November 2017

A Virtuous Cycle: Interdependence and the Ontogenesis of Social Justice

Nathan Carnes

A VIRTUOUS CYCLE: INTERDEPENDENCE AND THE ONTOGENESIS OF SOCIAL JUSTICE

A Dissertation Presented

by

NATHAN C. CARNES

Submitted to the Graduate School of the
University of Massachusetts Amherst in partial fulfillment
of the requirements for the degree of

DOCTOR OF PHILOSOPHY

September 2017

Social Psychology

A VIRTUOUS CYCLE: INTERDEPENDENCE AND THE ONTOGENESIS OF SOCIAL JUSTICE

A Dissertation Presented

by

NATHAN C. CARNES

Approved as to style and content by:

_____
Ronnie Janoff-Bulman, Chair

_____
Donald Tomaskovic-Devey, Member

_____
Brian Lickel, Member

_____
Aline Sayer, Member

_____
Caren Rotello, Department Head
Department of Psychological and Brain Sciences

DEDICATION

This dissertation is dedicated to my family, who taught me that love must be passed on rather than repaid. In many ways, the thesis of this work is rooted in that piece of wisdom shared from one generation to the next. I love the six of you very much.

ACKNOWLEDGMENTS

Social psychology teaches us about the power of the situation, so it follows that I would not be the scholar or person that I am today without a supportive environment and great mentors stretching from Hendrix College to the University of Massachusetts, Amherst. You ignited my imagination, encouraged me to think about the big picture, taught me how to test my ideas, and showed me how to teach others. You know what you did for me, and I want you to know that I will always be grateful for your time, wisdom, and friendship. Finally yet importantly, I also want to thank my graduate student peers and undergraduate research assistants; you motivated me, commiserated with me, helped me, and befriended me on this six-year journey.

ABSTRACT

A VIRTUOUS CYCLE: INTERDEPENDENCE AND THE ONTOGENESIS OF SOCIAL JUSTICE

SEPTEMBER 2017

NATHAN C. CARNES, B.A., HENDRIX COLLEGE

M.S., UNIVERSITY OF MASSACHUSETTS AMHERST

Ph.D., UNIVERSITY OF MASSACHUSETTS AMHERST

Directed by: Professor Ronnie Janoff-Bulman

Past research suggests that inequality has been the norm throughout most of human history (Piketty, 2014) until the shocks of the 20th century—especially mass mobilization for the world wars—spurred support for progressive public policies around taxation and social welfare (Scheve & Stasavage, 2010, 2012). The present research investigates why these collective events activated people's beliefs about fairness. We propose that these collective events imbued people with a powerful sense of interdependence, and that this feeling of interdependence gave motivational force to the belief that group members have a moral responsibility to share and contribute to the welfare of the group. We test this link between interdependence and social justice across five studies (N = 1,646) using diverse methodologies and quantitative techniques. As predicted, our analyses revealed that social justice is of greater importance in more interdependent groups of everyday life (Study 1) and for Americans whom feel a stronger sense of interdependence (Study 2). In fact, manipulating interdependence in well-mixed groups fosters social justice behavior, which in turn nurtures stronger perceptions of interdependence (Studies 3 – 5). The implications of this virtuous cycle between interdependence and social justice for social psychology, public policy, and contemporary events are discussed.

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Overview of Present Research

Much of the classic psychological research on social justice neglects one of its defining characteristics—social justice is something that happens in groups. It is not that they deny the presence of groups, for distributive justice is fundamentally an accounting of how goods ought to be distributed in a society (e.g., Deutsch, 1985). It is that other features of justice are given such weight as to make the social part seem trivial. In this stream of research, justice is something that depends only on the inherent features of actions (e.g., Turiel, 1983) or only on the balance of inputs and outputs (e.g., Walster, Walster, & Bersheid, 1978); it is something that people learn to reason about in a series of distinct hierarchical stages relatively divorced from social context (e.g., Kohlberg, 1981). However, in the following decades, researchers began to acknowledge the importance of groups for understanding social justice. In particular, work on procedural justice emphasizes how relational motives for belonging and respect can drive judgments of fairness (e.g., Lind & Tyler, 1988). Other work on social identity reveals how self-relevant ingroups can shape what criterion justice decisions are based on (e.g., Wenzel, 2002), when people are concerned about justice (e.g., Skitka, 2003), and the scope of justice or who it should and should not apply to (e.g., Clayton & Opotow, 2003).

My dissertation builds on this past work because it connects social justice to the underlying properties or dynamics of the groups in which it is embedded. I think this socially embedded approach can help explain when and why groups act on their members' convictions about social justice. Chapter 1 of my dissertation provides the

theoretical background and conceptual questions that underpin the present research. More specifically, I provide a brief history of inequality by drawing on recent work from economics and political science to put the present research in context. Then, I integrate this work with theory and research on morality from social psychology to guide the questions and hypotheses of my dissertation. Chapters 2 – 6 of my dissertation describe five studies I have conducted to better understand the group dynamics underlying social justice. Finally, Chapter 7 of my dissertation discusses the implications of the present research for our understanding of how inequality is perpetuated and remediated.

A Brief History of Inequality

Injustice takes many forms, but distributional inequality is a pernicious and enduring example. Piketty (2014) argues that inequality has been the norm throughout most of human history because it is the natural equilibrium for the distribution of wealth under historically typical conditions. His thesis is that inequality is a mathematical certainty when the return on capital exceeds the growth of incomes—barring major shocks or public policy interventions—because pools of wealth then grow faster than the economy as a whole. The earliest available data on capital to income ratios—which tell us something about the value of capital relative to the value of labor—show that this ratio was about 7:1 in 18th century France and Britain. The earliest available data on inheritances from 19th century France tell a similar story; the lifetime value of the top 1% of inheritances was three times larger than the lifetime value of the top 1% of earnings, meaning that a large inheritance offered the possibility of wealth simply unachievable through any form of work. In fact, there is evidence that this imbalance between capital and income characterized economic life stretching back from antiquity until the

beginning of the 20th century, when the Great Depression and two World Wars seem to have fostered unprecedented levels of equality in many industrialized nations. Piketty explains that these shocks destroyed more than half of all past wealth, and progressive taxation then prevented the wealth of the post-war years from aggregating once again among the few.

Piketty's work raises an important and timely question. Evidence from diverse disciplines suggests that people care about inequality (e.g., Dawes, Fowler, Johnson, McElreath, & Smirnov, 2007; Tricomi, Rangel, Camerer, & O'Doherty, 2010). In fact, even children (Shaw & Olson, 2012) and non-human primates (Brosnan & de Waal, 2003) seem to exhibit an aversion to inequality. Furthermore, equal societies are generally happier, healthier, and wealthier than unequal societies (e.g., Wilkinson & Pickett, 2009). So why was the middle of the 20th century the only time in the recorded history of humankind that we taxed the rich more than the rest? Furthermore, why have so many governments dismantled the progressive public policies of the 20th century and allowed a seemingly bygone era of radical inequality to return once more (for a review of rising inequality, see Piketty, Saez, & Zucman, 2016)?

A partial answer to these questions comes from the work of Scheve and Stasavage, who demonstrate that mass mobilization for warfare had an important effect on the rise of progressive public policies such as income taxation (2010) and inheritance taxation (2012; for a review, see Scheve & Stasavage, 2016). They find that progressive taxation as a government policy depends very little, if at all, on actual levels of inequality, the establishment of universal suffrage in democratic societies, or how liberal the political parties and institutions are. Instead, when governments mobilize for war by

asking tens of millions of their disproportionately younger and poorer citizens to serve in the military, it activates people's beliefs about fairness because the burden of the war effort is unequally shared among different segments of society. Thus, progressive taxation is seen as fair, even among the wealthiest members of society, because it compensates for the unequally distributed sacrifices of the war effort. "The conscription of wealth in the form of progressive taxation constituted part of a new social compact in which the mass of citizens agreed to fight while the rich agreed to bear a higher tax burden" (Scheve & Stasavage, 2010, p. 530). However, as the mass mobilization wars of the 20th century have faded from memory and into history, so too has the resonance of this compensatory social compact and the progressive public policies for which it laid the foundation.

The Moral Psychology of Fairness

People's beliefs about fairness play a pivotal role in how we explain the rise and fall of both distributional equality and progressive taxation. Despite this critical role, Scheve and Stasavage (2016) cite only a small sample of behavioral economic studies suggesting that people have prosocial preferences in addition to self-interested ones. If we want to understand what these prosocial preferences are and how they work, we should draw on the rich body of theory and research investigating the psychological mechanisms that underlie morality in general and fairness in particular. The most important insight—first proposed in this literature by Brickman, Folger, Goode, and Schul (1981)—is that there are two qualitatively different kinds of fairness. The basic idea is that we have a paucity of language with which to describe allocations that we dislike, we just say something is fair or unfair, but these words can have different

meanings. For example, consider the debate over affirmative action in which both sides claim to hold the only fair position. Opponents of affirmative action rely on a moral principle that Brickman et al. (1981) call microjustice; fairness is an empirical property of a set of individuals in which inputs like merit determine outputs like admission. In contrast, proponents of affirmative action rely on a moral principle that Brickman et al. (1981) call macrojustice; here, fairness is a desirable property of a distribution in which resources like education are awarded to produce a particular pattern of outcomes like equal representation.

The labels assigned to microjustice and macrojustice in the literature have changed over time and between authors. For clarity, I adopt the language prescribed by Janoff-Bulman and Carnes (2013), who use the terms fairness (i.e., microjustice) and social justice (i.e., macrojustice). In recent work, my colleagues and I have found support for making a meaningful distinction between these two constructs. For example, when participants rate the personal importance of these two moral principles, the items load on discrete latent factors (Janoff-Bulman & Carnes, 2016a). When participants rate how applicable these two moral principles are in various groups, we find that fairness and social justice are used in distinct types of social contexts (Carnes, Lickel, & Janoff-Bulman, 2015). Moreover, when participants rate how fair and how just different moral scenarios are, we find that these moral principles are sensitive to different kinds of social information (Bharadwaj & Janoff-Bulman, in preparation). As will soon become apparent, this distinction between fairness and social justice is important if we are to understand why levels of equality and progressive taxation surged only in the middle of the 20th century.

Fairness judgments are made within interpersonal contexts and are based on personal characteristics like deservingness and need (Janoff-Bulman & Carnes, 2013). Moral Foundations Theory classifies fairness as an individualizing moral principle because it ascribes value to these personal characteristics, motivating people to care about the rights of individuals (e.g., Graham et al., 2011; Haidt, 2007; Haidt & Kesebir, 2010). It is important for people to believe that fairness ought to apply to everyone (Turiel, 1983) and ought to be endorsed by everyone (Mullen & Skitka, 2006) because this helps groups avoid the risks of both exploitation and infighting (Descioli, 2016; Descioli & Kurzban, 2013; Thomas, De Freitas, DeScioli, & Pinker, 2016). Put differently, by protecting the rights of individuals, fairness helps people live together as autonomous social actors who can cooperate strategically with minimal personal risk (Haidt & Graham, 2009; Shweder, Much, Mahapatra, & Park, 1997). Thus, this moral principle emphasizes doing the right thing from a perspective of enlightened self-interest (e.g., Haidt, 2012; see also Deutsch, 1985), which fits with the idea that fairness is an evolutionary product of direct and indirect reciprocity (for a review, see Rand & Nowak, 2013).

In contrast, social justice judgments are made within group contexts and are based on a priori constraints on the pattern or shape of the distribution (Janoff-Bulman & Carnes, 2013). The group context is important here, as Wenzel (2000, 2002, 2004) emphasizes that social justice is more strongly rooted in social identity than enlightened self-interest. Classic research on social identity suggests that people exhibit a preference for the ingroup (Tajfel, Billig, Bundy, & Flament, 1971) because their conception of the self is intimately tied to the groups they belong to (Tajfel, 1978). These self-relevant

groups provide norms for what constitutes right behavior (Turner, Hogg, Oakes, Reicher, & Wetherell, 1987) that then guide the actions of members and how they expect others in the group to perceive them (Ellemers, Pagliaro, Barreto, & Leach, 2008; Pagliaro, Ellemers, & Barreto, 2011; for a review, see Ellemers & van den Bos, 2012). These ingroup norms help explain why group members are more likely to help other people in the group (e.g., Dovidio, 1984; Levine, Cassidy, Brazier, & Reicher, 2002) and cooperate with fellow group members (e.g., De Cremer & Van Vugt, 1999; Kramer & Brewer, 1984).

Janoff-Bulman and Carnes (2013) argue that social justice promotes moral behavior by leveraging this group psychology. More specifically, social justice is a group-based moral principle because it ascribes value to the roles and responsibilities accompanying group membership, motivating people to care about the welfare of the group. By committing individuals to the welfare of the group, social justice helps people live together as members of a cohesive community that can cooperate around their shared goals and values. Thus, this moral principle emphasizes doing the right thing from a perspective of parochial altruism (Carnes & Janoff-Bulman, 2016; Janoff-Bulman & Carnes, 2016b), which fits with the idea that some parts of our moral psychology are the evolutionary product of group selection (e.g., Haidt, 2012; Richerson & Boyd, 2005; Wilson, 2012).

It might seem counterintuitive to assert that social justice is somehow parochial or groupish when past work has linked social justice to inclusive categorization (Rock & Janoff-Bulman, 2010), political liberalism (Janoff-Bulman & Carnes, 2016a), and egalitarian social preferences (Janoff-Bulman, Sheikh, & Baldacci, 2008). Indeed,

Graham et al. (in press) describes social justice as a centrifugal force because it pushes our circle of moral concern outward to include more distant others as deserving of moral treatment. One potential answer to this apparent conflict is that groups bound together by interdependence—by their shared goals and outcomes—might be more inclusive than groups bound together by a common location, history, or physical characteristic. In fact, this is what Durkheim (1893/1984) predicted in distinguishing organic solidarity from mechanical solidarity more than a century ago.

How Convictions Bind Groups

We can think of "morality as the glue that binds" (Haidt, 2008, p. 65). Crediting Durkheim, Haidt notes, "morality constrains individuals and ties them to each other to create groups" (2007, p. 1000; for a review of binding morality, see Graham & Haidt, 2010). According to Durkheim (1893/1984), members of a society can be bound together by their fundamental sameness—possessing shared beliefs, values, and customs—and he calls this unifying force mechanical solidarity. Carnes and Lickel (under review) demonstrate that moral principles around hierarchy and authority (e.g., social order) can indeed foster this sense of sameness, providing group members with a visceral sense of oneness with the group called identity fusion (for a review of identity fusion, see Swann & Buhrmester, 2015). However, Durkheim (1893/1984) also argues that members of a society can be bound together by their complementarity or interdependence, and he calls this unifying force organic solidarity.

Interestingly, Durkheim (1893/1984) acknowledges that social justice is part of the collective consciousness or shared beliefs that foster organic solidarity. That is, social justice is more than a distributional principle; it is a kind of social glue that helps create

tight-knit groups out of collections of individuals. Psychological research on entitativity—the degree to which a group is perceived to be a real entity—has found support for the role of interdependence in binding people together apart from similarity (e.g., Brewer, Hong, & Li, 2004; Haslam, Rothschild, & Ernst, 2000). Drawing on Durkheim's insight, the thesis of my proposed dissertation is that social justice and interdependence are intimately linked (for a review, see Janoff-Bulman & Carnes, 2013). I think this link might be able to explain why levels of equality and progressive taxation surged in the middle of the 20th century.

