)	2.52** (0.52)	2.50** (0.52)	2.27** (0.47)	2.52** (0.52)
Intercept-2 (σ_{00kl})	3.75** (1.22)	4.18** (1.36)	2.98** (1.11)	3.68** (1.24)	4.08** (1.33)	2.92 (1.10)	2.89** (1.09)	3.55** (1.21)	2.83** (1.07)
Intercept-3 (ε_{000l})	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.01)	0.00 (0.00)	0.00 (0.00)

Note: Standard errors in parentheses. N = 3,480 observations for all models. The model includes 29 level one units, 120 level two units, and 40 level three units, and 20 level 4 units.

+ Significant at 10% level; * significant at 5% level; ** significant at 1% level.

when norm enforcement is possible. I examine the effects of conflict on contribution to the group in Table 3.6, below. Table 3.6 allows for an examination of the two hypotheses regarding contribution: H2a, which predicts that conflict alone will motivate contribution, and H2b, which predicts that conflict will motivate contribution to the extent that the outgroup behaves competitively. The models in Table 3.6 are identical to those in Table 3.4, with the exception that contribution to the group is the dependent variable.

Hypothesis H2a can be evaluated by examining row 1 (“Conflict”) of Table 3.6. Model 1 in row 1 shows the main effect of conflict. The main effect of conflict is positive, but not statistically significant, and thus fails to support hypothesis H2a. Individuals do not contribute to the group significantly more often when the group is in conflict with an outgroup than when it is not in conflict. Furthermore, adding additional control variables to the model does not change this substantive conclusion. As can be seen in models 2-9, the effect of conflict remains positive but not statistically significant across all model specifications.

In the analysis of the norm enforcement data, I found that conflict alone did not motivate norm enforcement. Instead, participants became more likely to enforce norms in conflict only when the outgroup aggressively pursued that conflict. Does a similar pattern emerge for contribution? Hypothesis H2b predicts that individuals will contribute more to the group in conflict to the extent that the outgroup behaves competitively.

We can answer this question by examining row 8 in Table 3.6 (“ConflictX Outgroup C”), which includes the coefficient for the interaction effect of the conflict manipulation and the level of outgroup member’s contribution to the conflict on the prior round. The results show that, even when coupled with outgroup competition, conflict does not appear to increase the likelihood of contribution to the group. In fact,

there appears to be a weak negative effect of conflict and outgroup contribution. The coefficients for the conflict/outgroup competition interaction are negative, and marginally significant in models 5 and 8. Contrary to predictions, in no case does conflict appear to increase contribution to the group. These findings will be considered in greater detail in the discussion.

As in the model of norm enforcement, some unpredicted effects appear in the results for contribution. While not predicted, these results do not contradict the hypotheses. First, study participants are more likely to contribute to the group if their own group members have contributed on the prior round, as shown by the positive and significant effect of ingroup contribution across all model specifications (row 3). This finding, that participants contribute more when their peers do, is consistent with goal/expectation theory (Pruitt and Kimmel 1977), which argues that people actually prefer the mutual cooperation outcome in prisoner's dilemmas, and thus will cooperate when others cooperate.

Second, study participants tended to contribute more when their peers enforced norms on the previous round, as shown by a significant and positive effect of ingroup norm enforcement reported in row 4 of table 3.6. Interpretation of this effect is straightforward: the more participants were punished for non-contribution in the past, the more likely they were to contribute in the present. Interestingly, this effect is weaker in the conflict condition, as shown by a negative and significant interaction of conflict and ingroup norm enforcement, shown in row 8 (labeled "Conflict X Ingroup NE"). This negative effect is of lesser magnitude than the main effect of ingroup norm enforcement, such that there is a reduced, but still positive overall, effect of norm enforcement on contribution to the group. Why would punishment be less effective in the conflict condition? This may be a ceiling effect. Rates of contribution are approximately 16% greater in the conflict condition (64%) than in the no-conflict

condition (55%), and so participants may have been less willing to increase their contribution rates even further. An implication of this finding is that punishment overall was effective, even if it was somewhat weaker in the conflict condition.

Third, the effect of round is negative and significant. This means that contributions declined over time in both conditions. As discussed earlier, this is a common finding usually attributed to learning effects. As there were no significant effects of conflict to be explained, I did not conduct a mediational analysis to test the framing hypothesis.

Preference for a Leader

I now turn to examining the effects of conflict on preferences for a leader. Does intergroup conflict increase willingness to relinquish one's choices to a leader? Because the preference for a leader variable was collected at the close of the study, I use individual-level data, rather than round-level data, to test this hypothesis. In addition, instead of using lagged measures of outgroup contribution, ingroup contribution, and ingroup norm enforcement, I use the standardized average levels of these variables over the course of the study as covariates, as well as the interaction of the covariates with conflict.

Table 3.7 shows the multilevel logit model coefficients for the regression of leader preference on conflict and the behavioral variables. The table includes eight models. Model 1 includes only the main effect of conflict. Models 2-4 include the main effect of conflict, and the main effects and interactions with conflict of ingroup contribution, ingroup norm enforcement, and outgroup contribution, respectively. Models 5-7 include the behavioral variables and their interactions with conflict in pairs: model 5 includes ingroup contribution and ingroup norm enforcement, model 6 includes ingroup contribution and outgroup contribution, and model 7 includes

ingroup norm enforcement and outgroup contribution. Model 8 includes all of the variables.

I begin by evaluating hypothesis H3a, which predicts that individuals will be more likely to relinquish some of their decision-making power to a leader when their group is in conflict with an outgroup than when it is not. We can evaluate this prediction by referencing row 1 of model 1, the main effect of conflict. This shows that conflict with an outgroup alone does not significantly affect preference for a leader, as evidenced by non-significant and substantively small ($b = 0.07$) main effect of conflict. Thus, the results fail to support hypothesis H3a.

Table 3.7: Multilevel Model Coefficients of Preference for a Leader

Fixed Effects	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
Conflict	0.07 (0.52)	0.17 (0.56)	0.09 (0.51)	-0.21 (0.51)	0.21 (0.53)	-0.18 (0.52)	-0.20 (0.50)	-0.14 (0.52)
Ingroup Contribution		0.06 (0.30)			-0.02 (0.33)	0.03 (0.28)		-0.06 (0.32)
Ingroup Norm Enf.			0.14 (0.39)		0.21 (0.42)		0.20 (0.34)	0.24 (0.4)
Outgroup Contribution				-0.17 (0.28)		-0.16 (0.28)	-0.16 (0.27)	-0.16 (0.28)
Conflict X Ingroup C		-0.87 (0.56)			-0.84 (0.58)	-0.18 (0.52)		-0.11 (0.57)
Conflict X Ingroup NE			-0.05 (0.50)		0.00 (0.53)		-0.19 (0.45)	-0.18 (0.5)
Conflict X Outgroup C				1.64** (0.54)		1.56** (0.58)	1.62** (0.54)	1.55** (0.58)
Constant	-0.48 (0.37)	-0.45 (0.40)	-0.49 (0.36)	-0.47 (0.34)	-0.47 (0.38)	-0.46 (0.34)	-0.47 (0.34)	-0.49 (0.34)
Random Effects								
Intercept-1 ($u_{0,ikl}$)	0.99 (0.82)	0.10 (0.71)	0.91 (0.81)	0.00 (0.00)	0.07 (0.70)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
Intercept-2 (σ_{00kl})	0.00 (0.00)	0.66 (0.68)	0.00 (0.00)	0.36 (0.42)	0.55 (0.64)	0.35 (0.42)	0.32 (0.42)	0.31 (0.41)

Note: Standard errors in parentheses. N = 120 observations for all models. The model includes 120 level one units, 40 level two units, and 20 level 3 units.

+ Significant at 10% level; * significant at 5% level; ** significant at 1% level.

Next, I turn to evaluating hypothesis H3b, which predicts that people will prefer a leader in conflict to the extent that the outgroup behaves competitively. In order to evaluate this hypothesis, I examine row 8 in table 3.7, which reports the interaction effects of conflict and outgroup contribution (labeled “Conflict X Outgroup C”). The findings show a significant, positive effect across all models. In the conflict condition, people were more likely to prefer a leader when the outgroup had behaved competitively. However, outgroup behavior did not significantly affect individuals’ leader preferences in the no conflict condition. In addition, the results are substantively large. Model 4 predicts that, in the conflict condition, a one standard deviation increase in the average level of outgroup competition over the course of the study will cause an individual will become approximately 4.3 times more likely to willingly relinquish some of their autonomy to a leader.¹⁵ As in the case of the norm enforcement variables, the significant effects were not mediated by the framing hypothesis questions. No other significant effects were predicted, and none were found. I also checked to see if there were differences in preference for a leader among high and low contributors (results not shown, available upon request). As creating a leadership hierarchy can serve as a form of social control, one might expect high contributors to be more likely to prefer a leader. However, none of the predicted effects varied substantially between high and low contributors. This may be due to the fact that norm enforcement is costly, but electing a leader is not. Apparently, only those willing to contribute to the group at high levels responded to conflict by enforcing norms, but all group members were equally willing to submit to initially costless top-down social control.

¹⁵ The 4.3 figure is the odds ratio obtained by exponentiating the difference between the coefficients of the interaction of conflict and outgroup contribution in model 4 (1.64), and the main effect of outgroup contribution in model 4 (-0.17), i.e. $e^{(1.64-0.17)}$.

Discussion

The results of Study 1 support the hypothesis that conflict leads people to enforce norms of contribution to the group to the extent that the outgroup inflicts costs on the ingroup. Participants in the conflict condition, but not in the no-conflict condition, responded to past outgroup contributions were becoming more likely to invest resources in enforcing ingroup contribution. On average, a standard deviation increase in outgroup contribution in the conflict condition increased participants' likelihood of investing resources in punishment by 19%, but had no significant effect in the no conflict condition. In fact, there was a non-significant tendency for outgroup contribution to reduce investment in punishment in the no conflict condition. The results also showed a similar interaction effect of conflict and outgroup competition on preference for a leader. The findings show that a standard deviation increase in outgroup contribution over the course of the study increased participants' likelihood of voting to create a centralized leadership structure by over 400% in the conflict condition, but had no significant effect in the no conflict condition. The results did not support the prediction that structural conflict, in the form of antagonistic group goals, increases norm enforcement or contribution, in the absence of outgroup contribution. The study also found no evidence that conflict, either alone or in conjunction with outgroup behavior, increases contribution to the group. The results do suggest an indirect link between conflict and contribution to the group, by which conflict increases punishment of non-contributors, and punishment increases contribution.

The primary contribution of the study to show that conflict can shape group structure, by causing group members to create rules for behavior and rudimentary systems of authority from previously unstructured groups. Historical case studies have suggested that group conflict and group structure are intertwined, but disentangling the direction of causality from case study data can be challenging. This study

complements existing work by showing that conflict can shape group structure. Furthermore, experimental work has shown that conflict can motivate contribution, but has yet to test whether conflict can produce *enforcement* of cooperation as well. The present study contributes to this literature by showing two ways in which conflict can produce enforcement: either directly, through the application of sanctions, or indirectly, by investing a leader with the authority to set and enforce behavioral standards.