The great depression and world wars of the last century gave societies a powerful sense of interdependence or interconnectedness—this feeling that we are all in this together for better or worse—because citizens became aware of their shared fate in both economics and conflict. The claim that societies become more communal in difficult times is supported by recent cross-temporal evidence that individualism swells during economically prosperous times while interdependence swells during recessionary times (Bianchi, 2016). The central thrust of my argument is that this feeling of interdependence gave motivational force to the belief that group members have a moral responsibility to share and contribute to the welfare of the group. Thus, members of the labor class gave soldiers to the war effort while members of the capital class gave wealth. Governments dismantled the progressive public policies of the 20th century because their citizens no longer felt interdependent; the wars eventually ended, the soldiers came home, and the prosperous post-war years fed into an upwelling of individualism. I think interdependence explains the ontogenesis of social justice in large modern societies, and my proposed dissertation will test this hypothesis using diverse methodologies.

CHAPTER 2

STUDY 1

Overview

We designed Study 1 to identify how basic group dynamics shape the contextual importance of different moral principles in daily social life. That is, the primary unit(s) of analysis in this experiment were the different groups people belong to and not so much the people themselves. We were particularly interested in the relationship between social justice and interdependence; our prediction was that social justice would be more important in more interdependent kinds of groups. That is, organic solidarity should predict social justice. However, we also assessed the relationship between tightness-looseness and social order. Tight groups have strong norms and low tolerance for deviant behavior, whereas loose groups have weak norms and high tolerance for deviant behavior (e.g., Gelfand et al., 2011); in other words, tight groups have homogenous beliefs and values. We predicted that to maintain such tight cohesion, groups would need moral values around authority, hierarchy, and tradition to enforce and maintain conformity. That is, mechanical solidarity should predict social order.

Method

Participants. We recruited 99 participants from the University of Massachusetts Amherst subject pool in exchange for extra credit (8.1% male; median age of 20). A majority of the participants identified as White (62.6%), and the remaining identified as Asian (17.2%), Black (4%), Hispanic (3%), or nominated another racial category (8%). No participants were excluded from analysis. A power analysis revealed that this sample

had sufficient power ($\rho = .80$) under realistic assumptions to detect a medium to small effect of our principal hypothesis ($\delta = .36$) (see Raudenbush et al., 2011).

Procedure. After reading detailed directions, participants completed the group listing task to generate self-relevant target groups. These groups were then piped into the group rating task, which asked participants to rate each target group along several dimensions. At the end of the experiment, participants completed the social value orientation measure and demographics, including political orientation.

Group Listing Task. We asked participants to think about the rich variety of groups they belong to, and we gave them some guidance to help get them started. We suggested that, "you might think of yourself as a member of a family, part of a university, a citizen of a country, having a particular race, gender, or religion, or even being in a group of strangers in an experiment." Participants then typed a list of 12 different groups they belong to now or sometime in the past. This generated an array of socially *in vivo* groups that were self-relevant to the participant. For example, one typical participant's anonymized list of groups included: family, university student, school friends, home friends, church, citizen of USA, employee of restaurant, female, Caucasian, member of experiment, resident of Massachusetts, and psychology major.

Group Rating Task. Participants rated the 12 target groups from the group listing task on 12 dimensions, one dimension at a time, following the procedure developed by Lickel et al. (2000; see also Carnes et al., 2015). Each dimension was presented in a randomized order and assessed on a 7-point rating scale using item-specific anchors. Interdependence ($\alpha = .826$), tightness-looseness ($\alpha = .745$), social justice ($\alpha = .825$), and

social order (α = .626) were sufficiently reliable across all groups when measured with three items each (see Appendix A for items).

Social Value Orientation Measure. This 9-item measure, originally developed by Van Lange (1999), asks participants to choose among three hypothetical distributions between the self and an anonymous other person. One option is prosocial (maximizing joint payoffs), another option is egoistic (maximizing one's absolute payoff), and the final option is competitive (maximizing one's relative payoff). Our sample was 54.5% prosocial, 24.2% egoistic, 10.1% competitive, and the remainder was inconsistent[1].

Political Orientation Measure. Participants rated two items using 7-point scales anchored from very liberal/strong democrat (1) to very conservative/strong republican (7). These items were sufficiently reliable ($r(97) = .702$).

Results

We analyzed the data using hierarchical linear modeling with full maximum likelihood estimation because the data had a nested structure with two levels (e.g., observations nested within persons). That is, the level-1 units were the different group targets that participants rated (i.e., observations), while the level-2 units were the individuals making the ratings (i.e., persons).

We first calculated the intraclass correlation[2] for each of the four measured dimensions to assess the proportion of variance that lies between group targets versus raters. We ran four one-way ANOVAs with random effects in which each dimension was a separate outcome to generate four intraclass correlations. This analysis revealed that a majority of the variance in interdependence (84.15%), tightness-looseness (77.77%), social justice (75.89%), and social order (72.44%) was due to differences between group

targets; the remainder, about 22.44% on average, was due to differences between individual raters.

We then assessed the relationship between the group property dimensions and the moral principle dimensions. We ran two random-coefficient regression models in which interdependence and tightness-looseness were the predictors[3] and either social justice or social order were the outcome. Our analyses revealed that interdependence strongly predicted social justice ($\gamma = .79$, $SE = .024$, $t(98) = 33.45$, $p < .001$), whereas tightness-looseness was only a weak predictor ($\gamma = .070$, $SE = .030$, $t(98) = 2.37$, $p = .020$). In fact, a multivariate contrast revealed that interdependence was a significantly stronger predictor ($\chi^2(1) = 248.95$, $p < .001$) and alone accounted for 71.62% of the level-1 variance in social justice. In contrast, tightness-looseness strongly predicted social order ($\gamma = .54$, $SE = .030$, $t(98) = 17.89$, $p < .001$), while interdependence was only a weak predictor ($\gamma = .17$, $SE = .028$, $t(98) = 6.07$, $p < .001$). A multivariate contrast revealed that tightness-looseness was a significantly stronger predictor ($\chi^2(1) = 61.43$, $p < .001$) and alone accounted for 46.25% of the level-1 variance in social order.

Finally, we investigated whether individual differences in politics or social value orientation moderate these relationships. We ran two intercepts-and-slopes-as-outcomes models in which politics and social value orientation were added as predictors to the prior models. These individual differences did not improve model fit to the observed data when predicting social justice scores ($\Delta\chi^2(6) = 4.66$, $p = .588$), suggesting no moderation, but provided a marginal improvement in model fit when predicting social order scores ($\Delta\chi^2(6) = 11.43$, $p = .076$). A closer inspection of the parameter estimates in this model revealed that social value orientation moderated the relationship between tightness-

looseness and social order (γ = .15, *SE* = .058, $t(85) = 2.57$, p = .012) such that this relationship was stronger for those with a proself rather than prosocial orientation[4].

Discussion

As predicted, social justice was more important in interdependent groups while social order was more important in tighter groups. These relationships had very large effect sizes, explaining almost 60% of the variance in morality on average, and did not appear to be moderated by politics. This is particularly interesting in light of extensive past research showing that conservatives endorse social order while liberals endorse social justice (e.g., Janoff-Bulman & Carnes, 2016; Janoff-Bulman et al., 2008). However, the relationship between social order and tightness-looseness was moderated by social value orientation; essentially, those who were more selfish thought that more constraints on self-interest were required to maintain tight groups.

Moreover, groups had a dominant effect on morality relative to individual differences, as indicated by the intraclass correlation, which replicates past research using this methodological approach (see Carnes et al., 2015). This suggests that group dynamics play an important role in guiding moral convictions, but that some group properties (e.g., interdependence) might have a unique relationship with some moral principles (e.g., social justice). We further explored this potentially important relationship between group dynamics and moral convictions in Study 2.

Notes

[1] We followed the convention, established by Van Lange and colleagues, of collapsing egoistic and competitive participants into a single category and dropping inconsistent participants for analyses with this construct.

[2] HLM calculates an intraclass correlation that quantifies the proportion of variance in an outcome variable that is between the level-2 units by dividing the level-2 variance component by the sum of both the level-1 and level-2 variance components (see Raudenbush & Bryk, 2002).

[3] Both predictors were group-mean centered because our primary research question involved the relationships at level-1 (see Enders & Tofighi, 2007). All random effects were included in the model.

[4] There was not a significant effect of the remaining parameter estimates (all $p \geq .112$).

CHAPTER 3

STUDY 2

Overview

We designed Study 2 to replicate and extend the findings of Study 1 while leveraging an alternative methodological approach. More specifically, we examined the relationship between group properties and moral principles for a single identity—Americans—rather than across the many groups of daily social life. Furthermore, we hoped to better infer the direction of this relationship by experimentally manipulating participants' sense of solidarity with other Americans.

We made a theoretical distinction between social justice—this sense of moral responsibility to the disadvantaged or the group writ large—and explicit support for equality in this study. This helped us unpack the psychological process linking interdependence with equality. We also assessed how inclusive people feel about other groups of humans, and even plants or animals, to test whether interdependence implies a parochialism that is less tribal and more global.

Method

Participants. We recruited 671 participants on Amazon Mechanical Turk via TurkPrime in exchange for payment (39.88% male; median age of 32) (see Litman, Robinson, & Abberbock, 2016). A majority of the participants identified as White (73.96%), and the remaining identified as Asian (5.66%), Black (9.67%), Hispanic (4.61%), or nominated another racial category (6.10%). We excluded 47 participants (7%) based on an American identity check, a manipulation check, and two catch questions designed to identify careless responders. The 624 included participants should

provide sufficient power to test our predicted multigroup structural equation model (45 estimated parameters) as well as alternative models, and the covariance matrices should provide sufficient information for this model to be identified.

Procedure. We randomly assigned each participant to one of two experimental manipulations (weak solidarity or strong solidarity) which involved a prompted writing task. Next, participants completed the group measures, the morality measures, the equality measures, and the inclusion measures in that order. We assessed demographics, including socioeconomic status, at the end of the experiment.

Experimental Manipulations. In both conditions, we first explained how "the social fabric of a community is a metaphor for the quality or closeness of relationships in that community." Participants randomly assigned to the weak solidarity condition wrote about "some of the ways in which America's social fabric has gotten weaker since the presidential election (even if you think it has gotten stronger in other ways)" for at least two minutes. Participants randomly assigned to the strong solidarity condition wrote about "some of the ways in which America's social fabric has gotten stronger since the presidential election (even if you think it has gotten weaker in other ways)" for at least two minutes. A coder read these essays to ensure participants responded to the prompt.

Group Measures. Identity fusion ($\alpha = .924$) and interdependence ($\alpha = .882$) were each measured with eight items presented in counterbalanced order. These items were all rated on a 7-point scale anchored from strongly disagree (1) to strongly agree (7). To measure identity fusion, we added an item to the identity fusion scale (Gómez et al., 2011). This item is: "I am nothing without my country." We developed and pilot tested the interdependence items for this study based on past research[5] (see Appendix A for

items). For analysis, we created four parcels for each construct. We used the item to construct balance approach for identity fusion and the a priori questionnaire construction approach for interdependence (see Little, Cunningham, Shahar, & Widaman, 2002).

Morality Measures. Social justice ($\alpha = .907$) and social order ($\alpha = .897$) were each measured with six items presented in counterbalanced order. Participants rated these items on a 7-point scale anchored from strongly disagree (1) to strongly agree (7). These items were adapted from Janoff-Bulman and Carnes (2016a) as well as the group rating task from Study 1 (see Appendix A for items). We used the item to construct balance approach to generate three parcels for each of these constructs.

Equality Measures. Equality was measured with three indicators: a 4-item self-report measure and two rating tasks presented in counterbalanced order (see Appendix A for self-report items). Participants rated these self-report items on a 7-point scale anchored from strongly disagree (1) to strongly agree (7). These self-report items were adequately reliable ($\alpha = .874$) so their composite acted as an indicator of equality. The first rating task presented participants with two bar charts depicting Black pay relative to White pay (65%) and female pay relative to male pay (80%). We asked participants to "imagine that you could decide how much (Black people/females) earn if you had control over the matter." Participants could manipulate the bar chart to indicate their preference. We created a single dummy variable indicating whether Black people and females should earn 100% of the pay of their counterparts. The second rating task presented participants with a bar chart depicting wealth inequality by quintile with the top 20% controlling 88% of all wealth. We asked participants to "imagine that you could decide how our country's wealth should be distributed if you had control over the matter." Participants could

manipulate the bar chart to indicate their preference. The Gini index of these wealth quintiles acted as the third and final indicator of equality[6].

Inclusion Measures. Inclusion was measured with two indicators: a subscale of the identification with all humanity scale (McFarland, Brown, & Webb, 2013) and the moral expansiveness scale (Crimston, Bain, Hornsey, & Bastian, 2016) presented in counterbalanced order. The identification with all humanity scale has nine items rated on a 5-point scale with item-specific anchors. Shortening the original scale, we only had participants rate "Americans" ($\alpha = .925$) and "people all over the world" ($\alpha = .921$). An example item is: "How close do you feel to each of the following groups?" We did not analyze the "Americans" version of the scale. The moral expansiveness scale asks participants to rate 24 targets[7] on a 4-point scale anchored from outside the moral boundary (1) to inner circle of moral concern (4). Example targets include: family member, dolphin, murderer, old-growth forest, and refugee.

Socioeconomic Status Measure. Participants answered two questions by entering values into numeric boxes. These questions asked about years of education and household income last year[8].

Results

We tested a multigroup structural equation model in LISREL 8.80 using full maximum likelihood estimation. We used a multigroup framework—with the groups being the strong solidarity and weak solidarity conditions—because this allowed us to test for moderation by our experimental manipulation. We predicted that interdependence would be positively associated with social justice while identity fusion would be positively associated with social order, and that these relationships would increase in

strength in the strong solidarity condition relative to the weak solidarity condition. We also predicted that social justice would be associated with greater support for equality and inclusion whereas social order would not. These predictions were made a priori and preregistered using the Open Science Framework (osf.io/34k4q).

Our baseline model, shown conceptually in the figure below, included 43 estimated parameters that were all held equivalent between the two groups. This included 19 error variances, 13 factor loadings, 2 factor variances, 4 disturbance variances, 6 paths, 2 disturbance covariances, and 1 factor covariance. The estimation process for this model converged on an admissible solution, and although the model chi-square was significant ($\chi^2(333) = 879.26$, $p < .001$)[9], the values for the other recommended fit indices were acceptable (*RMSEA* = .071, *CFI* = .963, $SRMR_1$ = .094, $SRMR_2$ = .114).

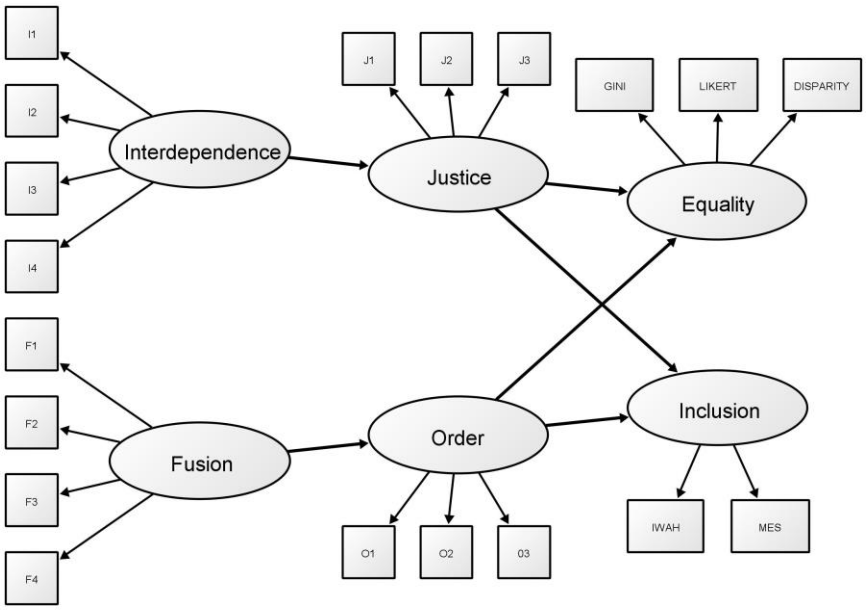

Figure 1: Conceptual Diagram of Predicted Model

Before testing for moderation, we investigated whether the model could be improved by estimating additional paths between latent constructs. A model with the interdependence to social order and identity fusion to social justice paths was a significantly better fit to the observed data than the baseline model ($\Delta\chi^2(2) = 39.07$, $p < .001$). We found that controlling for these unpredicted paths increased the predictive validity of our predicted paths. That is, controlling for identity fusion strengthened the relationship between interdependence and social justice, whereas controlling for interdependence strengthened the relationship between identity fusion and social order. Therefore, we included these additional paths in all subsequent models.

We also investigated whether all of the effects associated with morality are necessary; it is possible that the effect of social justice or social order alone sufficiently accounts for equality and inclusion. We tested two alternative models in which we trimmed the paths from social justice or social order to equality and inclusion. Interestingly, a model without the effects of social justice ($\Delta\chi^2(2) = 356.74$, $p < .001$) or social order ($\Delta\chi^2(2) = 95.25$, $p < .001$) was a significantly worse fit to the observed data. This suggests that these relationships are, in fact, important.

Next, we tested for moderation by experimental condition. We did not expect the measurement equations (e.g., error variances and factor loadings) to vary between experimental conditions, but a model in which we allowed these parameters to vary between groups was a significantly better fit to the observed data ($\Delta\chi^2(32) = 77.18$, $p < .001$). We tested a series of models to identify whether some or all of these measurement equations were moderated by the experimental manipulation. Our analyses revealed that a model with variant measurement for interdependence, social justice, and equality was a

significantly better fit to the observed data ($\Delta\chi^2(17) = 61.69$, $p < .001$). In contrast, a model with variant measurement for identity fusion, social order, and inclusion was not a significantly better fit to the observed data ($\Delta\chi^2(15) = 15.76$, $p = .399$). We retained variant measurement for interdependence, social justice, and equality in subsequent models because it is important to account for differential item functioning.

We then tested for the predicted moderation effect by allowing the paths from interdependence to social justice and identity fusion to social order to vary between experimental conditions. We also allowed the following associated parameters to vary between groups: the variance of interdependence, the variance of identity fusion, the disturbance variance of social justice, the disturbance variance of social order, the covariance between interdependence and identity fusion, and the disturbance covariance between social justice and social order. This model was a significantly better fit to the observed data than the model with only variant measurement ($\Delta\chi^2(8) = 30.13$, $p < .001$), which suggests that the experimental manipulation moderated these predicted paths.

We also tested a series of alternative models in which we allowed unpredicted paths to vary between experimental conditions against the prior predicted model. A model with variant paths from interdependence to social order and identity fusion to social justice was not a significantly better fit to the observed data ($\Delta\chi^2(2) = 2.69$, $p = .261$). Nor were models with variant effects emanating from social justice ($\Delta\chi^2(5) = 2.62$, $p = .758$) or social order ($\Delta\chi^2(5) = 5.08$, $p = .406$). This suggests that the experimental manipulation did not moderate these other relationships.

As can be seen in Table 1 (see Appendix B), our best fitting model had partially variant measurement, added paths from interdependence to social order and identity

fusion to social justice, and variant paths from interdependence to social justice and identity fusion to social order. Although the model chi-square was significant ($\chi^2(306) = 748.37$, $p < .001$), the values for the other recommended fit indices were pretty good (*RMSEA* = .066, *CFI* = .970, $SRMR_1$ = .077, $SRMR_2$ = .076). We tested a variety of nested alternative models to get this final model, but we also wanted to test a non-nested alternative model. It is possible that group dynamics are the most proximate to our outcome variables of equality and inclusion, and that morality is actually the antecedent driving how people feel about the group. We tested a model that was structurally similar to our best fitting model, except we switched interdependence with social justice and identity fusion with social order. However, this alternative model appeared to be a much worse fit to the observed data (*AIC* = 1307.44) relative to our final model (*AIC* = 864.27). We retained our final best fitting model, and not this alternative model, for analysis.

An inspection of the parameter estimates revealed that all of our predictions were supported (for factor parameters, see Table 2 in Appendix B; for indicator parameters, see Table 3 in Appendix B). Interdependence was positively associated with social justice but did not predict social order. In contrast, identity fusion was positively associated with social order and had a weak, negative association with social justice. Critically, the experimental manipulation of solidarity strengthened both of these relationships. The moderation of social justice appeared to be stronger, as these predictors explained more of the variance in social justice in the strong solidarity condition (27.67%) relative to the weak solidarity condition (12.76%). These predictors explained only a little more of the variance in social order in the strong solidarity condition (58.35%) relative to the weak solidarity condition (52.40%).

Social justice was a very strong, positive predictor of both equality and inclusion, whereas social order was a weak, negative predictor. In fact, morality explained an impressive 69.14% of the variance in equality and 46.46% of the variance in inclusion averaging across experimental conditions. Finally, our analyses revealed that there was a significant total indirect effect of interdependence on both equality (b = .602, SE = .079, t = 7.62, p < .001) and inclusion (b = .416, SE = .061, t = 6.86, p < .001) mediated through morality in the strong solidarity condition. These indirect effect on equality (b = .297, SE = .074, t = 4.01, p < .001) and inclusion (b = .204, SE = .051, t = 3.98, p < .001) were smaller but still significant in the weak solidarity condition. All of these effects can be seen in the figures below, which depict the unstandardized parameter estimates in both the strong solidarity and weak solidarity conditons.

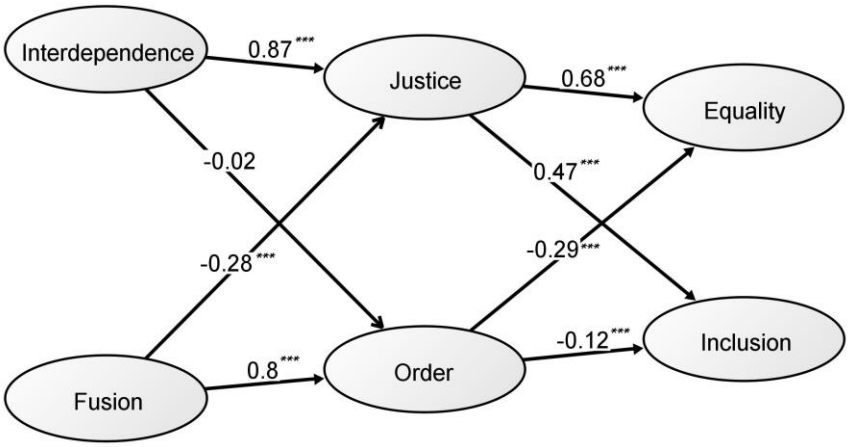

Figure 2: Unstandardized Structural Model in the Strong Solidarity Condition

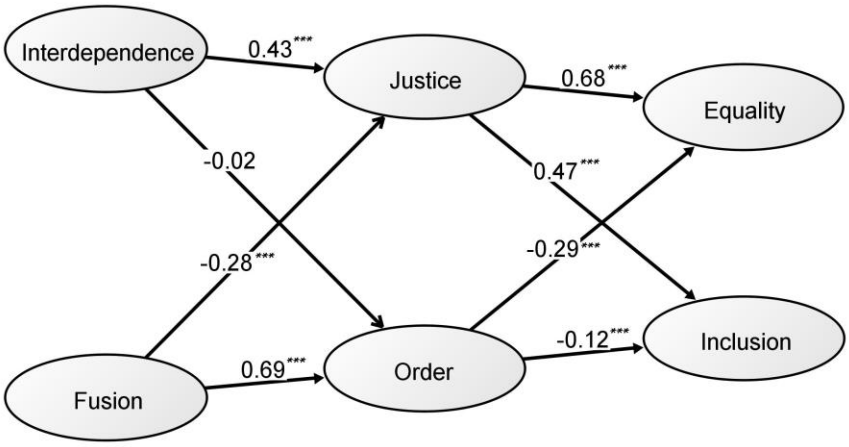

Figure 3: Unstandardized Structural Model in the Weak Solidarity Condition

Discussion

As predicted, interdependence was positively associated with social justice while identity fusion was positively associated with social order. This conceptually replicates our Study 1 findings in an intragroup context (as opposed to an intergroup context). Additionally, the powerful relationship between identity fusion and social order replicates past research with these constructs (see Carnes & Lickel, under review). Interestingly, our experimental manipulation of solidarity caused these relationships to grow stronger, but had no effect on other relationships. As theorized in past work, both social justice and social order appear to be binding moral principles that rely on different group properties to unify people (see Janoff-Bulman & Carnes, 2013).

A particular strength of this study was the use of multiple indicators and latent analytic methods to assess how people feel about equality and inclusion. As can be seen in Table 3, these different measures appeared to be valid and (mostly) reliable indicators of their respective latent constructs. Social justice was a strong predictor of support for

equality, and there was indeed a robust indirect effect of interdependence on equality through morality. Despite its parochial roots, social justice also predicted greater inclusivity (with its indicators of identification with all of humanity and an expansive circle of moral regard). In contrast to social justice, social order had independent, negative effects on support for equality and inclusion. One might not expect social order to have independent effects on these particular outcome variables, but the direction of these effects was sensible. In the next study, we further explored these relationships with real behaviors.

Notes

[5] These items emphasize common goals, shared interests, mutual dependence, and common fate based on past work on entitativity (e.g., Brewer, Hong, & Li, 2004), solidarity (e.g., Durkheim, 1893/1984), and interdependence theory (e.g., Kelley et al., 2003). These items were sufficiently reliable in a pilot study (α = .877).

[6] The Gini index is a measure of equality bounded between zero (perfect equality) and one (perfect inequality). The Gini index is the average absolute difference between all pairs of individuals divided by twice the mean, or equivalently it is twice the area between the Lorenz curve and the line of perfect equality (Allison, 1978).

[7] We removed the following targets: U.S. president, U.S. soldier, charity worker, chicken, fish, bee, redwood tree, apple tree, rose bush. We added the following targets: undocumented immigrant, person of different race, welfare recipient. The wording of some targets was slightly revised (e.g., intellectually disabled person instead of mentally challenged person).

[8] We excluded these covariates from analysis for the time being because the questions (and answers) are somewhat ambiguous. Participants described their income (ones versus thousands) and education (total versus after high school) in different metrics. We can infer what many participants meant and recode their responses, but the more ambiguous answers may have to be deleted.

[9] It is common for models with large samples or multigroup designs to have significant chi-square values even if the model fits the observed data (Kline, 2015).

CHAPTER 4

STUDY 3

Overview

We designed Study 3 to replicate and extend the findings of the prior two studies using yet another methodological approach. More specifically, we wanted to study the relationship between group dynamics and moral principles in the controlled setting of the lab with experimentally contrived groups. Behavioral economic games provide an ideal vehicle to test this question because the group is small enough to influence underlying group dynamics, and behavior in the game might represent morality in action. The essence of social justice is a concern for the welfare of the group, which clearly fits with the definition of cooperation—to pay a personal cost to produce a collective benefit (e.g., Rand & Nowak, 2013). Similarly, the essence of social order is a concern for conformity and obedience in the group, which clearly fits with the definition of coordination—to match or potentially mismatch on the actions of others (e.g., Thomas, DeScioli, Haque, & Pinker, 2014). Therefore, we asked participants to play a public goods game (i.e., a cooperation problem) and a matching game (i.e., a coordination problem) to assess moral behavior in real social groups.

An especially novel contribution of this study is our experimental manipulation of interdependence in which participants' payoffs were randomly shuffled at the end of the game. In effect, participants in this condition were actually interdependent because their outcomes in the game depended just as much on the behavior of others as themselves. Interestingly, this echoes Rawls' (1971/2009) original position wherein members of a society must decide how to organize the distribution of rights and resources behind a

"veil of ignorance" (i.e., without knowing their ultimate position or status in society). We also experimentally manipulated the norms in each game using a more typical external assignment condition (e.g., Van Huyck, Gillette, & Battalio, 1992) that told participants what to do. Our prediction was that the interdependence manipulation would promote cooperation (i.e., social justice) and the norm manipulation would promote coordination (i.e., social order).

Method

Participants. We recruited 249 participants from the University of Massachusetts Amherst subject pool in exchange for extra credit and prizes (28.5% male; median age of 20). A majority of the participants identified as White (66.7%), and the remaining identified as Asian (15.9%), Black (6%), Hispanic (4%), Multiracial (3.2%), or nominated another racial category (3.2%). No participants were excluded from analysis. A power analysis revealed that this sample had sufficient power ($\rho = .80$) under realistic assumptions to detect a medium to small effect of our principal hypothesis ($\delta = .31$) (see Raudenbush et al., 2011).

Procedure. Participants completed the experiment in groups of three to five people sitting at separated desks facing a wall. We randomly assigned these groups to one of three experimental manipulations (control, norm, or interdependence) and then participants played the public goods game and the matching game in counterbalanced order. These economic games were conducted using the z-Tree software package (Fischbacher, 2007). Participants read detailed directions[10] about each of these games immediately before playing. The experimenter read additional directions to the group depending on their experimental condition, and confirmed that each participant

understood the rules prior to beginning. Participants could exchange points earned in the economic games for a tiered system of prizes including candy, novelties (e.g., Frisbees, sunglasses, etc.), or a raffle for one of several $50 gift cards[11]. After playing both economic games, participants completed the morality measure and demographics, including checks on understanding and familiarity[12].

Experimental Manipulations. We randomly assigned each group of participants to the control condition, the norm condition, or the interdependence condition. In the control condition, the experimenter did not provide extra directions about the game; however, the experimenter read additional directions aloud before each game in the other two conditions (see Appendix A for manipulations). In the norm condition, the experimenter explained that there was an informal rule in the public goods game to contribute most of their endowment, and an informal rule in the matching game to pick the middle number; this norm tells participants "what most people should do in the game." In the interdependence condition, the experimenter explained that he or she would randomly shuffle participants' final payoffs at the end of the experiment; this rule tells participants that they "are just as likely to get someone else's payoffs as their own."

Public Goods Game. Each player received an endowment of 20 points, and could contribute any amount of that endowment to the common pool. All contributions to the common pool were increased by an efficiency factor of 1.6 and divided equally among the players. Players could see how much they had contributed, how much they had earned, and how much other players had contributed in each round; however, players could not identify which actual person was associated with which player number. Participants played 13 rounds, including one for practice. The directions framed the game

as “an example of a generic situation in which a group of people can try to achieve a common goal, but people in the group have the option of only benefitting themselves.” The primary outcome variable in this game was the number of points each player contributed to the common pool.

Matching Game. Each player could pick any number between one and seven. Everyone earned 32 points if all players matched on the same number; otherwise, no one earned any points. Players could see what number they had picked, how much they had earned, and what number other players had picked in each round; however, players could not identify which actual person was associated with which player number. Participants played 13 rounds, including one for practice. The directions framed the game as “an example of a generic situation in which a group of people are trying to reach the same outcome, but people in the group don't yet know how to get there.” The primary outcome variable in this game was a dummy code indicating whether each player matched the median number selected in the group.