The findings that the positive effect of conflict on norm enforcement and contribution to the group is contingent on competition from the outgroup speaks to longstanding theoretical differences in the literature on intergroup conflict and intragroup cohesion. Sherif (1966) argued that intragroup solidarity arises when two groups have incompatible goals, a condition that I have called structural conflict. In contrast, others have argued that the outgroup must pose a real threat to the ingroup in order for conflict to motivate solidarity (Coser 1956; Williams 1947). The findings thus far lend more support to the latter arguments. It appears that the potential for conflict alone does not motivate solidarity, instead, the outgroup must actively inflict costs on the ingroup before group members will increase their investments in norm enforcement and relinquish some of their autonomy to a leader.

The analysis of the behavior of high and low contributors added a further level of detail to the findings by showing that high contributors, but not low contributors, respond to conflict with a competitive outgroup by becoming more likely to punish non-contributors. Statistically and substantively, the effects of conflict and a competitive outgroup are strong for high contributors, but non-existent for low contributors. This dovetails with theoretical results obtained by Centola, Macy, and Willer (2005), which suggest that a core of “true believers” underlies norm enforcement. It also dovetails with results from public goods research indicating that cooperation is established by a

subpopulation of “strong reciprocators” who actively contribute to the group and punish those who do not (Gürerk et al 2006). The present study extends our understanding of this phenomenon, by showing that strong reciprocators are not only generally more likely to enforce norms, but are also more responsive to conflict with an outgroup. The results shown here indicate that, while groups may become stricter and more punitive in conflict, this tendency does not necessarily characterize the population of the whole. Instead, “true believers” who are more committed to the success of the group may drive the emergence of norms in conflict.

Interestingly, high and low contributors are equally likely to respond to conflict with a competitive outgroup by relinquishing some of their autonomy to a leader. As discussed earlier, this may be because enforcing norms is costly at the outset, whereas electing leaders is not. More generally, this distinction may have implications for understanding how people resolve the Hobbesian problem when faced with threatening outgroups (Ellis 1971). Hobbes proposed that humans could either exist in a state of nature, in which the strong would exploit the weak, or submit to the “leviathan,” a strong leader who would maintain order at the cost of freedom. Social scientists have pointed to the bottom-up emergence of norms as an alternative to either anarchy or domination from the top down (e.g., Macy 1993), what Garrett Hardin referred to as “mutual coercion, mutually agreed upon” (Hardin 1968:1247). Indeed, informal norms often provide a superior alternative to formal legal control (e.g., Ellickson 1994). However, this requires that some members of the group assume the costs of enforcing norms. The present study offers some insight into when different forms of social control may emerge. When facing a competitive outgroup, people had a generic tendency to resort to electing leaders. However, only high contributors were willing to invest in informal social control. This suggests that the presence of high

contributors may be an important determinant of a group's response to conflict with an outgroup.

The fact that conflict did not increase contribution to the group is intriguing, given that prior studies have found a relationship between conflict and contribution to the group (e.g. Bornstein 2003). There are differences between this study and earlier studies that may account for this discrepancy. This is the first study to make the use of sanctions available in an IPD setting. Sanctions may suppress contribution, particularly when they are perceived to be unfair (Fehr and Rockenbach 2003). If sanctions are, for some reason, seen as less fair in the conflict condition, this might explain why no effect of conflict on contribution was found. In addition, most prior studies conducted using the IPD were conducted in Israel, rather than the US (as is the case with the present study). Cultural differences between the two nations, particularly cultural differences relating to group conflict, could help explain why conflict did not appear to affect contribution in the present study.

Finally, the study included a basic test of the framing hypothesis, by including questions designed to measure whether the effect of conflict was mediated by a change in participants' preferences for group outcomes relative to individual outcomes. No mediational effect was found. The failure to support the mediating effect found by Bornstein might be explained by the same factors that led to the differing results for contribution compared to Bornstein's work. Further research is necessary to explore this possibility.

Study 1 examines the effect of conflict on group members' willingness to enforce punishment-based norms. The main finding is that conflict increases norm enforcement and preference for a leader, when the outgroup actively participates in the conflict. The norm enforcement finding is especially strong for high contributors to the group.

However, there is another side to norm enforcement. Groups often use rewards as well as punishments to enforce norms within their ranks. Chapter 4 explores how conflict affects reward-based norms, and examines whether the dynamics of reward-based norms differ from those of punishment-based norms.

CHAPTER 4
STUDY 2: THE EFFECT OF CONFLICT ON REWARD-BASED NORM
ENFORCEMENT

In chapter 3, I reported the results of study 1, which found that conflict promotes enforcement of punishment-based norms and the emergence of hierarchies, to the extent that outgroup members contribute to the conflict. Study 1 did not find an effect of conflict on contribution to the group, net of other factors. In chapter 4, I report the methods and results of study 2, in which norms may be enforced with rewards, rather than punishments. Study 2 also allows participants to enforce norms of non-contribution, as well as norms of contribution. These changes address potential limitations of study 1, described below.

One potential limitation of study 1 is that participants could only enforce norms by using punishment. While punishment is an important dimension of norm enforcement, rewards are another (Oliver 1980). Punishment encourages compliance by imposing a cost on deviants; rewards encourage compliance by supplying a bonus to conformists. An advantage of punishment is that it is potentially less costly, as it need only be used when deviance occurs. If the threat of punishment effectively induces conformity, then actual punishment need never be used. Rewards, in contrast, must be constantly supplied when conformity is high (Oliver 1980; Pruitt and Rubin 1986). This is an important consideration to the extent that the decision to enforce norms is purposive and incentive-driven.

One disadvantage of punishment is that it may engender resentment or counter-punishments (Oliver 1980; Pruitt and Rubin 1986), particularly when it is perceived as unfair (Fehr and Rockenbach 2003). However, at least one study has shown that it is possible for frequent use of punishment to motivate high levels of cooperative

behavior in social exchange without provoking negative affect (Molm 1994). Because rewards may differ from punishment in theoretically important ways, I conduct a follow-up study (Study 2) in which participants have the option of enforcing norms through rewards, rather than punishments. In this chapter, I do not predict differences in behavior across reward and punishment-based sanctioning systems. Instead, my interest is in determining the extent to which the hypotheses are robust to different methods of norm enforcement.

Another possible criticism of study 1 is that punishment could only be used to enforce, not deter, compliance. This artificially limits individuals' range of actions compared to the real world, in which people are free to enforce norms that either proscribe or prescribe a given behavior. In time of war, for example, pro and anti-war factions may enforce compliance among their members with equal vigor. This limitation suggests an alternative explanation for the findings of Study 1. In the conflict condition, outgroup aggression harms participants, and presumably serves as an aversive stimuli. Participants may have wanted to respond to this aversive stimuli by taking some action to change the behavior of their group members. It is possible that participants chose to enforce norms of contribution because it was the only apparent avenue for influencing their group's members. If anti-conflict norms were available, would participants have enforced norms of peace instead? This possibility is examined in study 2.

The norm enforcement hypotheses tested in chapter 4 are similar to those tested in chapter 3, but address enforcement of reward-based norms.

H1a: Conflict with an outgroup will lead group members to reward contributors at greater rates.

H1b: Conflict with an outgroup will lead group members to reward contributors at greater rates, *to the extent that the outgroup behaves competitively.*

The hypotheses on contribution to the group and preference for a leader remain unchanged from Study 1.

H2a: Conflict with an outgroup will lead group members to contribute to the group at greater rates.

H2b: Conflict with an outgroup will lead group members to contribute to the group at greater rates, *to the extent that the outgroup behaves competitively.*

H3a: Conflict with an outgroup will lead group members to create leadership hierarchies at greater rates.

H3b: Conflict with an outgroup will lead group members to create leadership hierarchies at greater rates, *to the extent that the outgroup behaves competitively.*

As in chapter 3, the “b” hypotheses predict an interaction between structural conflict and contributions from the outgroup, such that groups in conflict develop norms of contribution when the outgroup members contribute to their group, but not when the outgroup is passive.

Study Overview

Study 2 was largely similar to study 1, with several important changes. In both studies, participants are randomly assigned to participate in either a two-group, three-person-per-group IPD with sanctioning (conflict condition), or a three-person PD with sanctioning (no conflict condition). As in study 1, both the conflict and the no conflict condition included two groups, and individuals could observe the contribution and sanctioning decisions of the other group via their computer monitors. Also as in study 1, the two groups were only interdependent in the conflict condition. Study 2 differed

from study 1 in that participants could use rewards instead of punishments to enforce cooperation, and they could enforce pro-conflict or anti-conflict norms. Participants could also decide the amount of the reward, and the magnitude of the rewards increased.¹⁶ These changes are discussed in more detail below.

METHOD

Participants

Participants consisted of 204 paid undergraduate volunteers (112 women and 92 men) who were recruited via flyers offering payment for participation in an experimental study.¹⁷ There were 102 participants each in the conflict condition and no conflict conditions (each condition included 17 sessions, with two three-person groups in each session).

Procedure

The procedure was similar to that of study 2, but incorporated several changes. Instead of being able to enforce norms through punishments, participants could enforce norms through rewards, using the following procedure. Participants received a ten-point endowment in a “reward account” at the beginning of every round. After participants decided whether to invest their team account on a given round, they observed the investment decisions of their other team members (team members were identified by numbers only to protect confidentiality).

¹⁶ As discussed above, using rewards instead of punishments and allowing participants to enforce any behavior address potential criticisms of study 1. Allowing participants to decide the amount of the reward and increasing the magnitude of the sanctions bring the study in line with prior studies used to examine reward-based norm enforcement (e.g. Horne 2004). This has the advantage of making the study more comparable to other studies of reward-based norm enforcement, though it reduces the extent that study 2 is directly comparable with study 1. However, the differences between study 1 and 2 mean that similar findings across both studies would indicate a more robust relationship between conflict and norm enforcement that if the two studies were more similar.

¹⁷ As in study 1, no gender effects were predicted or found. Gender did not exert a main effect on any of the behavioral variables, nor was there a significant interaction effect of gender with conflict on any of the behavioral variables.

Participants could then send points from their reward accounts to other team members. Points sent to team members were doubled, so that sending five points to a team member cost the sender five points, but earned the recipient ten points. Participants were free to reward in any amount up to their ten point total per round, and could reward other participants regardless of whether they had contributed. Participants could therefore enforce norms of contribution, enforce norms of non-contribution, or not enforce norms. The norm enforcement procedure was adapted from a series of studies by Horne and colleagues (Horne 2001; Horne and Cutlip 2002; Horne 2004).¹⁸ Because the rewarding procedure lasted longer than the punishment procedure of study 1, study 2 lasted for 10 rounds to ensure that the studies could be completed within the available time frame. Following the study, participants filled out a survey that included demographic questions such as gender and year in school. Participants were then debriefed, paid, and thanked for their participation.

DATA AND MODEL

As in study 1, analysis of the experimental data focused on three dependent measures: norm enforcement, contribution to the group, and preference for a leader. These dependent measures are analyzed with four primary independent variables: the presence of conflict, the level of contribution from the outgroup, the level of contribution from the ingroup, and the level of sanctioning from the ingroup. The following paragraphs discuss the dependent and independent measures in more detail.