Morality Measure. We measured social justice (α = .819) and social order (α = .715) with five items each (see appendix A for items), which were presented in counterbalanced order on a single page. Participants rated these items using 7-point scales anchored from strongly disagree (1) to strongly agree (7). This measure was adapted from Janoff-Bulman and Carnes (2016a).

Results

We analyzed the data using hierarchical linear modeling because the data had a nested structure with three levels (e.g., observations nested within individuals nested within groups). That is, the level-1 units were the decisions participants made in the

games across 12 rounds of play (i.e., observations), the level-2 units were the individuals making the decisions (i.e., individuals), and the level-3 units were the experimental groups that participants played the games in (i.e., groups). We used full maximum likelihood estimation for the public goods game data and adaptive Gaussian quadrature for the matching game data.

Public Goods Game

We first calculated the intraclass correlation for contributions to the common pool to assess the proportion of variance that lies between observations, individuals, and groups. This is similar to the procedure in Study 1, but with three levels of nesting (see Raudenbush & Bryk, 2002). An unconditional model with random intercepts revealed that differences between observations—representing the underlying dynamics of the game itself—account for about half of the variance in contributions (50.42%). The rest of the variance, however, lies between both individuals (28.65%) and groups (20.93%).

Next, we estimated participants' baseline behavior in the game by including the round of play (1 – 12) as a level-1 linear predictor with random slopes at both level-2 and level-3. We also included understanding, familiarity, group size, and game order as covariates[13]. A participant in a typical group contributed 12.20 points (SE = .63, $t(59)$ = 19.25, $p < .001$) to the common pool in round one, but there was a significant amount of variation around this intercept at the individual level and the group level (all $p < .001$). These contributions declined over the course of the game (γ = -.33, SE = .079, $t(59)$ = -4.22, $p < .001$), but there was again significant variation around this slope at the individual level and the group level (all $p < .001$).

We then added social justice and social order as grand-mean centered level-2 predictors to test whether these individual differences affect cooperation[14]. The model revealed that while social justice was associated with marginally higher initial contributions (γ = .52, SE = .29, $t(117) = 1.78$, p = .077), social order was not (γ = .23, SE = .31, $t(117) = .75$, p = .454). Furthermore, social justice attenuated the decline in contributions over time (γ = .081, SE = .032, $t(117) = 2.51$, p = .013), while social order did not (γ = -.014, SE = .034, $t(117) = -.42$, p = .673). These individual differences improved model fit ($\Delta\chi^2(4) = 20.43$, $p < .001$), explaining 3.00% of the variance in the initial intercept, 10.68% of the variance in the endgame intercept, and 6.75% of the variance in the slope at level-2.

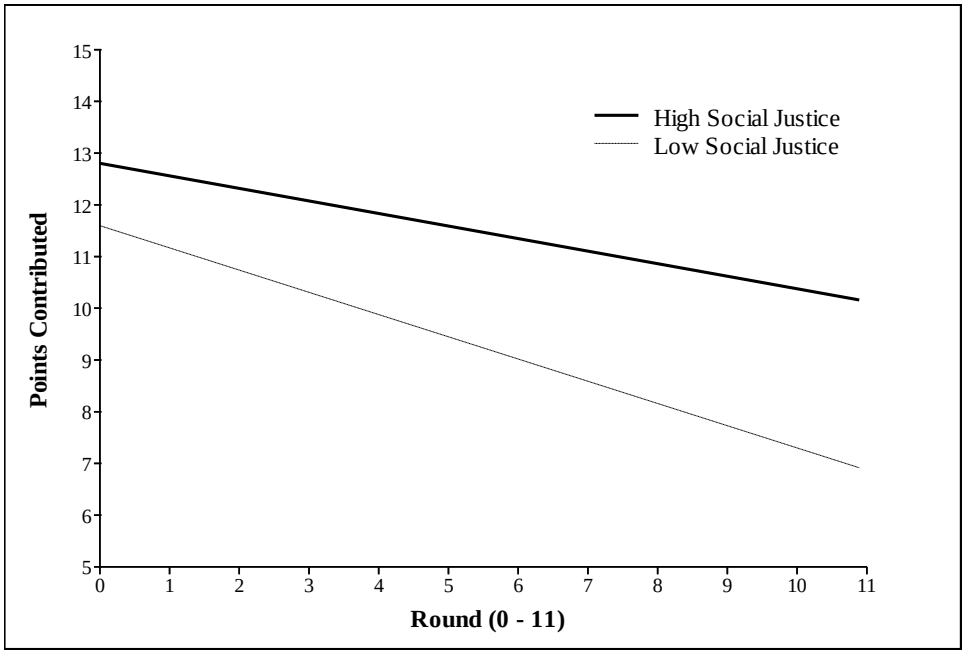

Figure 4: Effect of Social Justice on Cooperation.

Finally, we added the experimental manipulations as level-3 predictors to test whether these group dynamics affect cooperation (see Appendix C for this model's equation). The model revealed that initial contributions were marginally higher in the

norm condition (γ = 2.02, *SE* = 1.01, $t(57) = 1.99$, $p = .052$) relative to the control condition, but were not significantly different in the interdependence condition (γ = 1.43, *SE* = 1.02, $t(57) = .87$, $p = .169$). In contrast, the decline in contributions over time was significantly attenuated in the interdependence condition (γ = .26, *SE* = .13, $t(57) = 2.05$, $p = .045$) relative to the control condition, but was not significantly attenuated in the norm condition (γ = .12, *SE* = .13, $t(57) = .93$, $p = .359$). These experimental manipulations improved model fit ($\Delta\chi^2(4) = 9.61$, $p = .047$) and had a bigger effect size, explaining 13.43% of the variance in the initial intercept, 18.77% of the variance in the endgame intercept, and 11.20% of the variance in the slope at level-3.

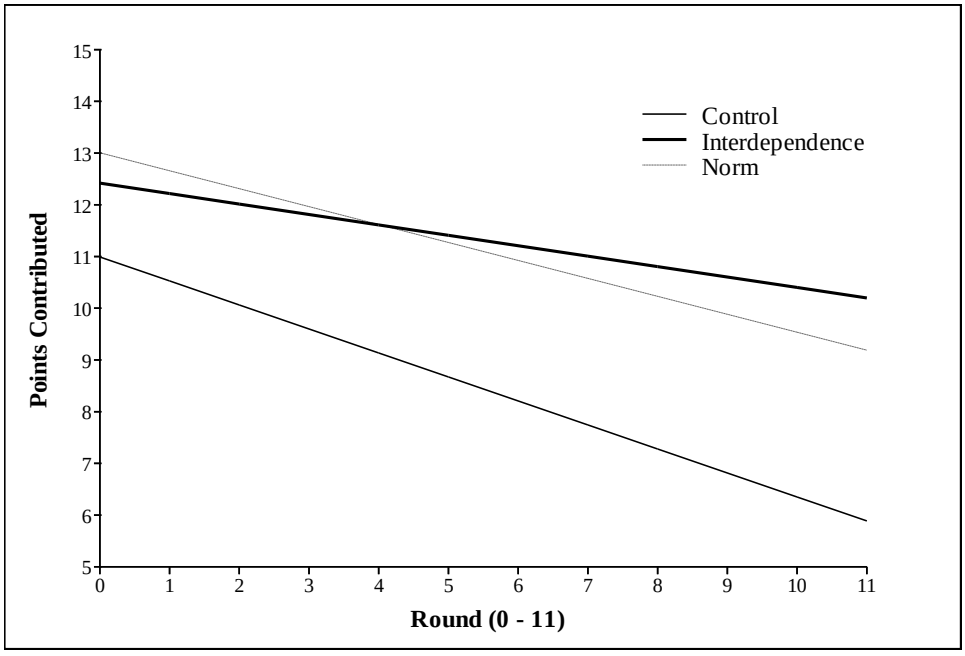

Figure 5: Effect of the Experimental Manipulations on Cooperation.

Matching Game

We used a logit link function to represent the likelihood of matching because the outcome variable was nominal. We first calculated the intraclass correlation for the likelihood of matching to assess the proportion of variance that lies between individuals

and groups. An unconditional model with random intercepts revealed that differences between groups accounted for most of the variance (82.29%) while differences between individuals accounted for a relatively small share of the variance (17.71%)[15].

Next, we estimated participants' baseline performance in the game by including the round of play (1 – 12) as a level-1 linear predictor with random slopes at both level-2 and level-3. We also included understanding, familiarity, group size, and game order as covariates[16]. A participant in a typical group matched the median number picked 51.91% of the time in round one ($\gamma = .036$, $SE = .38$, $t(59) = .096$, $p = .924$), but there was significant variation around this intercept at both the individual level and the group level ($p \leq .016$). In addition, the likelihood of matching increased each round ($\gamma = .57$, $SE = .093$, $t(59) = 6.10$, $p < .001$), but there was significant variation around this slope at the group level ($p < .001$) and not at the individual level ($p > .500$).

We then added social order and social justice as grand-mean centered level-2 predictors to test whether these individual differences affect coordination[17]. Surprisingly, the model revealed that social justice was associated with marginally higher chances of matching in round one ($\gamma = .25$, $SE = .14$, $t(117) = 1.73$, $p = .086$). There was also a trending effect of social order on the likelihood of matching over time ($\gamma = .040$, $SE = .030$, $t(117) = 1.40$, $p = .164$), but the remaining effects were not significant (all $p \geq .348$). These individual differences explained 4.01% of the level-2 variance in matching when assessed at the beginning of the game.

Finally, we added the experimental manipulations as level-3 predictors to test whether group dynamics affect coordination (see Appendix C for this model's equation). A participant in a typical group in the norm condition matched 77.65% of the time in

round one, which was significantly more than the 35.23% chance in the control condition (γ = 1.85, *SE* = .62, *t*(57) = 2.99, *p* = .004) and the 35.08% chance in the interdependence condition (γ = 1.86, *SE* = .61, *t*(57) = 3.07, *p* = .003). The likelihood of matching increased over time in all three conditions, but this rate was trending slower in both the norm condition (γ = -.25, *SE* = .15, *t*(57) = -1.69, *p* = .096) and the interdependence condition (γ = -.21, *SE* = .14, *t*(57) = -1.47, *p* = .146) relative to the control condition. These experimental manipulations explained 20.88% of the level-3 variance in matching when assessed at the beginning of the game.

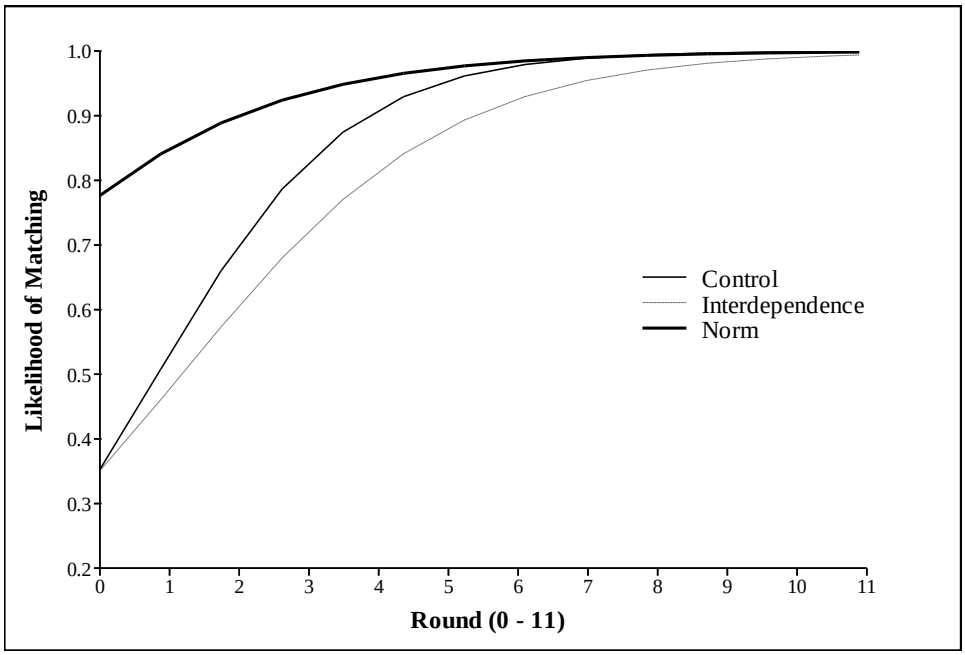

Figure 6: Effect of the Experimental Manipulations on Coordination.

Discussion

Our experimental manipulations were designed to simulate different group dynamics, and they had the predicted effects on moral behavior. Interdependence sustained cooperation over time in the public goods game and norms initiated coordination in the matching game. This conceptually replicates the findings from

Studies 1 and 2 using very different operationalizations, and allows us to make causal arguments because we had random assignment to experimental condition. We also found that self-reported endorsement of social justice predicted higher initial levels of cooperation and more sustained cooperation over time. It is interesting that attitudes about social justice in a societal context would generalize to behavior in a small group context, but the fact that they did strengthens our claim that cooperation is a form of social justice in action.

Our findings also illustrate an important point about enacting social justice. One might be tempted to impose social justice on a group by simply telling people what to do. However, this is precisely what we did in the norm condition, and although initial levels of cooperation improved, contributions to the common pool still collapsed over the course of the game just as in the control condition. This suggests that simply telling people to share or help might be an ineffective strategy; we need to build community and strengthen the underlying connections between people if we want to promote social justice over the long haul. Social justice is not simply an abstract moral principle; it is embedded in the structure or character of groups (see Carnes et al., 2015).

Interestingly, imposing a "veil of ignorance" (i.e., the interdependence manipulation) promoted our operationalization of social justice in the public goods game. The logic of Rawls' (1971/2009) thought experiment uses self-interest against itself—what would a reasonable person do if they cannot privilege their group or themselves? Structure society based on the principles of equality and justice so that it benefits the least advantaged members. Our manipulation worked in a similar manner, structuring the incentives of the game so that self-interest serves the common interest, but we note that

the original position also makes people interdependent. As Rawls might have predicted, the proportion of truly cooperative 20-point contributions rose from 19.8% in the control condition to 35.2% in the interdependence condition (and 26.8% in the norm condition).

There were, however, some problems with this study. First, many of the effects were relatively modest in their level of significance and size, suggesting that we need to replicate our findings. Second, there was a strong order effect wherein participants cooperated less and matched more in their second game. This suggests that our replication should include only one economic game. Third, based on participants' self-reported understanding, it was also clear that the directions could be made clearer in this replication study. It is important that participants understand the social problem confronting them in order for their moral behavior to be inferentially valid. Finally, we need more measures of psychological process to complete the conceptual leap from our first set of studies to our second set of studies. We give labels to our experimental manipulations, but we need evidence that they induce the psychological processes predicted from theory. We implemented these changes in Study 4.

Notes

[10] The directions described the rules for each game using neutral, simple language as is typical in the literature. We also provided a table of potential payoffs to illustrate how players' choices interact to determine the outcome of the games. In the public goods game, this table showed that "you earn more points when the common pool is larger. However, you can also earn more points by keeping more of your own endowment. The

same is true for all of the other players." In the matching game, this table showed that "you only earn points by matching on any of the seven numbers. You do not get any points if you fail to match on a number. The same is true for all of the other players."

[11] Past research suggests that increasing stakes, even up to several months' wages, has a surprisingly weak effect on behavior in economic games (e.g., Cameron, 1999; List & Cherry, 2000; Slonim & Roth, 1998).

[12] Most participants were unfamiliar with economic games (85.1%), but understood the public goods game (89.2%) and the matching game (98%).

[13] We included these variables as covariates, rather than cutting participants, for power reasons. No random effects were added for any covariates. Covariates were centered so that the intercepts represented individuals in 4-person groups who played the public goods game first, understood the game, and were not familiar with it from before. Participants' initial contributions were smaller when they played the public goods game second (γ = -3.05, *SE* = .87, $t(59) = -3.53$, $p < .001$) or did not understand the game (γ = -2.52, *SE* = 1.07, $t(121) = -2.34$, $p = .021$). These covariates did not have any additional significant effects (all $p \geq .137$).

[14] We did not include random effects for these variables because they should theoretically reflect individual differences.

[15] We cannot calculate the proportion of variance between level-1 units with this estimation method.

[16] We included these variables as covariates, rather than cutting participants, for power reasons. No random effects were added for any covariates. Covariates were centered so that the intercepts represented individuals in 4-person groups who played the matching game first, understood the game, and were not familiar with it from before. Participants were 38.39% more likely to match in round one when playing the matching game second (γ = 2.09, *SE* = .56, $t(59) = 3.75$, $p < .001$), and were marginally more likely to match as the rounds progressed when they were familiar with economic games (γ = .18, *SE* = .099, $t(121) = 1.79$, $p = .076$). These covariates did not have any additional significant effects (all $p \geq .313$).

[17] We did not include random effects for these variables because they should theoretically reflect individual differences.

CHAPTER 5

STUDY 4

Overview

We designed Study 4 to overcome the methodological problems identified in Study 3; participants only played one economic game (i.e., the public goods game), we clarified the directions for this game, and we added more psychological measures. We also replaced the norm condition with a novel experimental manipulation of prosociality. Participants in this condition were given a bonus for cooperating—making it more rewarding for them—and led to believe that they alone were receiving this bonus. The logic here is that prosocial individuals cannot necessarily infer that other people also have prosocial tendencies. We predicted that only the interdependence manipulation would sustain cooperation in the public goods game, replicating the results of our prior studies. Additionally, we predicted that this sustained cooperation would make people feel interdependent with their group, complementing past research suggesting that cooperation facilitates the formation of social ties (Sonnemans, van Dijk, van Winden, 2006; van Dijk, Sonnemans, & van Winden, 2002). This would suggest that there is a bidirectional relationship (or feedback loop) between interdependence and social justice.

Method

Participants. We recruited 299 participants from the University of Massachusetts Amherst subject pool in exchange for extra credit and prizes (16.5% male; median age of 20). A majority of the participants identified as White (66%), and the remaining identified as Asian (18.1%), Black (5.5%), Hispanic (3.9%), Multiracial (3.9%), or nominated another racial category (2.6%). No participants were excluded from analysis.

A power analysis revealed that this sample had sufficient power ($\rho = .80$) under realistic assumptions to detect a medium to small effect of our principal hypothesis ($\delta = .29$) (see Raudenbush et al., 2011).

Procedure. Similar to Study 3, participants completed the experiment in groups of three to five people sitting at separated desks facing a wall. We randomly assigned these groups to one of three experimental manipulations (control, prosocial, or interdependence) and then participants played the public goods game. Unlike Study 3, the experimental manipulations were embedded in the detailed directions presented immediately prior to play. In addition, the experimenter administered a questionnaire both before and after the game. Points earned in the economic game directly increased participants' chances of winning one of several $20, $50, or $100 gift cards. After answering the post-game measure, participants completed the morality measure and demographics, including checks on understanding and familiarity[18].

Experimental Manipulations. We randomly assigned each group of participants to the control condition, the prosocial condition, or the interdependence condition. The prosocial and interdependence manipulations were embedded in the directions for the public goods game (see Appendix A for manipulations). No additional directions were provided in the control condition. In the interdependence condition, participants were told that their payoffs would be shuffled at the end of the game. In the prosocial condition, participants were told that they alone would receive a bonus to their payoff for every point they invested in the common pool.

Public Goods Game. We made two small changes to this economic game, but it was otherwise the same as in Study 3. First, participants played for 15 rounds and did not

have a practice round. Second, the directions were clarified by providing examples, using analogies, and simplifying language (see Appendix A for directions).

Pre-Game Measure. We asked participants what they expect the other players in the economic game to be like. Participants rated a series of adjectives using 7-point scales anchored from strongly disagree (1) to strongly agree (7). Prosociality was assessed with the following five items (α = .877): kind, giving, generous, helpful, and nice. Interdependence was assessed with the following seven items (α = .746): interdependent with me, considerate, share my goals, dependable, connected to me, trustworthy, and in the same "boat" as me.

Post-Game Measure. We asked participants what they thought of the other players in the economic game. Participants rated a series of adjectives using 7-point scales anchored from strongly disagree (1) to strongly agree (7). Prosociality (α = .934) and interdependence (α = .859) were assessed with the same items as in the pre-game measure. Using the same rating scale, we also assessed participants' support for different distributional principles using four items. These distributional principles were equality, justice, and equity ($r(297)$ = .725) (see Appendix A for items)[19].

Morality Measure. We measured social justice (α = .795) and social order (α = .674) with the same items used in Study 3.

Results

Public Goods Game

Just as in Study 3, we analyzed the data using hierarchical linear modeling with full maximum likelihood estimation because the data had a nested structure with three levels (e.g., observations nested within individuals nested within groups). We first

calculated the intraclass correlation for contributions to the common pool to assess the proportion of variance that lies between observations, individuals, and groups as in Study 3. An unconditional model with random intercepts revealed that differences between observations account for about half of the variance in contributions (49.94%). The next largest share of variance lies between groups (35.40%) and the smallest share of variance lies between individuals (14.66%).

Next, we estimated participants' baseline behavior in the game by including the round of play (1 – 15) as a level-1 linear predictor with random slopes at both level-2 and level-3. We also included understanding, familiarity, and group size as covariates[20]. A participant in a typical group contributed an estimated 9.90 points (SE = .51, $t(72)$ = 19.44, $p < .001$) to the common pool in round one, but there was a significant amount of variation around this intercept at the individual level and the group level (all $p < .001$). Contributions then declined over the course of the game (γ = -.18, SE = .041, $t(72)$ = -4.34, $p < .001$), but there was again significant variation around this slope at the individual level and the group level (all $p < .001$).

We then added social justice and social order as grand-mean centered level-2 predictors to test whether these individual differences affect cooperation[21]. Our analyses revealed that while social justice was associated with significantly higher initial contributions to the common pool (γ = .68, SE = .23, $t(143)$ = 2.94, $p < .001$), it did not have a significant effect on the rate of contribution decline over time (γ = -.027, SE = .021, $t(143)$ = -1.31, p = .194). In contrast, social order was associated with significantly lower initial contributions to the common pool (γ = -.63, SE = .030, $t(65)$ = -2.08, p = .042) and had a marginally slower rate of contribution decline (γ = .040, SE = .023, $t(143)$

= 1.77, p = .079). These individual differences improved model fit ($\Delta\chi^2(4) = 14.31$, p = .007) and explained a small to moderate amount of the variance in the initial intercept (7.22%), endgame intercept (0.92%), and slope (4.90%) at level-2.

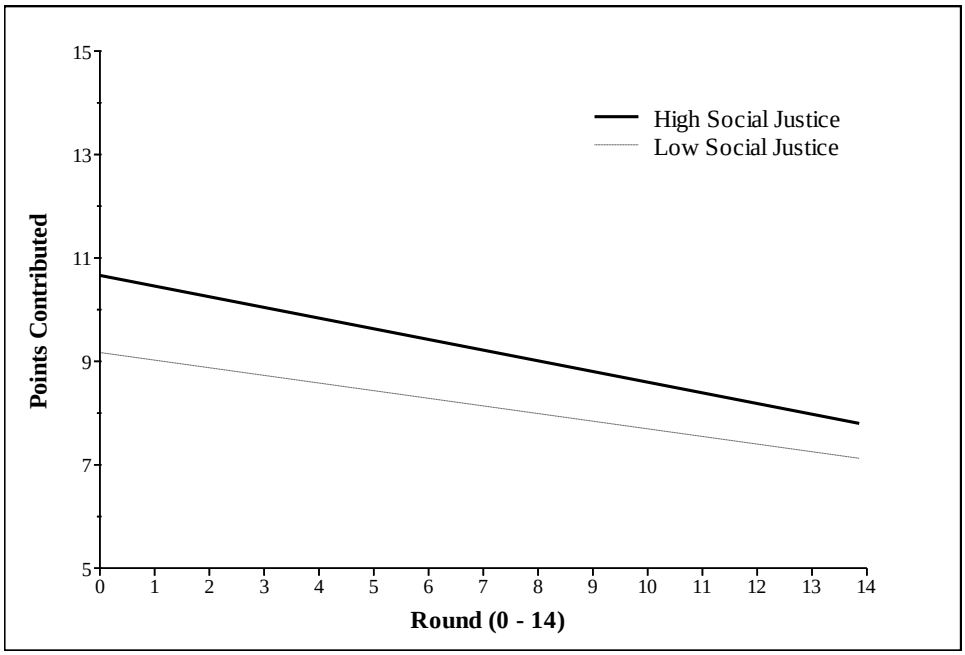

Figure 7: Effect of Social Justice on Cooperation.

Before further elaborating on our multilevel model, we conducted one-way ANOVAs testing for differences between the experimental conditions in both pre-game interdependence and prosociality pooled across level-2 units; that is, we examined the average pre-game ratings of groups and not individuals. As predicted, the experimental conditions differed significantly in their ratings of pre-game interdependence ($F(2, 71)$ = 5.91, p = .004, η_p^2 = .143) but not pre-game prosociality ($F(2, 71)$ = .79, p = .458). Post-hoc contrasts revealed that ratings of pre-game interdependence were significantly higher in the interdependence condition (M = 4.93) relative to either the prosocial condition (b = .39, SE = .11, p < .001) or the control condition (b = .23, SE = .11, p = .049).

We then added the experimental manipulations as level-3 predictors to test whether group dynamics affect cooperation (see Appendix C for this model's equation). The model revealed that participants' initial contributions were significantly higher in the prosocial condition relative to the control condition (γ = 2.39, *SE* = 1.13, $t(70) = 2.12$, p = .037), but were not significantly different in the interdependence condition (γ = .95, *SE* = 1.11, $t(70) = .86$, $p = .394$). In contrast, the decline in contributions over time was significantly attenuated in the interdependence condition relative to the control condition (γ = .24, *SE* = .086, $t(70) = 2.80$, $p = .007$), but was not significantly attenuated in the prosocial condition (γ = .0081, *SE* = .087, $t(70) = .093$, $p = .926$). These experimental manipulations improved model fit ($\Delta\chi^2(4) = 13.94$, $p = .008$) and had a bigger effect size, explaining 7.69% of the variance in the initial intercept, 11.50% of the variance in the endgame intercept, and 17.88% of the variance in the slope at level-3.

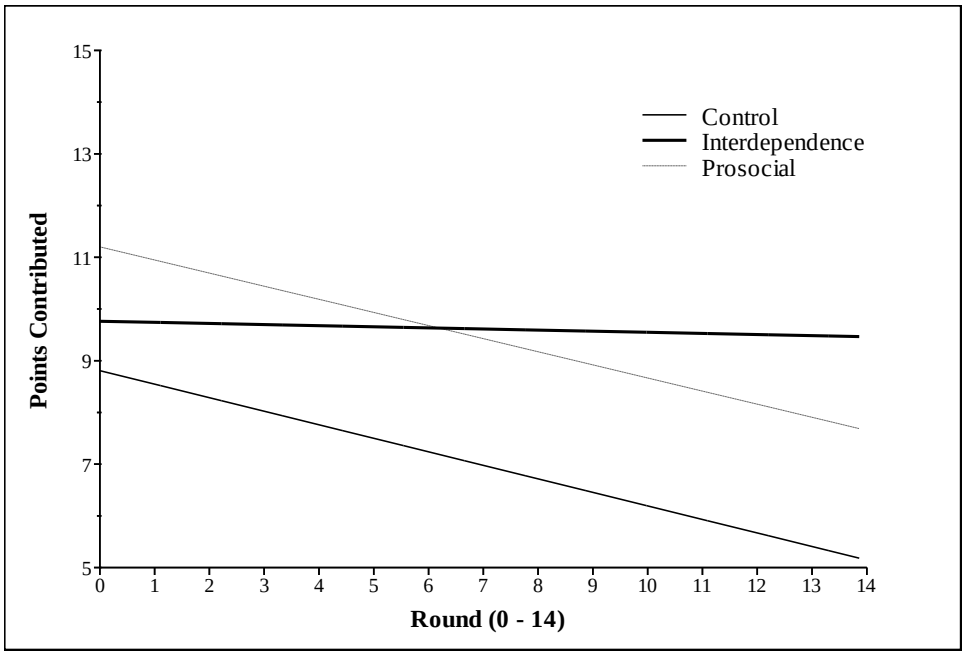

Figure 8: Effect of the Experimental Manipulations on Cooperation.

Post-Game Ratings

Next, we reran our baseline model twice—once with giving and once with getting as the outcome variable[22]—to estimate and export endgame intercept parameters for each individual; these parameters summarize the level of cooperation each participant displayed and experienced respectively at the end of the game. We then conducted a bootstrapping test of the serial indirect effect with the experimental manipulations (X) predicting post-game interdependence ratings (Y) mediated through both giving (M1) and getting (M2)[23]. Importantly, we controlled for pre-game ratings of interdependence on all outcomes. This analysis revealed a significant total indirect effect (*b* = .27, *SE* = .089, *CI*[.118, .469]) and a non-significant direct effect (*b* = .19, *SE* = .16, t(298) = 1.19, p = .237, *CI*[-.123, .497]) of the interdependence manipulation relative to the control condition. However, neither the total indirect effect (*b* = .12, *SE* = .064, *CI*[-.003, .250]) nor the direct effect (*b* = -.021, *SE* = .15, t(298) = -.13, p = .894, *CI*[-.322, .281]) were significant when comparing the prosocial manipulation to the control condition.

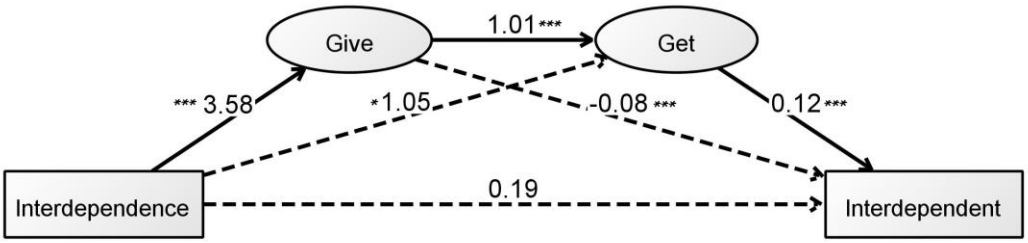

Figure 9: Indirect Effect of the Interdependence Manipulation on Post-Game Interdependence through Cooperation.

Finally, we tested a series of 2-level models in which the outcome was participants' rated support for one of the three different distributional principles[24]. The

level-1 units were the individuals making the ratings and the level-2 units were the experimental groups. We added social justice as a level-1 predictor with random effects and post-game interdependence as a level-2 predictor in each of these models (see Appendix C for these models' equations). Both social justice (γ = .26, *SE* = .091, $t(73)$ = 2.87, p = .005) and post-game interdependence (γ = .30, *SE* = .12, $t(72)$ = 2.56, p = .013) were positively associated with support for equality. However, only social justice was significantly associated with support for justice (γ = .20, *SE* = .086, $t(73)$ = 2.35, p = .021) or equity (γ = .19, *SE* = .086, $t(73)$ = 2.23, p = .029).