Dependent Measures

With the exception of norm enforcement, all dependent measures were identical to those used in study 1. Because participants in the study could enforce norms of either contribution or non-contribution, I created a single measure of norm

¹⁸ Note that because study participants are identified by numbers, they can observe which participants contributed to the group, and which participants rewarded them. The study thus falls within the scope of the signaling theory described in chapter 2.

enforcement by subtracting the number of points sent to non-contributors from the number of points sent to contributors. Greater values on this measure indicate greater levels of resources invested in *enforcing norms of contribution to the group*. These data were collected for every participant on every round of the study, resulting in 2,040 observations.

Independent Variables

The independent variables are identical to those used in Study 1, except that ingroup norm enforcement now measures the number of points invested in rewarding contribution on each round.

Data Structure and Model

The experiment produces a data set consisting of 2,040 participant-rounds, in which rounds of the study are nested within participants, participants are nested within teams, and teams are nested within sessions. Because this data structure is nearly identical to that analyzed in study 1, I use a similar multilevel modeling strategy. The only difference is that, because the variable for norm enforcement is now continuous rather than an indicator variable, and so the logit link function is no longer used to model this outcome.

RESULTS

Bivariate Relations

To begin analyzing the results of the study, To begin evaluating the findings, I first compare the amounts participants invested in to enforcing norms, and the proportions of participants choosing to contribute to the group and elect a leader. The bivariate relations are shown in Table 4.1.

Table 4.1: Means or Proportions for Norm Enforcement, Contribution, and Preference for a Leader, by Conflict Condition

	No Conflict	Conflict	Total
Norm Enforcement	4.51 (3.38)	3.79 (3.06)	4.15 (3.24)
Contribution	0.75 (0.31)	0.74 (0.29)	0.75 (0.30)
Leader	0.60 (0.49)	0.52 (0.50)	0.56 (0.50)

Note: SDs in parentheses. N = 204 (102 per condition).

The results do appear not support the prediction that norm enforcement, contribution to the group, and preference for a leader will increase when two groups experience structural conflict. Reading across the row labeled “Norm Enforcement” (row 1), one can see that participants invested an average of 4.51 points in enforcing contribution in the no conflict condition (column 1, labeled “No Conflict”), and an average of 3.79 points per round enforcing contribution in the conflict condition (column 2, labeled “Conflict”). The fact that the norm enforcement amounts are positive indicates that participants tended to reward contribution, rather than non-contribution. In fact, they did so overwhelmingly. Of 204 participants in study 2, only 6 tended to enforce non-contribution more than contribution over the course of the study.¹⁹

Table 4.1 also shows that participants, on average, contributed to the group on about three out of every four rounds in both the no conflict and conflict conditions (row 3, “Contribution”). Finally, the descriptive statistics also show that approximately 0.60 of participants voted to elect a leader in the conflict condition,

¹⁹ In other words, only 6 participants had negative scores, averaged across all rounds, for the rewarding contributors measure of norm enforcement. Interestingly, in the post-experimental interview, some of these participants volunteered that they rewarded non-contributors in an effort to integrate them into the group, with the hope that this would encourage them to contribute more. It thus appears that some rewarding of non-contribution actually occurred in an effort to increase contribution.

while 0.52 of participants voted to elect a leader in the no-conflict condition. None of the differences in means are statistically significant.²⁰

Just as in study 1, the descriptive statistics indicate that structural conflict may not be sufficient to motivate enforcement of cooperation, contribution to the group, or the emergence of leadership hierarchies. This analysis, however, does not tell us whether responses to conflict might be conditioned on the behavior of ingroup and outgroup members. Are participants in conflict more likely to enforce norms if the outgroup actively contributes, thereby harming the outgroup? I address this question using multivariate models in the next section.

Multivariate Models

In this section, I present an analysis of the experimental data that systematically considers the effects of conflict, outgroup contributions, ingroup contributions, and ingroup norm enforcement, using the multilevel model described above. I also control for the effects of round and the interaction of round and conflict, to account for the fact that participants' behavior might change over time (Bornstein, Winter, and Goren 1996; Rapoport and Chammah 1965).

Norm Enforcement

Table 4.2 presents the coefficients from a series of models estimating the effects of conflict and the standardized behavioral variables on norm enforcement. The first two models focus on the main effect of conflict: model 1 includes the conflict variable only, model 2 adds the effects of round and the interaction of round and conflict. Models 3-5 separately add the main effects and interactions with conflict of ingroup contribution, ingroup norm enforcement and outgroup contribution,

²⁰ Tested using t-tests of the data collapsed at the session level.

Table 4.2: Multilevel Model of Norm Enforcement

Fixed Effects	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9
Conflict	-0.72 (0.70)	-0.21 (0.70)	-0.22 (0.76)	-0.30 (0.61)	-0.40 (0.77)	-0.22 (0.62)	-0.33 (0.77)	-0.39 (0.62)	-0.31 (0.63)
Ingroup Contribution			0.19 (0.13)			-0.11 (0.14)	0.19 (0.13)		-0.10 (0.14)
Ingroup Norm Enf.				1.07+ (0.17)		1.11** (0.19)		1.05** (0.17)	1.09** (0.19)
Outgroup Contribution					0.06 (0.14)		0.07 (0.14)	0.06 (0.13)	0.05 (0.13)
Round		0.07** (0.027)	0.06+ (0.03)	0.04 (0.03)	0.06+ (0.03)	0.03 (0.03)	0.06* (0.03)	0.04 (0.03)	0.03 (0.03)
Conflict X Ingroup C			-0.39* (0.17)			-0.25 (0.18)	-0.43* (0.17)		-0.27 (0.18)
Conflict X Ingroup NE				-0.54 (0.21)		-0.43+ (0.23)		-0.57* (0.21)	-0.46* (0.23)
Conflict X Outgroup C					0.35+ (0.19)		0.36+ (0.19)	0.28 (0.18)	0.31+ (0.18)
Conflict X Round		-0.09* (0.038)	-0.09* (0.04)	-0.05 (0.04)	-0.06 (0.04)	-0.07 (0.04)	-0.08+ (0.04)	-0.04 (0.04)	-0.05 (0.04)
Constant	4.51** (0.50)	4.12** (0.50)	4.20** (0.54)	4.24** (0.43)	4.22** (0.55)	4.25** (0.44)	4.19** (0.54)	4.23** (0.44)	4.25** (0.45)
Random Effects									
Intercept-1 (u_{0jkl})	2.32** (0.36)	2.33** (0.36)	2.31** (0.37*)	3.04** (0.46)	2.42** (0.37)	2.94** (0.45)	2.29** (0.36)	3.00** (0.46)	2.90** (0.44)
Intercept-2 (σ_{00kl})	7.36** (1.43)	7.36** (1.43)	7.59** (1.51)	3.90** (1.01)	7.90** (1.53)	4.12** (1.05)	7.83** (1.55)	4.16** (1.06)	4.41** (1.10)
Intercept-3 (ε_{000l})	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)

Note: Standard errors in parentheses. N = 1,836 observations for all models. The model includes 10 level one units, 204 level two units, and 68 level three units, and 34 level 4 units.

+ Significant at 10% level; * significant at 5% level; ** significant at 1% level

respectively. Models 6-8 include the behavioral variables and their interactions with conflict in pairs: model 6 includes ingroup contribution and ingroup norm enforcement, model 7 includes ingroup contribution and outgroup contribution, and model 8 includes ingroup norm enforcement and outgroup contribution. Model 9, the full model, includes all of the variables.

I first evaluate hypothesis H1a, the prediction that participants would be more likely to reward contribution to the group in the conflict condition, relative to the no conflict condition. This hypothesis can be evaluated by examining the main effect of conflict, in row 1 (labeled “Conflict” of Table 4.2). Reading across row 1, one can see that this hypothesis is not supported by any of the models. In each model, the main effect of conflict is not statistically significant, and also is not in the predicted direction. This suggests that structural conflict between groups is not sufficient to promote the emergence of reward-based norms.

I now turn to assessing hypothesis H1b. Do people reward contribution in conflict to the extent that outgroup contribute more heavily? We can evaluate this prediction by examining row 8 in Table 4.2 (labeled “Conflict X Outgroup C”), which includes the coefficients for the interaction of conflict with outgroup contributions. Positive values of this coefficient indicate that participants increasingly reward contribution in the conflict condition to the extent that members of the outgroup contribute to the conflict.

Table 4.2 indicates that this hypothesis is supported by the data. Participants in the conflict condition invested more resources in enforcing norms of contribution when members of the outgroup contributed to the conflict, but not when the outgroup did not contribute. This finding is shown by a positive, marginally significant interaction effect of conflict and outgroup contributions on the amount of resources invested in supporting norms of contribution among ingroup members. The

significance of this effect holds across the models, except model 8. The effect is present in the full model (model 9), and so appears to be generally robust.

The coefficient also differs little in magnitude across the model. In the full model, the coefficient is 0.31, meaning that, in the conflict condition, approximately 0.31 more points are invested in norm enforcement for every standard-deviation increase in the number of outgroup members contributing. With a baseline level of contribution of approximately 4.25, this suggests that each standard deviation increase in contribution by outgroup members increases norm enforcement by approximately 7.2% of baseline levels. This supports hypothesis H1b, suggesting that conflict increases norm enforcement, but only when the outgroup behaves aggressively. As expected, outgroup behavior had no effect on norm enforcement in the no conflict condition.

Some additional findings, while not theoretically predicted, are also worth noting. In particular, norm enforcement by team members tends to increase the amount individuals invest in rewarding contribution. This trend is present in both conditions, but is stronger in the no conflict condition. This finding is shown in row 4 (“Ingroup Norm Enf.”) of Table 4.2, which shows the coefficients for the main effect of ingroup norm enforcement. The coefficient for this variable is positive and significant across all models, indicating that participants tend to reward contribution to a greater extent when their group members invested in enforcing norms on the prior round. The slope of 1.09 in the full model, for example, indicates that participants invested an average of 1.09 more points in norm enforcement for every standard deviation increase in resources invested in norm enforcement by their group members.

This effect is qualified by a negative and significant interaction effect of conflict with ingroup norm enforcement. This effect is approximately half the magnitude of the main effect of norm enforcement, indicating that the effect of

ingroup norm enforcement on the amount participants invested in enforcing norms was present, though weaker, in the conflict condition. The implications of this finding will be considered in the discussion section.

Study 1 found that punishment of defectors is largely conducted by high contributors. This is consistent with earlier research, in addition, study 1 provided novel evidence that high contributors are also more responsive to conflict. Research has not examined whether high and low contributors differ in the extent to which they apply positive sanctions. An analysis of the data from study 2 found no major substantive in the behavior of high and low contributors on key theoretical variables. This is not surprising. In study 1, low contributors who chose to punish would be punishing themselves, whereas low contributors in study 2 cannot be harmed by rewarding others, and may indirectly benefit from doing so if others reciprocate the rewards.

Contribution to the Group

While the present study focuses on the emergence of norms and hierarchies, it will also allow us to examine the effect of conflict on contribution to the group. Earlier studies have found that people are more willing to make costly contributions to their groups when they are in conflict with an outgroup (Bornstein 2003). However, as discussed in chapter 3, prior work does not examine contribution to the group in conflict when a system of norm enforcement is also available.