Discussion

Replicating Study 3, we again found that the interdependence manipulation sustained cooperation in the public goods game, but we were also able to show that this manipulation actually made people feel interdependent with their fellow group members. This is important because it shows that our experimental manipulation influenced what we theoretically predicted and not just prosociality. We did not replicate the effects of social justice on cooperation, finding that individuals who endorse this moral principle only cooperate more at the beginning of the game. Social justice did however predict the distributional principles of equality, justice, and equity controlling for feelings of interdependence; post-game ratings of interdependence only predicted support for equality. Cohering with these independent effects on the distributional principles, we also found no evidence of an interaction between individual differences in social justice and our experimental manipulation of interdependence in the public goods game[25].

Most importantly, we found compelling evidence for a feedback loop or virtuous cycle between interdependence and cooperation; there was a significant indirect effect of

the interdependence manipulation on post-game interdependence through cooperation controlling for pre-game interdependence. This finding helps us understand the emergence of social justice, operationalized as cooperation here. It is not simply that some precipitating event makes the group feel interdependent and suddenly people strongly endorse social justice; feelings of interdependence nudge people toward cooperative action, cooperative action shifts how they feel about the group, and—through many iterations of this virtuous cycle—social justice as a moral sentiment is nurtured or grown. Seeing that social justice is embedded in groups is the key to understanding when and why people care about equality.

A final novel finding involved disentangling giving (i.e., contributions to the common pool) from getting (i.e., compensation from the common pool) in the public goods game. Giving represents the level of cooperation each participant chose to display, while getting represents the level of cooperation each participant experienced. The latter was outside of participants' direct control—instead representing the actions of fellow group members—because we removed the part each participant contributed. We found that each point contributed was associated with about one point in compensation, suggesting that it pays to give because it induces others to give in reciprocation. Interestingly, an inspection of the parameter estimates in our mediational model revealed that—when controlling for one another's influence—getting was positively associated with post-game interdependence while giving was negatively associated. This suggests that what we get from others, and not just what we do for others, is important in fostering interdependence in groups. In Study 5, we investigate some of the conditions that could potentially break this virtuous cycle between interdependence and social justice.

Notes

[18] Most participants were unfamiliar with the public goods game (77.3%) but understood it (92.2%). This is similar to the rates we found in Study 3.

[19] We originally planned to create a composite with the equality and justice items, but they were only weakly correlated ($r(297) = .255$, $p < .001$).

[20] We included these variables as covariates, rather than cutting participants, for power reasons. No random effects were added for any covariates. Covariates were centered so that the intercepts represented individuals in 4-person groups who understood the game but were not familiar with it from before. Those familiar with this game contributed significantly more in the first round ($\gamma = 1.30$, $SE = .61$, $t(147) = 2.13$, $p = .035$) but their contributions declined at a marginally faster rate over the course of the game ($\gamma = -.099$, $SE = .054$, $t(147) = -1.85$, $p = .066$). These covariates did not have any additional significant effects (all $p \geq .743$).

[21] We did not include random effects for these variables because they should theoretically reflect individual differences.

[22] Compensation was simply participants' share of the common pool minus the part they contributed to the common pool.

[23] We used the PROCESS macro to generate bias-corrected bootstrapping 95% confidence intervals with 20,000 resamples (see Hayes, 2013).

[24] As before, we analyzed the data using hierarchical linear modeling with full maximum likelihood estimation because the data had a nested structure with two levels (e.g., individuals nested within groups).

[25] We freed the random effects associated with social justice and added the experimental manipulations as level-3 predictors of it. Adding these predictors did not improve model fit ($\Delta\chi^2(4) = 2.75$, $p > .500$). As can be seen in the graph below, the effect of social justice on the intercept is somewhat attenuated in the interdependence (γ = -.98, *SE* = .56, $t(71)$ = -1.75, p = .085) and prosocial (γ = -.81, *SE* = .72, $t(71)$ = -1.14, p = .260) conditions relative to the control condition.

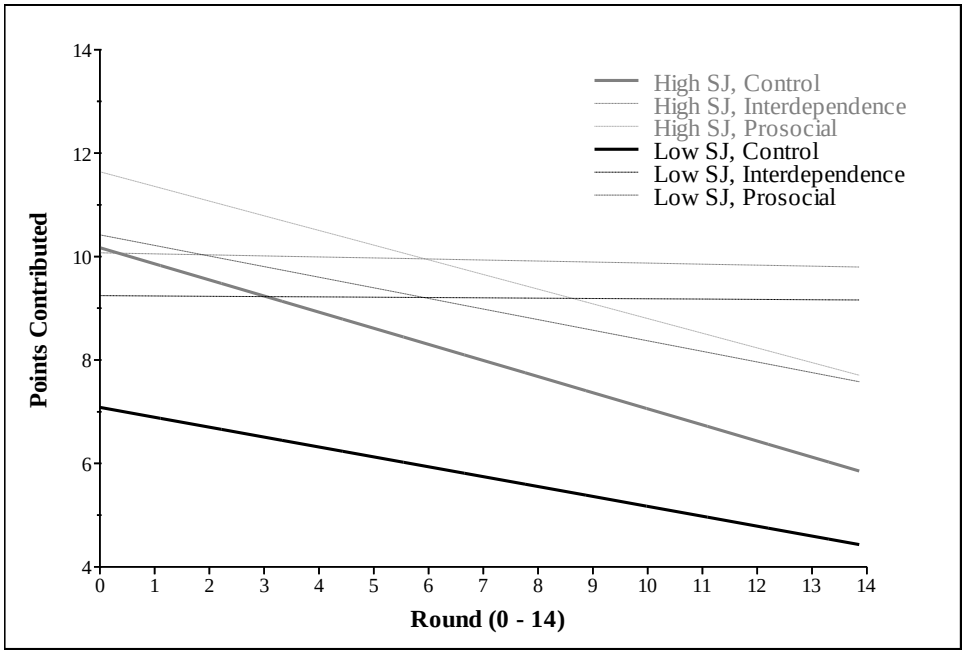

Figure 10: Effect of Social Justice and Interdependence on Cooperation.

CHAPTER 6

STUDY 5

Overview

Across four studies, and three very different methodological approaches, we find consistent evidence for a causal relationship between interdependence and social justice. In fact, we found evidence for a virtuous cycle between these constructs that helps us understand the ontogenesis of social justice as a moral sentiment. However, as noted by one of the participants in the comments section, "economic games are a metaphor for society." Even the most equal of societies are more unequal than the average experimental group in Study 3 (*Gini index* = .060, *SD* = .032) or Study 4 (*Gini index* = .046, *SD* = .025)[26]. This is because the structural design of a typical public goods game—with every participant receiving the same endowment and the same share of the public good—strongly pulls for equality. We designed Study 5 to assess the relationship between interdependence and social justice, but in the context of a structurally unequal society (or at least experimental group).

The design was similar to Study 4, but participants played an unequal public goods game in which one rich player received a larger endowment than the other players (*Baseline Gini Index* = .25). Past research strongly suggests that unequal endowments lead to lower levels of cooperation (Anderson, Mellor, & Milyo, 2008; Aquino, Steisel, & Kay, 1992; Cherry, Kroll, & Shogren, 2005; Colasante & Russo, 2016; Kroll, Cherry, & Shogren, 2007; Nishi, Shirado, Rand, & Christakis, 2015; Rapoport, Bornstein, & Erev, 1989; Tavoni, Dannenberg, Kallis, & Löschel, 2011). Inequality could potentially break the virtuous cycle between interdependence and social justice, but we predicted that the

interdependence manipulation would nonetheless overcome this obstacle and sustain cooperation in the public goods game—leading to lower levels of inequality.

We also replaced the prosociality condition with a more typical experimental manipulation of punishment. On the one hand, past research suggests that punishment is effective at sustaining cooperation (e.g., Fehr & Gächter, 2000, 2002; Ostrom, Walker, & Gardner, 1992), so we might expect inequality to be attenuated in the punishment condition (see also Masclet & Villeval, 2008). However, punishment might instead exacerbate inequality because punishment is associated with lower total payoffs (Dreber, Rand, Fudenberg, & Nowak, 2008) and the people most likely to engage in punishment are also the ones who care the most about equality (Johnson, Dawes, Fowler, McElreath, & Smirnov, 2009). It could also increase inequality if rich and poor players engage in counter-punishment (e.g., Nikiforakis, 2008) or antisocial punishment (e.g., Rand & Nowak, 2011) at different rates. For these reasons, we predicted that punishment might sustain cooperation in the public goods game but would ultimately exacerbate inequality.

Method

Participants. We recruited 328 participants from the University of Massachusetts Amherst subject pool in exchange for extra credit and prizes (24.1% male; median age of 20). A majority of the participants identified as White (60.1%), and the remaining identified as Asian (17.4%), Black (5.8%), Hispanic (3.7%), Multiracial (3.7%), or nominated another racial category (3%). No participants were excluded from analysis. A power analysis revealed that this sample had sufficient power ($\rho = .80$) under realistic assumptions to detect a medium to small effect of our principal hypothesis ($\delta = .27$) (see Raudenbush et al., 2011).

Procedure. The procedure was similar to Study 4. Participants completed the experiment in groups of three to four people sitting at separated desks facing a wall. We randomly assigned these groups to one of three experimental manipulations (control, punishment, or interdependence) and the experimenter administered a questionnaire both before and after the game. Unlike Study 4, we asked each participant to roll a pair of dice at the beginning of the experiment and to keep the result a secret; the participant with the highest result was assigned to the role of the wealthy player in an unequal public goods game, and everyone else was a poor player. Participants played this game for 15 rounds. After answering the post-game measure, participants completed the justice sensitivity measure and demographics, including checks on understanding and familiarity[27].

Experimental Manipulations. We randomly assigned each group of participants to the control condition, the punishment condition, or the interdependence condition. The punishment and interdependence manipulations were embedded in the directions for the public goods game (see Appendix A for the punishment manipulation). No additional directions were provided in the control condition. In the interdependence condition, participants were told that their payoffs would be shuffled at the end of the game just like in Study 3. In the punishment condition, participants were allowed to sanction other players at a cost to themselves after each round of play.

Unequal Public Goods Game. In this iteration of the public goods game, one player receives a larger endowment than the other players do—hence the unequal part of the name—but it was otherwise similar to the game played in Study 4. The poor players always received a 10-point endowment in each round; however, the rich player received a 25-point endowment in the 3-person version and a 30-point endowment in the 4-person

version[28]. The directions were slightly different for the two types of players, and explained why one player received a larger endowment than the other players did[29]. In the experimental condition with punishment, each round of the game included an extra phase. This phase allowed players to assign sanction points to one another. Each sanction point cost one payoff point to assign but destroyed three of the targeted player's payoff points[30]. Participants could see how many sanction points they received, which player sanctioned them, and the effect of these sanction points on their payoff total.

Pre-Game Measure[31]. This questionnaire was the same as in Study 4, except we added the following additional adjectives: moral, righteous, and virtuous. Prosociality (α = .885), interdependence (α = .735), and morality (α = .822) were all adequately reliable. We also assessed social justice (α = .827) and social order (α = .583) on this measure using the same items from Study 3 and Study 4.

Post-Game Measure. This questionnaire was the same as in Study 4, except we added the following additional adjectives: moral, righteous, and virtuous. Prosociality (α = .920), interdependence (α = .822), and morality (α = .924) were all adequately reliable. Using the same rating scale, we also assessed participants' support for two different distributional principles using six items. These distributional principles were equality (α = .722) and equity (α = .760) (see Appendix A for added items).

Justice Sensitivity Measure. This construct was assessed with eight items from an abbreviated version of the justice sensitivity inventory (Schmitt, Baumert, Gollwitzer, & Maes, 2010; Baumert et al., 2014). There are four types of justice sensitivity: victim sensitivity, perpetrator sensitivity, observer sensitivity, and beneficiary sensitivity. Justice sensitivity was adequately reliable (α = .803) (see Appendix A for items).

Results

Public Goods Game

As in the prior two studies, we analyzed the data using hierarchical linear modeling with full maximum likelihood estimation because the data had a nested structure with three levels (e.g., observations nested within individuals nested within groups). We first calculated the intraclass correlation for contributions to the common pool to assess the proportion of variance that lies between observations, individuals, and groups. An unconditional model with random intercepts revealed that differences between observations (41.15%) and between individuals (44.53%) accounted for most of the variance in contributions. Unlike Study 4, we found that the smallest share of variance lies between groups (14.32%).

Next, we estimated participants' baseline behavior in the game by including the round of play (1 – 15) as a level-1 linear predictor with random slopes at both level-2 and level-3. We also included group size as a level-3 covariate[32]. A participant in a typical group contributed an estimated 6.45 points (SE = .34, $t(87) = 18.91$, $p < .001$) to the common pool in round one, but there was a significant amount of variation around this intercept at the individual level and the group level (all $p < .001$). Contributions then declined over the course of the game (γ = -.091, SE = .024, $t(87) = -3.81$, $p < .001$), but there was again significant variation around this slope at the individual level and the group level (all $p < .001$).

We then added wealth (rich or poor) as a level-2 predictor with random effects. A typical poor player contributed an estimated 5.17 points (SE = .27, $t(87) = 19.51$, $p < .001$) to the common pool in round one, whereas a typical rich player initially contributed

an estimated 9.56 points (*SE* = .68, $t(87) = 14.12$, $p < .001$). Rich and poor players' trajectories did not otherwise differ ($\gamma = -.025$, *SE* = .053, $t(88) = -.47$, $p = .637$). Although rich players contributed more in absolute terms ($\gamma = 4.40$, *SE* = .59, $t(88) = 7.40$, $p < .001$), a rerun of this model with contributions as a proportion of endowment size revealed that rich players contributed less in relative terms ($\gamma = -.18$, *SE* = .026, $t(88) = 6.73$, $p < .001$). Accounting for wealth improved model fit ($\Delta\chi^2(2) = 102.69$, $p < .001$) and explained a substantial amount of the level-2 variance in the initial intercept (38.26%) and endgame intercept (23.28%), but explained little of the variance in the slope (0.25%) at level-2[33].