Table 4.3 allows for an examination of the two hypotheses regarding contribution: H2a, which predicts that conflict alone will motivate contribution, and H2b, which predicts that conflict will motivate contribution to the extent that the outgroup contributes to conflict. The models in Table 4.3 are identical to those in Table 4.2, with the exception that contribution to the group is the dependent variable.

Because contribution is a binary variable, I estimate the models in GLLAMM using a logit link function.

Hypothesis H2a can be evaluated by examining row 1 (“Conflict”) of Table 4.3. This row shows the coefficients for the main effect of conflict. Table 4.3 reveals that H2a is not supported by the data: the main effect of conflict does not reach statistical significance in any of the models. Furthermore, the direction of the coefficient fluctuates between positive and negative across the models, indicating that the data does not trend in the direction of the hypotheses. As in the case of norm enforcement, this suggests that structural conflict between groups is not sufficient to motivate contribution to the group, net of the behavior of the participants in the conflict. In addition, this finding closely matches the results of study 1.

In the analysis of the norm enforcement data, I found that the effects of conflict are contingent on competitive behavior from the outgroup. Does a similar pattern emerge for contribution? We can begin to answer this question by examining row 8 (“Conflict X Outgroup C”) in Table 4.3, which includes the coefficients for the interaction of conflict with outgroup contribution.

The findings show that there is some support for H2b, although it is weaker than in the case of norm enforcement. The conflict x outgroup contribution interaction effect, which measures the extent to which participants in the conflict condition enforce norms to a greater extent as the outgroup behaves more aggressively, is in the predicted direction and marginally significant in some models (models 5 and 8). However, the coefficient moves slightly outside of the range of significance in models 7 and 9, which control for the effects of ingroup contribution. While there is some evidence that outgroup contribution spurs ingroup contribution in the conflict.

Table 4.3: Multilevel Model of Contribution

Fixed Effects	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9
Conflict	-0.26 (0.54)	0.02 (0.60)	-0.16 (0.57)	-0.13 (0.55)	-0.22 (0.64)	-0.15 (0.53)	0.21 (0.57)	-0.19 (0.55)	-0.20 (0.52)
Ingroup Contribution			0.55** (0.15)			0.45** (0.16)	0.54** (0.15)		0.44** (0.16)
Ingroup Norm Enf.				0.63** (0.21)		0.44+ (0.23)		0.65** (0.21)	0.47* (0.23)
Outgroup Contribution					-0.23 (0.17)		-0.22 (0.17)	0.26 (0.17)	-0.24 (0.17)
Round		-0.02 (0.03)	0.00 (0.04)	-0.01 (0.04)	-0.01 (0.04)	0.00 (0.04)	0.00 (0.04)	-0.01 (0.04)	0.00 (0.04)
Conflict X Ingroup C			-0.24 (0.19)			-0.19 (0.20)	-0.24 (0.19)		-0.17 (0.20)
Conflict X Ingroup NE				-0.37 (0.24)		-0.23 (0.27)		-0.40 (0.24)	-0.26 (.27)
Conflict X Outgroup C					0.40+ (0.23)		0.34 (0.23)	0.38+ (0.22)	0.32 (0.22)
Conflict X Round		-0.05 (0.05)	-0.02 (0.06)	-0.02 (0.05)	-0.02 (0.06)	-0.02 (0.06)	-0.02 (0.06)	-0.02 (0.06)	-0.01 (0.06)
Constant	2.07** (0.39)	2.16** (0.43)	2.06** (0.42)	2.05** (0.41)	2.17** (0.47)	1.99** (0.39)	2.07** (0.42)	2.05** (0.40)	2.00** (0.38)
Random Effects									
Intercept-1 (u_{0jkl})	2.55** (0.56)	2.58** (0.57)	2.97** (0.65)	2.32** (0.53)	2.33** (0.54)	2.84** (0.62)	2.97** (0.65)	2.32** (0.52)	2.82** (0.61)
Intercept-2 (σ_{00kl})	3.34** (0.97)	3.36** (0.98)	1.99* (0.82)	1.88* (0.79)	3.54** (1.02)	1.14 (0.72)	1.98* (0.83)	1.79* (0.77)	1.07 (0.71)
Intercept-3 (ϵ_{000l})	0.00 (0.00)								

Note: Standard errors in parentheses. N = 1,836 observations for all models. The model includes 10 level one units, 204 level two units, and 68 level three units, and 34 level 4 units.

+ Significant at 10% level; * significant at 5% level; ** significant at 1% level

condition, it generally appears to be weaker and less reliable than the interaction effect of conflict and outgroup contribution on norm enforcement.

As in the model of norm enforcement, some unpredicted effects appear in the results for contribution. Ingroup contribution and ingroup enforcement of contribution both increase participants' contribution to the group. These effects can be seen in rows 2 and 3, respectively, of Table 4.3, and will be considered further in the discussion section.

Preference for a Leader

I now turn to examining the effects of conflict on preferences for a leader. Does intergroup conflict increase willingness to relinquish one's choices to a leader? Because the leader variable was collected at the close of the study, I use individual-level data, rather than round-level data, to test this hypothesis. In addition, instead of using lagged measures of outgroup contribution, ingroup contribution, and ingroup norm enforcement, I use the standardized average levels of these variables over the course of the study as covariates, as well as their interaction with conflict.

Table 4.4 shows the multilevel logit model coefficients for the regression of leader preference on conflict and the behavioral variables. The table includes eight models. Model 1 includes only the main effect of conflict. Models 2-4 include the main effect of conflict, and the main effects and interactions with conflict of ingroup contribution, ingroup norm enforcement, and outgroup contribution, respectively. Models 5-7 include the behavioral variables and their interactions with conflict in pairs: model 5 includes ingroup contribution and ingroup norm enforcement, model 6 includes ingroup contribution and outgroup contribution, and model 7 includes ingroup norm enforcement and outgroup contribution. Model 8 includes all of the variables.

Table 4.4: Multilevel Model of Preference for a Leader

Fixed Effects	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
Conflict	-0.32 (0.28)	-0.33 (0.29)	-0.40 (0.29)	-0.33 (0.29)	-0.37 (0.29)	-0.35 (0.29)	-0.41 (0.29)	-0.37 (0.29)
Ingroup Contribution		-0.18 (0.19)			0.22 (0.33)	-0.23 (0.19)		0.10 (0.34)
Ingroup Norm Enf.			-0.30 (0.19)		-0.48 (0.33)		-0.32 (0.19)	-0.40 (0.34)
Outgroup Contribution				-0.24 (0.19)		-0.29 (0.19)	-0.25 (0.19)	-0.23 (0.20)
Conflict X Ingroup C		-0.27 (.32)			-0.79 (0.49)	-0.17 (0.31)		-0.66 (0.50)
Conflict X Ingroup NE			0.06 (0.29)	0.50 (0.31)	0.66 (0.48)		0.10 (0.30)	0.60 (0.49)
Conflict X Outgroup C						0.47 (0.32)	0.47 (0.32)	0.43 (0.32)
Constant	0.39* (0.20)	0.41* (0.21)	0.44* (0.21)	0.41* (0.20)	0.46* (0.21)	0.43* (0.21)	0.46* (0.21)	0.47* (0.21)
Random Effects								
Intercept-1 (u_{0jkl})	0.00 (0.00)	0.04 (0.32)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
Intercept-2 (σ_{00kl})	0.00 (0.00)							

Note: Standard errors in parentheses. N = 204 observations for all models. The model includes 204 level one units, and 68 level two units, and 34 level 3 units.

+ Significant at 10% level; * significant at 5% level; ** significant at 1% level

The first hierarchy hypothesis, H3a, predicts that conflict will motivate the election of group leaders. This hypothesis can be evaluated by examining row 1 (“Conflict”) of Table 4.4, which includes the coefficients for the main effect for the conflict. The results show that this hypothesis is not supported. The main effect of conflict is negative across all of the models, and does not reach statistical significance.

The second hierarchy hypothesis, H3b, predicts that conflict will interact with levels of outgroup contribution such that participants in the conflict condition will be more likely to elect leaders when the outgroup has contributed much than when the outgroup has contributed little. This hypothesis can be evaluated by examining row 6 (“Conflict X Avg. Outgroup C”) of Table 4.4. The results indicate are in the predicted

direction, but do not reach statistical significance. Other than the constant term, no other coefficients in the model are statistically significant. As in study 1, the results for high and low contributors did not differ greatly.

Discussion

Like Study 1, the results of Study 2 support the hypothesis that conflict leads people to enforce norms of contribution to the group to the extent that the outgroup contributes to conflict. Participants in the conflict condition, but not in the no-conflict condition, responded to past outgroup contributions by investing more resources in enforcing ingroup contribution. The results also showed a similar, though less reliable, interaction effect of conflict and outgroup aggression on contribution to the group. The results did not support the prediction that structural conflict – in the form of antagonistic group goals – would increase norm enforcement or contribution, absent outgroup aggression. The study also found no significant effects of conflict on group hierarchies, although the trends were in the predicted direction.

The results of Study 2 provide further support for Study 1's key finding – namely, that conflict, when coupled with outgroup aggression, can increase enforcement of social norms. The fact that similar results were found using a very different method of norm enforcement speaks to the robustness of the findings, and suggests they are not simply the result of the particular experimental design used in Study 1. The findings reinforce the conclusion that conflict can be an important source of social norms. Furthermore, it is noteworthy that participants overwhelmingly chose to enforce pro-conflict, rather than anti-conflict norms.

The lack of findings for the leader variable raise interesting questions, given that this effect was quite strong in Study 1. Why did a similar effect not obtain in Study 2? One possible explanation for this outcome is that levels of contribution during the study were quite high – approximately 75% in both conditions. Given that

prior IPD studies have found levels of cooperation closer to 50% (Bornstein 2003), it may be the case that participants felt that little more coordination of the group's resources was necessary or possible, and therefore saw no need for electing a leader. This explanation could be tested with a follow-up study in which participants are given systematically varying feedback, indicating either high or low levels of contribution from their team members.

Several unpredicted effects, which do not bear on the hypotheses, were also found. One finding was that levels of norm enforcement from ingroup members motivated participants' investment in norm enforcement in both conditions. This is similar to the findings in Study 1 that punishment by an actor's peers seems to motivate punishment by a that actor. However, in Study 1, the contagious effect of punishment appeared only in the conflict condition. In Study 2, peers' rewarding significantly increases an individuals rewarding in both the conflict and no conflict conditions, and the effect is stronger in the no conflict condition. One possibility is that participants rewarded those who rewarded them, exchanging "approval for approval" (Flache and Macy 1996:19). Indeed in the postexperimental interview, a number of participants reported that they rewarded both contributors and those who rewarded them. Such a pattern would not have been possible in Study 1, because participants were constrained to punish (or not) non-contributors, and could not selectively punish individuals. However, it is not clear why the magnitude of this effect was reduced by approximately half in the conflict condition. The question of when and how norm enforcement might be contagious, and how the "trading favors" aspect of rewarding might affect the social control aspect of rewarding, merits future investigation.

Additional findings showed that participants tended to contribute more to the group as their teammates enforced norms of contribution and as their teammates

contributed more. The former finding is intuitive: participants contributed more to the extent that group members were rewarded for contributing in the past. The latter finding, that participants contribute more when their peers do, is consistent with goal/expectation theory (Pruitt and Kimmel 1977), which argues that people actually prefer the R outcome in prisoner's dilemmas, and thus will cooperate when others cooperate.