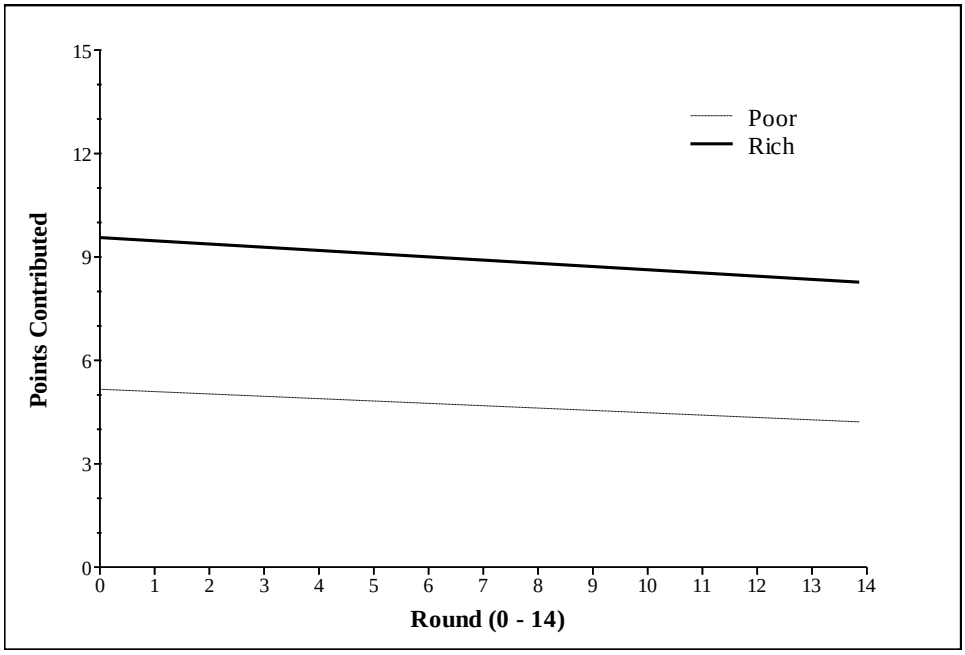

Figure 11: Effect of Wealth on Cooperation

Before further elaborating on our multilevel model, we conducted one-way ANOVAs testing for differences between the experimental conditions in pre-game interdependence, prosociality, and morality pooled across level-2 units[34]. As predicted, the experimental conditions differed significantly in their ratings of both pre-game

interdependence ($F(2, 43) = 5.74$, $p = .006$, $\eta_p^2 = .211$) and pre-game morality ($F(2, 43) = 4.93$, $p = .012$, $\eta_p^2 = .187$) but not pre-game prosociality ($F(2, 43) = 1.07$, $p = .352$). Post-hoc contrasts revealed that ratings of pre-game interdependence were significantly higher in the interdependence condition ($M = 4.72$) relative to the control condition ($b = .52$, $SE = .15$, $p = .002$) and marginally higher than the punishment condition ($b = .28$, $SE = .15$, $p = .074$). Similarly, post-hoc contrasts also revealed that ratings of pre-game morality were significantly higher in the interdependence condition ($M = 4.69$) relative to either the punishment condition ($b = .48$, $SE = .18$, $p = .009$) or the control condition ($b = .49$, $SE = .18$, $p = .009$).

We then added the experimental manipulations as level-3 predictors to test whether group dynamics affect cooperation (see Appendix C for this model's equation)[35]. The model revealed that participants' initial contributions were significantly higher in the interdependence condition relative to the control condition ($\gamma = 1.51$, $SE = .56$, $t(85) = 2.69$, $p = .009$), but were not significantly different in the punishment condition ($\gamma = .51$, $SE = .59$, $t(85) = .91$, $p = .367$). In contrast, the decline in contributions over time was significantly attenuated in the punishment condition relative to the control condition ($\gamma = .10$, $SE = .043$, $t(85) = 2.39$, $p = .019$), but was not significantly attenuated in the interdependence condition ($\gamma = -.017$, $SE = .043$, $t(85) = -.40$, $p = .692$). These experimental manipulations improved model fit ($\Delta\chi^2(4) = 15.92$, $p = .004$) and had a moderate to large effect size, explaining 12.70% of the variance in the initial intercept, 14.12% of the variance in the endgame intercept, and 19.75% of the variance in the slope at level-3.

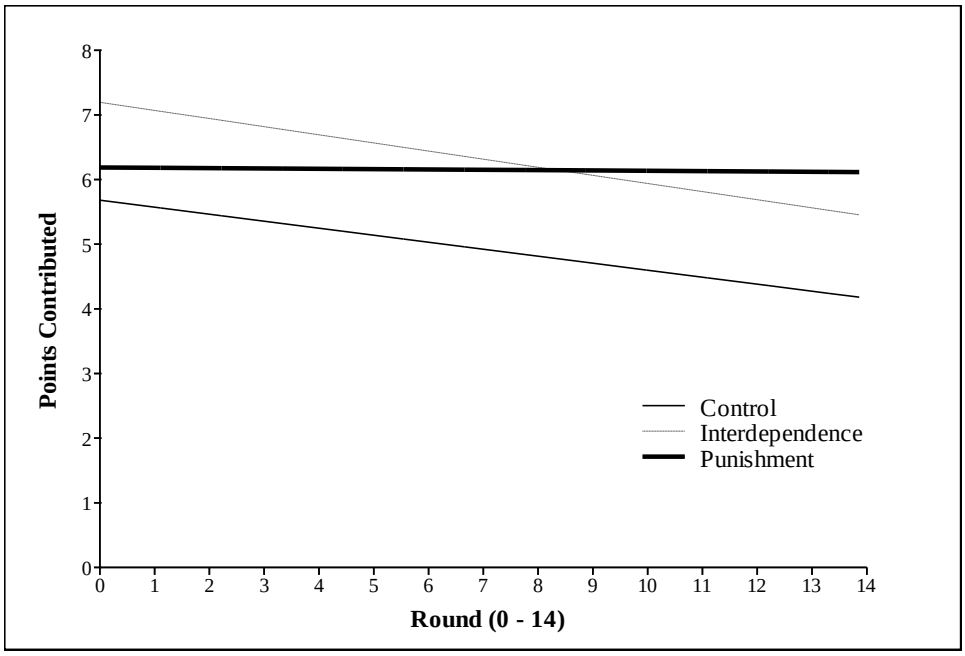

Figure 12: Effect of the Experimental Manipulations on Cooperation

Next, we calculated a measure of inequality—the Gini index—for each of the groups in our study (k = 89)[36]. The groups in this study appeared to be substantially more unequal (*Gini index*$_{S5}$ = .17, *SD* = .064) than in our prior studies (*Gini index*$_{S3}$ = .060, *SD* = .032; *Gini index*$_{S4}$ = .046, *SD* = .025). We conducted a one-way ANOVA testing for differences between the experimental conditions on this Gini index, but there were no significant differences between the control, punishment, or interdependence conditions in their level of inequality (F(2, 86) = .27, p = .763).

We also calculated a proportional measure of participants' contributions to the common pool that accounts for wealth[37]. These contributions ultimately determined both final payoffs and the Gini indices, so we can think of proportional contributions as a measure of relative influence over the outcome of the game. We tested a general linear model with these proportional contributions (X) predicting the inequality of each persons' group (Y) conditional on wealth (M). Our analyses revealed a significant wealth

by giving interaction ($F(1, 324) = 31.65$, $p < .001$, $\eta_p^2 = .089$). Poor players could make the group less unequal by contributing to the common pool ($b = -.067$, $SE = .012$, $t(324) = -5.44$, $p < .001$, $\eta_p^2 = .084$), but this relationship was significantly stronger for rich players ($b = -.15$, $SE = .026$, $t(324) = -5.63$, $p < .001$, $\eta_p^2 = .089$). In other words, rich players' behavior was the strongest predictor of each groups' level of inequality ($b = -.212$, $SE = .023$, $t(324) = -9.34$, $p < .001$, $\eta_p^2 = .212$).

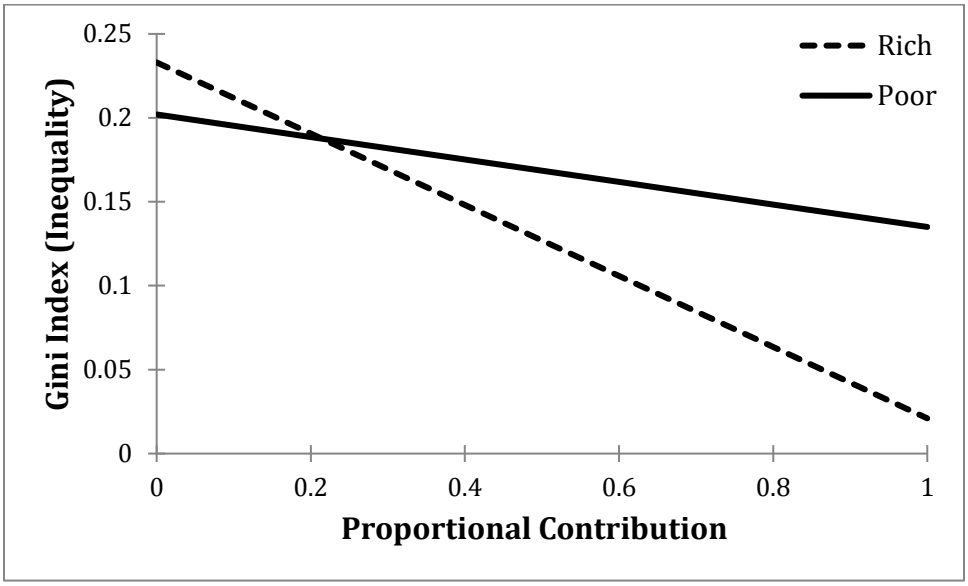

Figure 13: Effect of Wealth on Inequality

We wanted to identify the characteristics that predict whether rich players bring about inequality, so we tested a series of general linear models with wealth, an individual difference predictor, and their interaction predicting inequality[38]. Only pre-game interdependence was associated with inequality. More specifically, there was a significant wealth by pre-game interdependence interaction ($F(1, 169) = 6.02$, $p = .015$, $\eta_p^2 = .034$); greater pre-game interdependence predicted lower levels of inequality for rich players ($b = -.023$, $SE = .009$, $t(169) = -2.73$, $p = .007$, $\eta_p^2 = .042$) but not poor players ($b = .003$, $SE = .006$, $t(169) = .46$, $p = .645$).

Post-Game Ratings

Next, we investigated the effect of inequality, the experimental manipulations, and wealth on post-game ratings of interdependence, prosociality, and morality[39]. We ran a series of 2-level hierarchical linear models with all random effects using full maximum likelihood estimation (see Appendix C for these models' equations). The level-1 units were the individuals making the ratings and the level-2 units were the experimental groups. Across all three models, the only significant predictor was inequality (all other $p \geq .320$)[40]. Members of unequal groups felt less interdependence (γ = -2.92, *SE* = 1.11, $t(85) = -2.63$, $p = .010$), less prosociality (γ = -4.20, *SE* = 1.33, $t(85) = -3.15$, $p = .002$), and less morality (γ = -3.27, *SE* = 1.01, $t(85) = -3.24$, $p = .002$) in their groups. Inequality alone explained 17.65% of the variance in interdependence, 22.89% of the variance in prosociality, and 25.21% of the variance in morality at level-2.

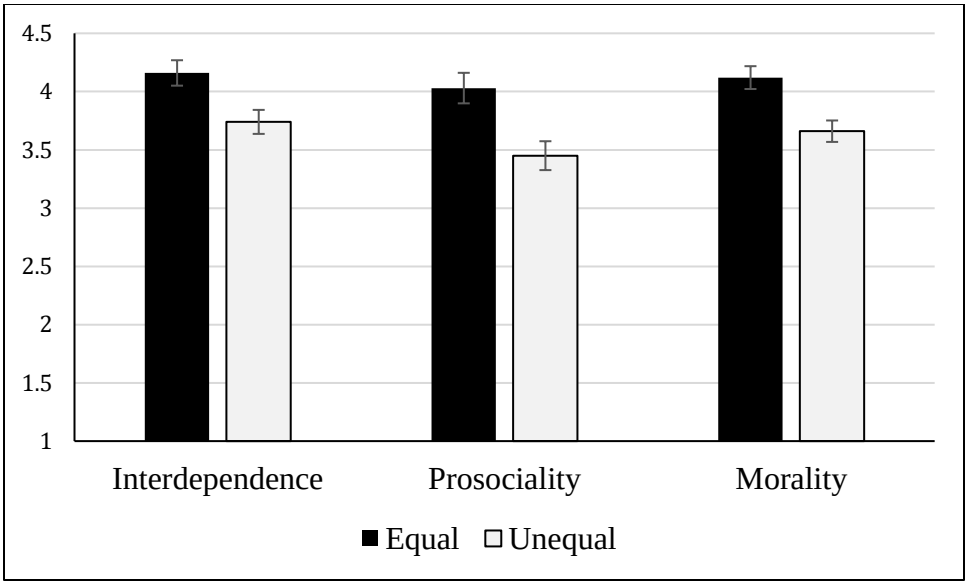

Figure 14: Effect of Inequality on Interdependence, Prosociality, and Morality

Finally, we tested the effect of inequality, the experimental manipulations, wealth, interdependence, and justice sensitivity[41] on support for equality and equity using a

similar 2-level model with all possible random effects (see Appendix C for these models' equations). The model revealed that rich players were less supportive of equality than poor players (γ = -.63, *SE* = .17, $t(75) = -3.66$, $p < .001$). In addition, justice sensitivity was positively associated with support for both equality (γ = .45, *SE* = .092, $t(75) = 4.92$, $p < .001$) and equity (γ = .24, *SE* = .10, $t(75) = 2.32$, $p = .023$). None of the remaining effects reached significance (all $p \geq .147$).

Discussion

Although the interdependence manipulation increased contributions to the common pool, it did not sustain cooperation as we found in Studies 3 and 4. The reason why, however, was clear from our analyses; rich players contributed less than poor players in proportion to endowment size. Rich players appeared to feel entitled to their larger endowment—even though it was literally determined by chance—because they were less supportive of equality as a distributional principle. The problem is that rich players exerted the most influence over the outcome of the game. They had the largest endowment and served as a focal point or correlating device for the other players' actions. If rich players would not cooperate, nor would the other players and the inescapable result was heightened inequality. Our most important finding was that inequality weakened participants' sense of community; in the face of modest inequality by real-world standards, they did not feel interdependent and actually thought the other players were less prosocial and less moral.

The punishment manipulation sustained cooperation, but the costs of punishment actually decreased payoffs and increased inequality (depending on how you calculate the Gini index). This might explain why increased rates of cooperation in this condition did

not translate into greater interdependence. We thought the interdependence manipulation would still work; after all, there is little incentive to horde one's wealth if you are unlikely to keep it. In fact, a game theoretic account would suggest that full cooperation is a strictly dominant strategy. However, many participants expressed a preference for an equal division heuristic when describing their strategy in this condition (see Allison & Messik, 1990). This might reflect a tendency for people to use tacit coordination rules in strong situations like this (see Kwaadsteniet, van Dijk, Wit, & de Cremer, 2006).

Interestingly, the only characteristic that appeared to induce rich players to bring about greater equality was pre-game interdependence. Individual differences in justice sensitivity and social justice were associated with support for equality as a distributional principle, but they did not have a significant effect on the behavior of rich players. This finding hints at a potentially important problem; people need to feel interdependent to do something about inequality, but inequality makes it difficult for them to feel this sense of interdependence.

Notes

[26] The Gini index is a measure of equality bounded between zero (perfect equality) and one (perfect inequality). The Gini index is the average absolute difference between all pairs of individuals divided by twice the mean, or equivalently it is twice the area between the Lorenz curve and the line of perfect equality (Allison, 1978).

[27] Most participants were unfamiliar with the public goods game (68.8%) but understood it (87.8%). This is similar to the rates we found in Study 3 and Study 4. Unfortunately,

due to a computer malfunction, 49 participants did not have data on demographics, justice sensitivity, familiarity, or understanding. For this reason, we did not include familiarity or understanding as covariates in subsequent analyses, and only preliminary analyses were conducted with justice sensitivity.

[28] We adjusted the rich player's endowment so that there would be 15 points per person in each version of the game. We tried to recruit only 4-person groups, but this proved to be extremely difficult—despite overscheduling sessions—because of no-shows.

[29] We avoided the terms "rich" and "poor" in the directions, instead using neutral language around larger and smaller endowments.