In sum, Study 2 offered additional support for the finding that conflict with a competitive outgroup increased contribution. It also offered somewhat weaker support for the hypothesis that conflict with an competitive outgroup increases contribution to the group. Predictions that conflict with a competitive outgroup would motivate the creation of leadership hierarchies trended in the positive direction, but did not reach significance. In the discussion and conclusion, I discuss Studies 1 and 2 in relation to each other and their contribution to theory and research more generally.

CHAPTER 5

DISCUSSION

How does conflict between groups shape social relations within groups? In this dissertation, I argue that intergroup conflict can lead to the emergence of norms and hierarchies, causing groups to become both more strict and more centralized than they were prior to conflict. In Chapter 5, I briefly summarize the theory, methods, and findings of the dissertation. I then discuss the contributions of the dissertation, including its direct contributions to theory and research, as well as the broader implications of the work. I also consider limitations of the research before moving on to discuss a range of possible follow-up studies. I conclude with a brief reflection on the relevance of group conflict for shaping social life.

Summary

This dissertation seeks to make a theoretical and empirical contribution to understanding the relationship between intergroup conflict and the emergence of intragroup norms and hierarchies. To develop an explanation for the causal link between conflict, norms, and hierarchies, I show how theories of signaling, evolved preferences, and framing predict a link between conflict, norm enforcement, and in some cases, group hierarchies.

Research on signaling behavior supports the theory that people enforce norms in order to enhance their reputation and maintain positive relations with other group members (Barclay 2006; Horne 2004). In conflict, this motivation may increase, either because group identity becomes more salient (Horne and Chen 2005), or because selfish behavior is interpreted more negatively. Evolved preference accounts argue that humans have an intrinsic tendency to respond to conflict by enforcing norms. In support of this, neurological studies have shown that punishing defectors (de Quervain

et al 2004) and rewarding contributors (Rilling et al 2002), both activate regions of the brain associated with pleasure and satisfaction. In addition, the theory of stability-dependent cooperation explains why incentives for norm enforcement might be activated by intergroup conflict, without resorting to group selection accounts (Lahti and Weinstein 2005). Finally, tests of social psychological theories of framing have shown that, in conflict, people are more likely to value their groups' outcomes over their own, and to seek to maximize the difference in payoffs between their group and an outgroup (Bornstein and Ben-Yossef 1994). This mechanism predicts greater levels of norm enforcement and preference for a leader in conflict, as well as greater contribution to the group.

In addition to developing theoretical accounts for the link between conflict and norm enforcement, I also seek to provide a clearer explanation of what type of conflict is necessary to trigger norm enforcement (Coser 1956). I distinguish between structural conflict, in which groups have conflicting goals, and whether the outgroup actually pursues the goals and inflicts costs on the ingroup. I then test whether one or both of these conditions increase norm enforcement, contribution to the group, and leader preferences.

To develop a setting in which to test these ideas, I drew on a body of prior work conceptualizing conflict as an intergroup prisoner's dilemma (IPD, see Bornstein 2003 for a review). The incentive structure of the IPD corresponds to that of many real-life conflicts, in that contribution to the group is costly, and helps the ingroup while harming the outgroup. The IPD served as the conflict condition in both experimental studies conducted for the dissertation. For the control condition, I employed a setting in which contribution did not affect the outgroup, but was otherwise identical. In both conditions, participants were given the option of enforcing costly norms and electing leaders. Adding these options to the IPD

extended the setting in a novel way. Previous studies have used a single group public goods dilemma to study costly norm enforcement (Barclay 2006; Horne 2001; Horne and Cutlip 2002; Horne 2004; Yamagishi 1986) and leader preferences (Messick et al. 1983; Samuelson et al. 1984; Rutte and Wilke 1985; Samuelson and Messick 1986), but these are the first studies to do so in an IPD context.

Study 1 examined the effect of intergroup conflict on the emergence of punishment-based norms. The study found that, in the conflict condition, participants enforced norms to the extent that the outgroup contributed to the conflict. This effect was largely driven by frequent cooperators, who were highly responsive to outgroup behavior in the conflict condition. Study 2 implemented a reward-based sanctioning system, and allowed participants to reward non-contribution as well as contribution. Study 2 found a positive interaction effect of conflict and outgroup contribution on norm enforcement, reinforcing the findings of Study 1.

Study 1 also found that conflict can cause groups to create leadership hierarchies. As in the case of norm enforcement, this effect is contingent on the competitive behavior from the outgroup, although the effect was not driven by high contributors. This effect is quite strong in the Study 1, but does not reach significance in Study 2. The much weaker effect in Study 2 may be due to the higher rates of contribution – given that groups were more successful, they may have felt less need to respond to outgroup competition by electing a leader.

Study 1 found no evidence that conflict increases contribution to the group. This trend appeared in Study 2, though only weakly. This contrasts with earlier work (Bornstein 2003), which found that conflict increases contribution to the group. This difference might be explained by either the availability of sanctioning in the present studies, which can undermine cooperation (Fehr and Rockenbach 2003), or structural or cultural differences in contribution or enforcement behavior between the U.S.,

where the present studies were conducted, and Israel, where most prior IPD studies have been conducted.

Discussion

The primary contribution of this dissertation is to develop and systematically test theoretical accounts predicting a relationship between group conflict and norm enforcement. The results show that people respond to a threat from an outgroup by enforcing stricter codes of conduct within their group, which require ingroup members to commit a greater proportion of their resources to the conflict. This finding is consistent across two studies employing different procedures for norm enforcement, suggesting that the findings are robust to the method of norm enforcement. The results also show that, at least under some conditions, conflict can cause groups to create leadership hierarchies, at the expense of group members' autonomy. The findings contribute to two fundamental problems in the social sciences: the relationship between intergroup conflict and intragroup cohesion, and the emergence of social norms.

The question of whether conflict increases group cohesion is a classic problem in sociology and related disciplines (Sumner 1960 [1906]; Simmer 1955 [1908]; Coser 1956; Stein 1971). Previous research has shown that conflict tends to increase the extent to which people contribute to their groups (e.g., Bornstein 2003). However, little is known about whether conflict will motivate people to *enforce* contribution, especially when doing so is costly. Field studies suggest that a correlation between conflict and normative control exists (e.g., Black-Michaud 1975; Coben 1964; Gibson 1988; Gould 2003; Stein 1971) but this is the first study to show direct, causal evidence for the link between conflict and costly norm enforcement. Study 1 also shows that conflict can motivate the emergence of group hierarchies. These findings

indicate that conflict shapes behavioral dimensions of group cohesion beyond contribution to the group.

More broadly, the present study shows how conflict can alter the structure of the group itself, causing it become more complex and internally differentiated. Conflict may thus serve as a point of departure for the emergence of more complex groups or even simple organizational forms. The latter point has been suggested by organizations scholars studying the effect of external threat on the emergence of hierarchies in firms (Staw et al 1981), and the present results lend weight to that hypothesis.

The results also offer insight into the emergence of social norms. The origin of social norms is a fundamental question in the social sciences because norms shape a diverse array of human behaviors (Hechter and Dieter-Opp 2001). Norms structured human societies long before formal codes of law (de Quervain et al 2004), and often supercede laws in practice (Ellickson 1994). Explaining norms is a puzzle, however, because enforcing norms requires that someone pay the cost of monitoring the group and distinguishing between norm-followers and norm-breakers (Hechter 1987), as well as administering sanctions (Oliver 1980). The costs of administering sanctions can include time, effort, or material resources. The costs can also be relational – people may avoid enforcing norms among their friends, for fear of damaging the friendship (Flache and Macy 1996). Because the benefits of enforcing norms to the individual are usually outweighed by the costs (Oliver 1980), it is not clear how norms could arise. Efforts to answer this question have produced an array of formal and computational models (e.g. Axelrod 1986; Bendor and Swistak 2001; Boyd 2003 et al; Centola, Macy, and Willer 2005; Coleman 1990; Heckathorn 1988, 1989; Macy 1993), as well as empirical studies (e.g. Barclay 2006; Horne 2004; Fehr and Gächter 2002; Yamagishi 1986, 1988).

However, scholars have yet to empirically examine the effect of intergroup conflict on costly norm enforcement (Fehr and Fischbacher 2003: 6). Pro-conflict norms are especially difficult to explain, because, in addition to being costly for the enforcer, they mandate behaviors that are both individually *and* collectively costly. The present study shows that intergroup conflict is one factor in the emergence of social norms. Furthermore, given the prevalence of intergroup conflict in early human societies (Bowles 2006), conflict may have been a key source of the emergence of social structures historically.

A similar argument underlies Tilly's (1989) account of the development of modern European states, which he views as driven by efforts to mobilize for war. To survive, states needed to create, maintain, and fund a military. This created a coordination problem, namely, how to ensure that group resources were directed towards conflict, despite individual incentives to divert those resources towards other ends. Emerging European states solved this problem by developing courts and tax collection systems to sanction non-contributors, as well as establishing more hierarchical systems of government in the form of centralized state bureaucracies. These innovations echo the patterns of norm-enforcement and hierarchy-building found in the present studies. However, determining whether this resemblance is more than superficial will require further investigation.

Broader Implications

In addition to the present work's contribution to the literatures on group conflict and cohesion and the origin of social norms, the work has broader implications for understanding the role of conflict in social life. For example, the results may help us explain the apparent durability of many conflicts. The findings showed that (1) conflict motivates enforcement of contribution, and that (2) enforcement motivates future contributions to the group. This suggests that conflict

and social norms may become mutually reinforcing, if conflict motivates the emergence of pro-conflict norms, and such norms motivate further conflict. This may create a positive feedback loop that could help explain why conflict sometimes appears to spiral out of control, and why, once out of control, conflict can be so difficult to extinguish.

The results also have implications for conflict resolution. The finding that conflict can change relations within groups in ways that make future conflict more likely suggests that attempts at conflict resolution may fail if they are directed only at changing relations between groups. By the time an intervention is attempted, conflict may have already led to the development of intra-group norms mandating that group members contribute to the conflict, and centralized hierarchies that facilitate mobilization. Parties interested in reducing conflict between groups may need to first defuse conflict-producing mechanisms within groups.

One approach could be to make it more difficult or costly for groups to enforce pro-conflict norms. The ability of groups to enforce norms depends, in part, on their members' dependence on the group (Hechter 1987; Homans [1950] 1992). Individuals who depend on the group for resources, protection, or approval are more likely to acquiescence to the group's demands than those who are not dependent. Decreasing members' dependence on the group could weaken the group's ability to enforce pro-conflict norms. For example, individuals who join radical militant groups might quickly become dependent on those groups, because they face prosecution or other sanctions if they are captured. Providing amnesty for at least some group members might make them more able to exit the group, and thus undermine the group's ability to extract costly contributions from their members.

On a theoretical level, the results caution against reifying groups in conflict. Observers often perceive groups in conflict as entities capable of acting with a single

mind, assuming that each side makes unanimous decisions about how best to defeat its enemy or ensure its survival. Axelrod's (1984) discussion of trench warfare during World War I, for example, portrays British and German units as single participants in an iterated prisoner's dilemma. Each unit, he argues, collectively decides "to shoot to kill or deliberately shoot to avoid causing damage (Axelrod 1984:75)." This belies the fact that actors within a group may have very different ideas about what goals their group should pursue and how it should pursue them.

Furthermore, even when group members agree on a goal, this does not guarantee that all will contribute to pursuing it (Olson 1965). Groups can act as unitary entities only insofar as they solve internal problems of collective action. A group can operate cohesively in conflict if it can enforce norms of behavior among its members or if its members willingly defer to the group's goals. This is true for antagonism as well as conciliation. Convincing displays of aggression require a majority of group members to display a readiness to fight; convincing displays of pacifism can be quickly undermined by a few uncompromising group members. As shown in the present studies, there is substantial variance in the extent to which members are willing to contribute and enforce norms of contribution, and their willingness to do so may be affected by the behavior of their own and other group members.

Limitations

One limitation of the study is the fact that it uses a sample of undergraduate participants. As with any research using a non-random sample, it is unclear to what extent the results will generalize to other populations. This may or may not limit the generality of the results, depending on whether the present sample responds to the independent measures differently than a general population sample. The theoretical mechanisms developed here are based on fundamental aspects of human psychology,

such as the desire to have positive interactions with peers or to see perceived injustices punished. We therefore do not have theoretical reasons to expect the results to vary widely across populations. Nevertheless, it would be useful to test this assumption.

While the sample may limit the external validity of the findings, the goal of this research is not to prove that conflict and cohesion are correlated in the real world. Existing field studies have already shown a correlation between conflict, norms and hierarchies (Black-Michaud 1975; Gould 2003). Instead, the goal of the research is to test whether conflict plays a *causal* role in the emergence of norms. To test the causal link between conflict and norm enforcement, it is necessary to employ an experimental design in which groups are randomly assigned to be in conflict with another group or not. Only by experimentally manipulating the level of conflict can we be sure that conflict produces changes in group structure, rather than the other way around.

Given that the goals of the research require an experimental design, there are practical reasons for using an undergraduate sample. It would be prohibitively costly to select a random sample of participants from a population of any meaningful size and transport them to the laboratory. For obvious reasons, systematically creating conflict in existing real-world groups is also not an option. Thus, in order to test the hypotheses in a controlled setting, I must rely on this more limited sample. Some of the future work described in the next section will address ways to improve upon the limitations of this sample.

An additional limitation is the fact that the present studies use small groups exclusively. The consistent patterns of norm enforcement lend confidence that the results are robust for small groups, but we do not know whether or how the results will scale up to larger groups. Examining this possibility will require further theory development and testing. I suggest some directions this work might take in the “future research” section, below.

Additionally, because the study uses the intergroup prisoner's dilemma to model conflict, the scope of the conclusions are limited, pending further research, to conflicts that conform to the IPD's assumption. The IPD was chosen because it is considered to be the model that most realistically accounts for the incentive structure of intergroup conflict, and thereby has the broadest applicability (Bornstein 2003). Nevertheless, other models of conflict exist, and might produce different results. I discuss this issue further in the "future research" section.

A final limitation is the fact that, while the hypotheses examined the interaction of conflict and outgroup behavior, only the level of conflict was manipulated experimentally. For the measure of outgroup behavior, I used levels of outgroup contribution. This approach has the advantage of allowing us to measure how actual groups behave in conflict. However, the manipulation of outgroup conflict is not experimental because individuals are not randomly assigned to interact with competitive or cooperative outgroups. In the current setting, this is unlikely to present a problem for assuming a causal effect, because participants do not voluntarily select the type of outgroup they interact with, and so the results would likely be very similar if the level of outgroup contribution was manipulated.

Alternatively, I could have manipulated the level of outgroup aggression by having study participants interact with simulated partners, with the outgroup programmed to contribute either a low or high amount. This would bolster support for the causal effect of outgroup contribution. I decided against this approach in the initial studies for two reasons. First, it is useful to know how live groups actually behave when participating in the IPD with sanctioning, as there are no prior studies of this type. Second, simulating partners in the IPD would require a number of arbitrary assumptions, such as how much simulated ingroup members should contribute and enforce, how they should respond to the participant's behavior, and how the outgroup

should respond to the ingroup's behavior. Without previous studies in this area, it would be difficult to have confidence in whatever assumptions I made to answer these questions.

Having now conducted an initial set of studies using the IPD with sanctioning, it may make sense to shift to using simulated partners in future work. This would provide the added advantage of making individuals an independent unit of analysis, by avoiding of clustering of standard errors at the group and session levels. The resulting gain in statistical power would facilitate adding additional conditions.

Future Research

The present work suggests a number of directions for future research. In general, these directions include addressing limitations of the current studies, further refining and testing the mechanisms underlying the theories presented here, and pushing the theory in new directions.

Addressing Limitations

As noted, one limitation of the present studies is that they use an undergraduate sample. While it is impractical to transport a sample of the general population into the lab, it may be possible to partially address this concern using data collected via the internet. As the experimental setting is already computer-based, it would be possible, given sufficient resources, to create a similar setting on-line and recruit large, diverse samples of participants. I have recently begun collaborating on a project to study intergroup conflict and conflict resolution on-line, which I hope will lead to online studies of conflict and norm enforcement. The online settings do pose unique risks and challenges, such as verifying the demographic information of participants and maintaining a controlled setting. For this reason, they are unlikely to replace laboratory experiments, at least in the near future. However, online studies could provide a useful source of additional data.

Online studies may also address another limitation of the present study by allowing for much larger sample sizes. This would help determine the extent to which the findings for small groups obtained in the present studies scale up to larger groups or organizations. However, it is also important to develop theory about the ways in which large groups may differ from small groups, and the implications of these differences for conflict. While a number of variables are likely to be theoretically relevant, here I briefly discuss one difference – internal heterogeneity – that may determine whether large groups respond to conflict similarly to small groups.

Large groups likely differ from small groups in their relative degree of internal differentiation. It is probably safe to assume that small groups typically have a homogenous internal structure, while larger groups tend to be comparatively more heterogeneous. This is true along a number of dimensions, but is particularly relevant in the case of individual and group preferences. In small groups, preferences may differ widely in magnitude, but there is little difference in the direction of preferences. In contrast, large groups are comprised of subgroups that vary widely in both the direction and magnitude of their preferences. This distinction has important implications for understanding when conflict may make states more or less cohesive and centralized.

I suggest that when large groups resemble small groups on this dimension – when they are internally homogenous – then conflict with an external rival will tend to make them more cohesive and centralized. As heterogeneity increases, conflict becomes increasingly likely to have disintegrating, rather than integrating, effects on states. In other words, I expect that conflict and heterogeneity negatively interact when degree of internal cohesion is the dependent measure. In large heterogeneous groups, cohesion and centralization are unlikely to result from conflict because the interests of subgroups will differ. The subgroups will therefore likely differ in their

estimation of the best response to conflict, which may produce tension between the groups, rather than cohesion. This argument could be tested in an online setting by assigning participants a distribution of preferences. It might also make sense to develop this idea theoretically by using an agent-based model to explore the effects of different distributions of preferences on outcomes at the group level.

It would also be informative to compare the results for the IPD to other models of conflict. One particularly interesting model for comparison is the intergroup public goods dilemma (IPG, Bornstein 1992). An IPG is similar to an IPD, but has a winner-take-all incentive structure. In the IPG, the team that invests more in conflict on a given round wins the entire public good, rather than a larger share, as in the IPD. As a result, in the IPG it can sometimes be rational to contribute to conflict, if one's contribution moves the group from a loss to a tie or a tie to a win, and the public good is sufficiently large relative to the cost of contribution. As a result, the incentive for contributing in the IPG is greater than the incentive for contributing in the IPD. In practice, average contribution rates of approximately 60% have been found in the IPD, relative to rates of about 27% in the IPD (Bornstein 1992: 601).²¹

Given these differences, it would be interesting to see how patterns of norm enforcement would differ between the IPD and the IPG. On one hand, the greater incentive to contribute in the IPG might translate to greater levels of norm enforcement, as the stakes of maintaining contribution are greater. On the other hand, the already-high stakes might encourage sufficient contribution to render norm enforcement unnecessary.

²¹ These proportions are for the case in which participants cannot communicate with one another. Contribution increases with communication, but is still greater in the IPG.

Testing the Mechanism

An important step for further research is to attempt to distinguish between the three theoretical mechanisms discussed in this dissertation. This dissertation has established that conflict can cause norm enforcement, but more must be done to explain why it has this effect. Some initial steps to address the framing mechanism did not yield evidence supportive of that mechanism, suggesting the signaling or intrinsic preference perspectives may provide a better account of the data. Nevertheless, each theory should be addressed more directly, with an experiment designed to test the specific underlying mechanism. I outline some possible studies that could accomplish this in the next section.

To test the signaling theory, it is necessary to systematically vary the signaling value of norm enforcement, or the extent to which participants can use their present behavior to communicate their future behavior to other group members. One way to do this would be to include a condition in which the groups are reconstituted on every round, so that participants have different teammates for every new interaction (Fehr and Gächter 2002). In this way, participants would have no knowledge of how their teammates behaved in the past, and thus no incentive to attempt to send signals about their own future behavior. This condition could be compared with a condition in which the groups remained the same on each round. In addition, conflict could be varied as in the present studies, and outgroup contribution could be varied by using simulated partners, with the outgroup programmed to either contribute often or rarely.

The key prediction would be that the two-way interaction between conflict and outgroup contribution would be qualified by a three-way interaction between conflict, outgroup contribution, and group composition, such that the two-way interaction would be significantly weakened when groups were reconstituted every round. One potential drawback of this method is that reconstituting the groups might also weaken

the effect of outgroup contribution, because participants do not know if they will be facing a hostile outgroup from round to round. This could potentially be addressed by varying only the ingroup composition, while the outgroup remained constant. Since partners are simulated, this does not pose the problem of only some ingroup members being aware of the outgroup's behavior.

Alternatively, another approach would be to allow team members to track one another's behavior by using reputation scores. In this method, in one condition participants would see a running total of the amount of resources each of their teammates invested in norm enforcement, and in other condition, participants would be unaware of their teammates' investments. The study would be otherwise similar to the partner switching study described above, and would similarly predict that the two-way interaction of conflict and outgroup contribution would be of lesser magnitude in the no-reputation score condition than in the reputation score condition. To heighten the incentive to signal, participants could be informed that either that they would interact with their team members in some form after the study (e.g. Barclay 2006; Willer 2006), or that their team members would have the opportunity to reward them for their behavior following the study (Horne 2004). This is similar to a theoretical argument made by Willer (2006), suggesting that status rewards for contribution may be greater during intergroup conflict. For reasons noted in chapter 2, norm enforcement may have even stronger signaling potential in intergroup conflict.

Testing the evolved preference hypothesis may prove more challenging. The most tractable approach could be to test the proximate mechanism, that neurological rewards for norm enforcement are greater in conflict. To test this idea, one could conduct a study using simulated partners, in which conflict (conflict or no conflict) is crossed with level of outgroup contribution (high or low). Participants could be scanned using PET or fMRI to determine if levels of activation in reward centers

during norm enforcement is greater when the ingroup is in conflict with a competitive outgroup. If this is the case, it would indicate that participants derive greater satisfaction from enforcing norms when the group is under threat than when the group is not under threat. While this cannot test the evolved preferences theory directly, it would provide evidence for the proximate mechanism derived from the theory.

Addressing the framing theory raises some new questions. I tested the theory using a questionnaire approach similar to that used earlier (Bornstein and Ben-Yossef 1994), but did not find support for this mechanism. As the theory is not well specified – for example, it is not explained in the theory *why* conflict would change an individual's preferences – it is not clear why the results would diverge from past findings. The answer may be related to the fact that conflict did not also substantially motivate contribution to the group, contrary to what Bornstein has found previously. Earlier, I suggested that this might be due to either the addition of sanctioning in the current studies, or to structural or cultural differences between the study samples. Investigating these possibilities in depth (for example, with a study that manipulates the availability of a sanctioning system), may help account for the differences between this and earlier studies.

New Directions

In addition to addressing the limitations of the current studies and refining the theoretical mechanism, I also plan to push the work in new theoretical directions. My current research focuses on pairs of groups in conflict. More generally, this is true of all prior experimental research on intergroup conflict of which I am aware. However, in the real world, the effects of conflict are never limited to just two actors. Instead, actors' behavior in conflict resonates beyond their immediate in- and outgroups. Aggressive or conciliatory actions send signals to third party observers about how the actor might behave in the future, which in turn shapes third parties' future interactions

with that actor. For example, contemporary debates about troop levels in Iraq have been dominated by attempts to predict how prospective allies and antagonists might interpret and react to changes in those levels.

Accordingly, my next step will be to study conflict between multiple individuals and groups embedded in networks. Moving from a two-group setting to a multi-actor or multi-group setting introduces new possibilities for signaling, reputation-building, and coalition formation. Studying conflict in networks will also allow me to address recently converging theories in sociology, psychology, economics, and theoretical biology exploring the role of reputations in conflict and aggression. I will begin with the more tractable, but still challenging, problem of studying networks of individuals, before moving studying networks of groups.

For example, one puzzling aspect of human behavior is the apparent senselessness of much violence. People are often documented displaying irrational levels of aggression, in which the costs of these actions often far outweigh any immediate benefits. Someone may respond to a harmless verbal insult with a physical attack, risking violent retaliation as well as legal sanctions (Gould 2003). Similarly, people may punish petty theft by bringing an expensive lawsuit to bear against the perpetrator, which costs them far more than simply replacing the stolen item (Frank 1988). Why would people act aggressively when they know that they themselves will suffer for it?

Theory and research from a number of fields has converged on the idea that people will behave aggressively, even when it is costly in the short term, in order to build a reputation for aggression that will protect them from future exploitation (Frank 1988; Gould 1999, 2003; Johnstone and Bshary 2004; Nisbett and Cohen 1996). Others may be less likely to attack or offend someone if that person has a history of responding to such actions with irrationally high levels of aggression. The long-term

gains of warding off potential threats may outweigh the short-term costs of building a reputation for aggression.

The general claim that reputation systems can foster aggression by increasing its long-term returns to individuals has never been directly tested, because existing research does not manipulate the existence of a reputation system. For example, Nisbett and Cohen (1996) show that southerners tend to respond to certain situations more aggressively than northerners, but can only infer that this is because reputations for violence are more important in the south, because northerners and southerners may differ on a wide range of other variables, in addition to region of origin. I plan to address this question with a series of studies designed to determine whether reputation systems play a causal role in increasing costly aggression. These studies will involve participants interacting in a setting in which they must choose whether to behave aggressively towards others. The studies will also experimentally manipulate the presence of a reputation system while controlling for other variables, thereby allowing for a comparison of levels of aggression in settings with and without reputation systems.

One key prediction is that actors will engage in short-term hostility, even when costly, in order to build a reputation for aggression that will protect them from future attacks. A related prediction is that weak actors (those more likely to lose a confrontation) will engage in disproportionately high levels of aggression, because they face substantial incentives to deter antagonism by sending a (false) signal of their readiness for conflict.

Findings in support of the predictions would suggest important negative effects of reputation systems, which have thus far been generally viewed positively for their cooperation enhancing benefits (Nowak and Sigmund 1998). The findings would also help explain apparently irrational behaviors, such as aggression on the part of

weak social actors who have apparently nothing to gain from their actions. On a practical level, the results could help explain the prevalence of interpersonal aggression, and highlight ways in which it could be addressed by policy.

Conclusion

This dissertation builds on a tradition of research addressing the role conflict plays in shaping social life. It provides an explanation for how two forms of social structure - bottom-up informal rules and top-down systems of authority - can emerge from previously unstructured groups. In doing so, it shows one route by which an undifferentiated collection of strangers can become a group with clearly demarcated hierarchies and rules of conduct. Indeed, because participants in the study were literally strangers to one another, they have no group history or identity to draw on, no preexisting ties between participants, and no country or family to fight for. Nonetheless, even under these minimal conditions, people enforce norms and create hierarchies, sometimes at substantial cost. Outside of the lab – when the conflict is more dangerous, groups are more valued, and sanctions more serious – these patterns might be stronger. Conflict may thus serve as a source of “groupness” and, more broadly, play an important role in structuring human societies.

Furthermore, the norms and hierarchies that conflict produces are those that are likely to exacerbate conflict. As others have observed, conflict and social structure appear to be mutually sustaining (Black-Michaud 1975; Coser 1956). Conflict provides a motive for norms and hierarchies; norms and hierarchies provide a mechanism for waging conflict. This highlights the network-driven aspect of mobilization: people use local influence to turn their peers into partisans; they support leaders who will do the same on a larger scale (Hardin 1995). As William Manchester observed, “[m]en do not fight for flag or country, for the Marine Corps or glory or any other abstraction. They fight for one another.” While Manchester may have been

overly confident in his proposal – surely there are many reasons why humans fight – Studies 1 and 2 show that peer influence is an important component of understanding behavior in conflict. Conflict drives efforts to influence others, which in turn drive conflict. Perhaps, much participation conflict is driven not only by fears of what our enemies might do to us, but also fears (and hopes) about what our *friends* might do. Understanding how conflict affects social norms and leadership hierarchies addresses fundamental questions about the way conflict shapes and is shaped by the fabric of social life.

REFERENCES

- Abrams, Dominic and Michael A. Hogg. 1990. "An Introduction to the Social Identity Approach." Pp. 1-9 in *Social Identity Theory: Constructive and Critical Advances*. edited by Dominic Abrams and Michael A. Hogg. London: Harvester-Wheatsheaf.
- Allison, Paul D. 1992. "The Cultural Evolution of Beneficent Norms." *Social Forces* 71:279-301.
- Arrow, Kenneth J. 1950. "A Difficulty in the Concept of Social Welfare." *Journal of Political Economy* 58:328.
- Axelrod, Robert M. 1984. *The Evolution of Cooperation*. New York: Basic Books.
- 1986. "An Evolutionary Approach to Norms." *The American Political Science Review* 80:1095-1111.
- Barclay, Pat. 2006. "Reputational Benefits for Altruistic Punishment." *Evolution and Human Behavior* 27:325-44.
- Barclay, Pat and Stephen Benard. 2008. "Stability-Dependent Cooperation, Status, and 'Upping the Threat Level'". Paper presented at the Human Behavior and Evolution Society Conference, Kyoto, Japan.
- Baron, Jonathan. 2001. "Confusion of Group Interest and Self-Interest in Parochial Cooperation on Behalf of a Group." *Journal of Conflict Resolution* 45:283-96.

- Benard, Stephen. 2008. "Norms." Pp. 536-8 in *International Encyclopedia of the Social Sciences*, edited by William A. Darity Jr. Detroit: MacMillan Reference USA.
- Bendor, Jonathan and Piotr Swistak. 2001. "The Evolution of Norms." *American Journal of Sociology* 106:1493-545.
- Black-Michaud, Jacob. 1975. *Cohesive Force : Feud in the Mediterranean and the Middle East*. Oxford Eng.: B. Blackwell.
- Blake, Robert R. and Jane S. Mouton. 1961. "Reactions to Intergroup Competition Under Win-Lose Conditions." *Management Science* 7:420-35.
- Blake, Robert R., Herbert A. Shepard and Jane S. Mouton. 1964. *Managing Intergroup Conflict in Industry*. Houston, Tex.: Gulf Pub. Co.
- Bornstein, Gary. 1992. "The Free-Rider Problem in Intergroup Conflicts Over Step-Level and Continuous Public Goods." *Journal of Personality and Social Psychology* 62:597-606.
- 1992. "The Free-Rider Problem in Intergroup Conflicts Over Step-Level and Continuous Public Goods." *Journal of Personality and Social Psychology* 62:597.
- 2003. "Intergroup Conflict: Individual, Group, and Collective Interests." *Personality and Social Psychology Review* 7:129-45.
- Bornstein, Gary and Meyrav Ben-Yossef. 1994. "Cooperation in Intergroup and Single-Group Social Dilemmas." *Journal of Experimental Social Psychology* 30:52-67.

- 1994. "Cooperation in Intergroup and Single-Group Social Dilemmas." *Journal of Experimental Social Psychology* 30:52-67.
- Bornstein, Gary, Eyal Winter and Harel Goren. 1996. "Experimental Study of Repeated Team-Games." *European Journal of Political Economy* 12:629-39.
- Bowles, Samuel. 2006. "Group Competition, Reproductive Leveling, and the Evolution of Human Altruism." *Science* 314:1569-72.
- Boyd, Robert, Herbert Gintis, Samuel Bowles and P. J. Richerson. 2003. "The Evolution of Altruistic Punishment." *Proceedings of the National Academy of Sciences of the United States of America* 100:3531-5.
- Brewer, Marilynn B. 2001. "Ingroup Identification and Intergroup Conflict: When does Ingroup Love Become Outgroup Hate?" Pp. 17-41 in *Social Identity, Intergroup Conflict, and Conflict Reduction*, edited by Richard D. Ashmore, Lee Jussim and David Wilder. New York: Oxford University Press.
- Centola, Damon, Robb Willer and Michael W. Macy. 2005. "The Emperor's Dilemma: A Computational Model of Self-Enforcing Norms." *American Journal of Sociology* 110:1009-40.
- Coben, Stanley. 1964. "A Study in Nativism: The American Red Scare of 1919-20." *Political Science Quarterly* 79:52-75.
- Coleman, James S. 1990. *Foundations of Social Theory*. Cambridge, Mass.: Harvard University Press.
- Coser, Lewis A. 1956. *The Functions of Social Conflict*. Glencoe, Ill.: Free Press.

- Dawes, Robyn M. 1980. "Social Dilemmas." *Annual Review of Psychology* 31:169-93.
- de Quervain, Dominique J. -, Urs Fischbacher, Valerie Treyer, Melanie Schellhammer, Ulrich Schnyder, Alfred Buck and Ernst Fehr. 2004. "The Neural Basis of Altruistic Punishment." *Science* 305:1254-8.
- Ellickson, Robert C. 1994. *Order without Law : How Neighbors Settle Disputes*. Cambridge, Mass.: Harvard University Press.
- Ellis, Desmond P. 1971. "The Hobbesian Problem of Order: A Critical Appraisal of the Normative Solution." *American Sociological Review* 36:692-703.
- Elster, Jon. 1989. *Nuts and Bolts for the Social Sciences*. Cambridge ; New York: Cambridge University Press.
- Erasmus, Charles J. 1967. "Obviating the Functions of Functionalism." *Social Forces* 45:319-28.
- Erikson, Kai. 1966. *Wayward Puritans; a Study in the Sociology of Deviance*. New York: Wiley.
- Fehr, Ernst and Urs Fischbacher. 2003. "The Nature of Human Altruism." *Nature* 425:785-91.
- Fehr, Ernst and Simon Gächter. 2002. "Altruistic Punishment in Humans." *Nature* 415:137-40.
- Fehr, Ernst and Bettina Rockenbach. 2003. "Detrimental Effects of Sanctions on Human Altruism." *Nature* 422:137-40.

Fiske, Susan T. 2002. "What we Know Now about Bias and Intergroup Conflict, the Problem of the Century." *Current Directions in Psychological Science* 11:123.

Flache, Andreas and Michael W. Macy. 1996. "The Weakness of Strong Ties: Collective Action Failure in a Highly Cohesive Group." *Journal of Mathematical Sociology*.

Frank, Robert H. 1988. *Passions within Reason: The Strategic Role of Emotions*. New York: Norton.

Gibson, James L. 1988. "Political Intolerance and Political Repression during the McCarthy Red Scare." *The American Political Science Review* 82:511-29.

Gintis, Herbert, Samuel Bowles, Robert Boyd and Ernst Fehr. 2003. "Explaining Altruistic Behavior in Humans." *Evolution and Human Behavior* 24:153-72.

Gintis, Herbert, Eric A. Smith and Samuel Bowles. 2001. "Costly Signaling and Cooperation." *Journal of Theoretical Biology* 213:103-19.

Gould, Roger V. 1999. "Collective Violence and Group Solidarity: Evidence from a Feuding Society." *American Sociological Review* 64:356-80.

-----2003. *Collision of Wills : How Ambiguity about Social Rank Breeds Conflict*. Chicago: University of Chicago Press.

Gurerk, Ozgur, Bernd Irlenbusch and Bettina Rockenbach. 2006. "The Competitive Advantage of Sanctioning Institutions." *Science* 312:108-11.

Hamblin, Robert L. 1958. "Leadership and Crises." *Sociometry* 21:322-35.

- Hardin, Garrett. 1968. "The Tragedy of the Commons." *Science* 162:1243-8.
- Hardin, Russell. 1995. *One for all : The Logic of Group Conflict*. Princeton, N.J.: Princeton University Press.
- Hechter, Michael. 1987. *Principles of Group Solidarity*. Berkeley: University of California Press.
- Hechter, Michael and Christine Horne. 2003. *Theories of Social Order : A Reader*. Stanford, Calif.: Stanford Social Sciences.
- Hechter, Michael and Karl-Dieter Opp. 2001. *Social Norms*. New York: Russell Sage Foundation.
- Heckathorn, Douglas D. 1988. "Collective Sanctions and the Creation of Prisoner's Dilemma Norms." *The American Journal of Sociology* 94:535-62.
- 1990. "Collective Sanctions and Compliance Norms: A Formal Theory of Group-Mediated Social Control." *American Sociological Review* 55:366-84.
- Hertzler, J. O. 1940. "Crises and Dictatorships." *American Sociological Review* 5:157-69.
- Hewstone, Miles, Mark Rubin and Hazel Willis. 2002. "Intergroup Bias." *Annual Review of Psychology* 53:575-604.
- Homans, George C. 1950. *The Human Group*. New York: Harcourt, Brace.
- 1961. *Social Behavior: Its Elementary Forms*. New York: Harcourt, Brace & World.

- Horne, Christine. 2001a. "The Enforcement of Norms: Group Cohesion and Meta-Norms." *Social Psychology Quarterly* 64:253-66.
- 2001b. "Sociological Perspectives on the Emergence of Social Norms." Pp. 3-34 in *Social Norms*, edited by Michael Hechter and Karl-Dieter Opp. New York: Russell Sage Foundation.
- 2004. "Collective Benefits, Exchange Interests, and Norm Enforcement." *Social Forces* 82:1037-62.
- Horne, Christine and Chien-fei Chen. 2005. "Can any Behavior Become Normative? Conflict, Coordination, and Norm Enforcement." *Unpublished Manuscript*.
- Horne, Christine and Anna Cutlip. 2002. "Sanctioning Costs and Norm Enforcement: An Experimental Test." *Rationality and Society* 14:285-307.
- Johnstone, Rufus A. and Redouan Bshary. 2004. "Evolution of Spite through Indirect Reciprocity." *Proceedings of the Royal Society B: Biological Sciences* 271:1917-22.
- Komorita, Samuel S. and Craig D. Parks. 1996. *Social Dilemmas*. United States: Westview Press.
- Lahti, David C. and Bret S. Weinstein. 2005. "The Better Angels of our Nature: Group Stability and the Evolution of Moral Tension." *Evolution and Human Behavior* 26:47-63.
- Macy, Michael W. 1993. "Backward-Looking Social Control." *American Sociological Review* 58:819-36.

- McAdams, Richard H. 1997. "The Origin, Development, and Regulation of Norms." *Michigan Law Review* 96:338-433.
- Messick, David M., Henk Wilke, Marilyn B. Brewer, Roderick M. Kramer, Patricia E. Zemke and Layton Lui. 1983. "Individual Adaptations and Structural Change as Solutions to Social Dilemmas." *Journal of Personality and Social Psychology* 44:294-309.
- Molm, Linda D. 1994. "Is Punishment Effective? Coercive Strategies in Social Exchange." *Social Psychology Quarterly* 57:75-94.
- Nisbett, Richard E. and Dov Cohen. 1996. *Culture of Honor : Psychology of Violence in the South*. Boulder, Colo.: Westview Press.
- Nowak, Martin A. and Karl Sigmund. 1998. "Evolution of Indirect Reciprocity by Image Scoring." *Nature* 393:573-7.
- Oberschall, Anthony. 1978. "Theories of Social Conflict." *Annual Review of Sociology* 4:291-315.
- Oliver, Pamela. 1980. "Rewards and Punishments as Selective Incentives for Collective Action: Theoretical Investigations." *The American Journal of Sociology* 85:1356-75.
- Olson, Mancur. 1965. *The Logic of Collective Action; Public Goods and the Theory of Groups*. Cambridge, Mass.: Harvard University Press.
- Palfrey, Thomas R. and Howard Rosenthal. 1983. "A Strategic Calculus of Voting." *Public Choice* 41:7-53.

- Posner, Richard A. 1997. "Social Norms and the Law: An Economic Approach." *The American Economic Review* 87:365-9.
- Pruitt, D. G. and M. J. Kimmel. 1977. "Twenty Years of Experimental Gaming: Critique, Synthesis, and Suggestions for the Future." *Annual Review of Psychology* 28:363-92.
- Pruitt, Dean G. and Jeffrey Z. Rubin. 1986. *Social Conflict: Escalation, Stalemate, and Settlement*. New York: Random House.
- Rabe-Hesketh, S. and Anders Skrondal. 2005. *Multilevel and Longitudinal Modeling using Stata*. College Station, Tex.: Stata Press.
- Rapoport, Anatol and Albert M. Chammah. 1965. *Prisoner's Dilemma; a Study in Conflict and Cooperation*. Ann Arbor: University of Michigan Press.
- Rilling, James K., David A. Gutman, Thorsten R. Zeh, Giuseppe Pagnoni, Gregory S. Berns and Clinton D. Kilts. 2002. "A Neural Basis for Social Cooperation." *Neuron* 35:395-405.
- Rutte, C. G. and Henk Wilke. 1985. "Preference for Decision Structures in a Social Dilemma Situation." *European Journal of Social Psychology* 15:367-70.
- Samuelson, Charles D. and David M. Messick. 1986. "Inequities in Access to and use of Shared Resources in Social Dilemmas." *Journal of Personality and Social Psychology* 51:960-7.

- Samuelson, Charles D., David M. Messick, Christel Rutte and Henk Wilke. 1984. "Individual and Structural Solutions to Resource Dilemmas in Two Cultures." *Journal of Personality and Social Psychology* 47:94-104.
- Sherif, Muzafer. 1966. *In Common Predicament; Social Psychology of Intergroup Conflict and Cooperation*. Boston: Houghton Mifflin.
- Sherif, Muzafer and Gardner Murphy. 1936. *The Psychology of Social Norms*. New York and London: Harper & brothers.
- Simmel, Georg. [1908]1955. *Conflict*. Glencoe, Ill.: Free Press.
- Simpson, Brent. 2003. "Sex, Fear, and Greed: A Social Dilemma Analysis of Gender and Cooperation." *Social Forces* 82:35-52.
- Singer, Tania, Ben Seymour, John P. O'Doherty, Klaas E. Stephan, Raymond J. Dolan and Chris D. Frith. 2006. "Empathic Neural Responses are Modulated by the Perceived Fairness of Others." *Nature* 439:466-9.
- Staw, Barry M., Lance E. Sandelands and Jane E. Dutton. 1981. "Threat Rigidity Effects in Organizational Behavior: A Multilevel Analysis." *Administrative Science Quarterly* 26:501-24.
- Stein, Arthur A. 1976. "Conflict and Cohesion: A Review of the Literature." *The Journal of Conflict Resolution* 20:143-72.
- Sumner, William G. and Albert G. Keller. 1906. *Folkways; A Study of the Sociological Importance of Usages, Manners, Customs, Mores, and Morals*. Boston, New York: Ginn and company.

- Tajfel, H. 1982. "Social Psychology of Intergroup Relations." *Annual Review of Psychology* 33:1-39.
- Tilly, Charles. 1989. "Cities and States in Europe, 1000-1800." *Theory & Society* 18:563.
- Turner, Jonathan H. 1982. *The Structure of Sociological Theory*. Homewood, Ill.: Dorsey Press.
- van den Berg, Axel. 1998. "Is Sociological Theory Too Grand for Social Mechanisms?" Pp. 204-37 in *Social Mechanisms*, edited by Peter Hedstrom and Richard Swedberg. United States: Cambridge University Press.
- Wagner, Thomas. 2006. "U.S. Strategy Threatened as Sunnis Attack Sunnis in Iraq." *Houston Chronicle*.
- Willer, Robb. 2006. "A Status Theory of Collective Action." Unpublished Ph.D. Dissertation, Cornell University.
- Willer, Robb, Ko Kuwabara and Michael W. Macy. 2008. "The Emperor's Dilemma II: The False Enforcement of Unpopular Norms." Manuscript in preparation. Berkeley, CA.
- Williams, Robin M. 1947. *The Reduction of Intergroup Tensions: A Survey of Research on Problems of Ethnic, Racial, and Religious Group Relations*. New York: Social Science Research Council.
- Yamagishi, Toshio. 1986. "The Provision of a Sanctioning System as a Public Good." *Journal of Personality and Social Psychology* 51:110-6.

-----1988. "The Provision of a Sanctioning System in the United States and Japan."

Social Psychology Quarterly 51:265-71.

-----1988. "Seriousness of Social Dilemmas and the Provision of a Sanctioning

System." *Social Psychology Quarterly* 51:32-42.