[30] This ratio is typical in the behavioral economic literature. However, experiments in this literature typically cap the losses from punishment at zero; we relaxed this assumption and allowed punishment to produce negative payoffs.

[31] Unfortunately, we can only perform preliminary analyses with these items because 47.26% of the data was lost (n = 155); a research assistant took original paper questionnaires out of the lab to conduct data entry and summarily lost the raw data.

[32] Covariates were centered so that the intercepts represented individuals in 3.5-person groups. None of the effects associated with group size were significant (all $p \geq .379$).

[33] To calculate these effect size estimates accurately, we excluded the random effects associated with wealth. Adding these random effects pushes almost all of the level-2 variance to level-3 and further improves model fit ($\Delta\chi^2(7) = 292.29$, $p < .001$).

[34] For this analysis, we used the 46 groups with complete pre-game data.

[35] We also added the experimental manipulations as level-3 predictors of wealth to test whether they have different effects for rich and poor players. The punishment manipulation had a significantly stronger effect on the contribution rate of rich players relative to poor players (γ = .25, *SE* = .13, $t(86) = 2.00$, $p = .049$). However, the contribution rate in the punishment condition was flat or rather non-significantly different from zero for both rich players (γ = .13, *SE* = .099, $t(85) = .129$, $p = .200$) and poor players (γ = .012, *SE* = .032, $t(85) = .37$, $p = .714$). Furthermore, including these predictors did not improve model fit ($\Delta\chi^2(4) = 5.02$, $p = .284$).

[36] We calculated the Gini index using participants' final income from the public goods game. We excluded points lost due to punishment, but find similar results if we include these lost points in the calculation. The primary difference is that the post-punishment Gini index reveals even higher levels of inequality in the punishment condition (*M* = .31) relative to the control condition (*b* = .13, *SE* = .026, $p < .001$) or the interdependence condition (*b* = .14, *SE* = .026, $p < .001$). Even though rich players are sanctioned more than poor players (γ = 1.32, *SE* = .42, $t(87) = 3.16$, $p = .002$), the destruction of wealth nonetheless yields rich players a larger proportion of all payoff points.

[37] We reran our baseline model with round of play (1 – 15) predicting contributions to the common pool relative to endowment size so we could estimate and export endgame intercept parameters for each person that represent endgame proportional contributions.

[38] These analyses are somewhat preliminary because of missing data present to varying degrees in all these individual difference variables: pre-game interdependence, pre-game prosociality, pre-game morality, social justice, and justice sensitivity.

[39] We did not include getting (or giving) as we did in Study 4 because these variables cause inequality, and thus are highly collinear. To illustrate this, we tested a general linear model with wealth, giving (group-mean centered), and their interaction predicting inequality. There was a massive interaction effect ($F(1, 324) = 471.24$, $p < .001$, $\eta_p^2 = .593$) such that getting predicted greater equality for poor players ($b = -.087$, $SE = .005$, $t(324) = -16.93$, $p < .001$, $\eta_p^2 = .469$) but greater inequality for rich players ($b = .057$, $SE = .004$, $t(324) = 13.59$, $p < .001$, $\eta_p^2 = .363$). Getting still predicts interdependence in this study ($r(326) = .289$, $p < .001$), but poor players get fewer points in unequal groups.

[40] However, there was also a marginal effect of wealth such that poor players felt the other participants were less prosocial ($\gamma = -.29$, $SE = .16$, $t(88) = -1.84$, $p = .069$) and moral ($\gamma = -.27$, $SE = .15$, $t(88) = -1.72$, $p = .089$) than rich players did.

[41] We used justice sensitivity because it had less missing data, but we find the same pattern of effects if we instead use social justice as a predictor.

CHAPTER 7

GENERAL DISCUSSION

Answering the central question of the present research, we found consistent evidence for a link between interdependence and social justice across different samples and methodologies. Study 1 revealed that social justice was more important in the groups of everyday life that are interdependent. Study 2 demonstrated that Americans who feel interdependent endorse social justice more, and this relationship was even stronger when we primed their sense of solidarity. Studies 3 and 4 manipulated well-mixed groups' level of interdependence, which sustained our operationalization of social justice behavior—cooperation in a public goods game. Just as Durkheim (1893/1984) predicted more than a century ago, there is a special connection between interdependence—what he more formally labeled as organic solidarity—and social justice.

Our experimental manipulation in the economic games simulated interdependence by giving participants' shared outcomes, but a more nuanced psychological construction of interdependence was evident in Study 2. These items emphasized four characteristics of interdependence based on past work on entitativity (e.g., Brewer, Hong, & Li, 2004), solidarity (e.g., Durkheim, 1893/1984), and interdependence theory (e.g., Kelley & Thibaut, 1978): possessing common goals, having shared interests, needing to depend on one another, and being affected by what happens to each other. Much of our work has been implicitly couched in social identity theory (e.g., Tajfel et al., 1971), but interdependence theory gives us a different perspective. "Interdependence theory analyzes the relations between people in terms of situation structure… [which] is the interpersonal reality within which cognition, affect, and motivation transpire" (Rusbult &

Van Lange, 2008, p. 2050). What this emphasis on situation structure makes clear is that social justice is elicited by particular types of social situations because it solves problems inherent to these specific interactions (see Carnes et al., 2015; for a taxonomy of situation structure, see Kelly et al., 2003). That is, social justice makes possible certain patterns of human social organization that might not otherwise be tractable. An important question for future research is to identify the situation characteristics that drive different moral principles.

The consequences of social justice were also highlighted in several of these studies. In Study 2, we found that social justice predicted support for economic, racial, and gender equality; in fact, social justice predicted inclusiveness—operationalized as identification with all of humanity and an expansive circle of moral regard—despite its seemingly parochial roots in interdependence. However, our potentially most interesting finding was in Study 4; our analyses suggested that social justice behavior fed back into group members' sense of interdependence with others. This suggests there might be a bidirectional relationship or virtuous cycle between these psychological processes.

This virtuous cycle resolves a lingering question about moral binding. The original conceptual framework for binding morality proposed by Haidt and colleagues suggests that some moral principles encourage prosocial behavior indirectly by motivating people to care about the roles, duties, and obligations of their group (e.g., Haidt & Graham, 2009; Haidt & Joseph, 2004; Haidt & Kesebir, 2010). Carnes and Lickel (under review) formalized this idea with a mediational framework in which morality (*X*) shapes prosocial behavior (*Y*) through identity (*M*). In the present research, however, identity (*X*) generally shapes prosocial behavior (*Y*) through morality (*M*).

These different pathways appear to be incompatible at first glance, but a virtuous cycle or feedback loop means that both pathways can simultaneously be correct.

Given this virtuous cycle, Study 5 left us with a puzzle. Interdependence appears to play an important role in fostering the moral convictions that oppose inequality, but inequality itself undermines perceptions of interdependence. Our interdependence manipulation did not sustain cooperation in the same way when we exogenously imposed inequality on experimental groups, and endogenous inequality was associated with less interdependence, prosociality, and even morality. The breaking of this virtuous cycle is not certain; American participants in Study 2—who live in an incredibly unequal society—still exhibited an association between interdependence and social justice. However, our findings suggest that this cycle may at least weaken as inequality worsens.

One way of understanding how inequality undermines support for equality comes from the impartiality account of fairness. This account suggests that people prefer equitable outcomes because humans value impartial actors (DeScioli & Kurzban, 2009) and distributing resources equitably signals one's impartiality to others (for a review, see Shaw, 2013). Both adults (Choshen-Hillel, Shaw, & Caruso, 2015) and children (Shaw, Choshen-Hillel, & Caruso, 2016; Shaw et al., 2014; Shaw & Olson, 2012, 2014) exhibit a self-disadvantaging effect that supports the impartiality account. That is, people would rather destroy a resource than create self-advantaging inequity because they do not want to appear partial. Perhaps some people oppose equality when they are nested in an unequal system because an equal distribution of resources would be self-advantaging, and thus would make them appear to be partial. This might explain why interdependence is so important for motivating broad support for equality; actions that benefit the self also

benefit other group members—and vice versa—in an interdependent group, so concerns about appearing partial or impartial are alleviated.

Putting the present research back into the historical context of progressive taxation and policy, our work suggests that there might be nothing particularly special about mass mobilization for warfare in inducing support for equality (see Scheve & Stasavage, 2016). Mass mobilization might have yielded support for social justice because people felt a renewed sense of interdependence. If some other mass event—like climate change or globalization—makes people feel interdependent, then those events should yield support for social justice as well. Based on the findings of Study 5, whether these events actually make people feel interdependent may critically depend on the inequality of our economic system.

Many media outlets have noted how the rise of right-wing populism in America and elsewhere was fueled by the downtrodden—those left behind by outsourcing or automation—so perhaps this work sheds a new light on a contemporary problem. Downtrodden communities should not feel interdependent because they have been left behind by radical inequality. Considering the link between interdependence and inclusion, it should come as no surprise that they oppose the "others." People of a different race, people of a different religion, people of a different political party, homosexual people, immigrants, and refugees—these groups were all represented in our measure of inclusion in Study 2, and these are the "others" in public debate. Furthermore, if this political coalition does not feel interdependent, then it is also predictable that they would oppose engines of interdependence; climate change, globalization, and government all make us more interdependent—they have collective benefits and costs—

but people with a depleted sense of interdependence might feel dissonance between how they feel and what they hear from scientists or policymakers. So they ignore what they hear, instead trusting what they feel, and attack the partiality of disadvantaged groups advocating for equality.

We learned about more than social justice in the present research; our findings also implicate social order. Study 1 revealed that social order was more important in the groups of everyday life that have strong norms and low tolerance for deviant behavior (i.e., tightness-looseness). Study 2 demonstrated that Americans who feel a visceral sense of oneness with the group (i.e., identity fusion) endorse social order more, and this relationship was even stronger when we primed their sense of solidarity. Finally, Study 3 manipulated the norms in well-mixed groups, which initiated our operationalization of social order behavior—coordination in a matching game. These findings, particularly those of Study 2, are consistent with past research on moral binding (see Carnes & Lickel, under review). Taken together with our social justice findings, this work suggests that there are at least two distinct types of moral binding. These two types of binding correspond remarkably well with Durkheim's (1893/1984) conception of organic solidarity and mechanical solidarity.

Turning to politics, Graham, Haidt, and Nosek (2009) argue that political conservatives rely on a more diverse set of moral principles than liberals (for a review, see Haidt, 2012); liberals endorse only the individualizing moral principles of care/harm and fairness/cheating, while conservatives also endorse the binding moral principles of loyalty/betrayal, authority/subversion, and purity/degradation. Janoff-Bulman and Carnes (2013) assert that this imbalance results from an incomplete map of the moral domain

rather than something intrinsically less groupish about the moral psychology of liberals. That is, social justice is the group-based moral principle representing the concerns of liberals not represented by Moral Foundations Theory. Perhaps unsurprisingly, Graham (2013) disagrees with this assessment. The present research provides compelling evidence that social justice is indeed group-based and has a binding effect of its own.

The present research was illuminating, but it also had its limitations. First, as might be expected on a project spanning three years, there were some inconsistencies in measurement between studies. This could, however, be construed as a strength considering that we found consistent results regardless of measurement differences. Second, we also had some data problems, particularly in Study 5, due to research assistants departing from lab protocol. This is regrettable, but one potential remediation would be to use multilevel multiple imputation to replace the missing data (see Enders, Mistler, & Keller, 2016). A third problem we encountered in Studies 4 and 5 involved how to predict group-level dependent variables with individual-level independent variables. This is because hierarchical linear modeling only uses group-level data as predictors. Our solution was to treat group-level data as individual-level data, but this is not ideal. In the future, we could solve this problem by reanalyzing the data using multilevel structural equation modeling with latent growth curves. A fourth limitation was that all of our data was collected in a single cultural context (i.e., American participants). Although such samples are the norm in psychological research, these WEIRD cultures (i.e., Western, educated, industrialized, rich, and democratic) are in fact only a narrow slice of the cultural variation seen in human societies (Henrich, Heine, &

Norenzayan, 2010). Collected data from non-WEIRD cultural contexts could help resolve this final problem.

There are many potential future directions for this work. Some questions are more applied, like understanding the effects of climate change, globalization, or inequality on public opinion. Others are more oriented around understanding psychological process, such as how inequality undermines interdependence or how the virtuous cycle works. There are also clear links to other theoretical accounts (e.g., interdependence theory, the impartiality account of fairness, moral foundations theory) that could benefit from conceptual integration. President Barack Obama called inequality the defining issue of our time, so our hope is that whatever course this work takes will help ameliorate a generation-defining problem.

APPENDICES

APPENDIX A

MATERIALS

Study 1

1. **Interdependence 1:** In some groups, members can depend on each other no matter what, whereas members of other groups don't depend on each other at all. How much can members of the following groups depend on one another?
2. **Interdependence 2:** Members of some groups are strongly affected when something, good or bad, happens to other people in the group, while members of other groups are weakly affected. How strongly affected are members of the following groups by what happens to other people in the group?
3. **Interdependence 3:** Most groups have common goals that their members want to attain. Members of some groups need each other in order to attain these goals, whereas members of other groups don't really need one another at all. How much do members of the following groups need each other?
4. **Tightness 1:** Most groups have norms or informal rules to guide behavior. Some groups have very strong norms that members must comply with, while other groups have very weak norms that members can choose whether or not to follow. How strong are the norms in the following groups?
5. **Tightness 2:** In some groups, members are strictly punished for failing to conform to the rest of the group, whereas members of other groups are never punished for acting differently. How strictly are members of the following groups punished if they don't conform and act differently?

6. **Tightness 3:** Most groups have norms or informal rules to guide behavior. Some groups have very many norms for members to follow, whereas other groups have very few norms for members to follow. How many norms are there for members of the following groups to adhere to?
7. **Social Justice 1:** In some groups, members have a responsibility to help the neediest members of their group, while members of other groups have no such responsibility. How important is it for members of the following groups to help their neediest members?
8. **Social Justice 2:** Members of some groups are expected to share what they can with other members, whereas members of other groups don't need to share with each other at all. How important is it for members of the following groups to share with one another?
9. **Social Justice 3:** In some groups, members care very much about the well-being of others simply because of their shared group membership, whereas in other groups members may not care much at all unless they have a personal relationship with someone. How much do members of the following groups care about the well-being of other group members?
10. **Social Order 1:** Most groups have some kind of hierarchy or social ladder. In some groups, the social ladder is very steep with big differences between those at the top and bottom, while the social ladder is very flat in other groups with almost no difference between those at the top and bottom. How steep is the social ladder in the following groups?

11. **Social Order 2:** Most groups have traditions and customs of some kind. Some groups have more elaborate or complex traditions, whereas other groups have plainer or simpler customs. How complex are the traditions and customs in the following groups?
12. **Social Order 3:** Most groups have both leaders and followers. In some groups, leaders will frequently tell followers what to do and what not to do, while leaders in other groups will rarely offer such guidance. How often do leaders in the following groups tell followers what to do?

Study 2

1. **Interdependence 1:** There is a common thread that connects the hopes and dreams of Americans.
2. **Interdependence 2:** There are common goals that unite us as Americans.
3. **Interdependence 3:** Despite our differences, we have fundamentally shared interests as Americans.
4. **Interdependence 4:** I feel interdependent with my fellow Americans no matter our divisions.
5. **Interdependence 5:** We, the American people, are all in this together.
6. **Interdependence 6:** We have to depend on each other, no matter what, because we are all Americans.
7. **Interdependence 7:** It affects me when something, good or bad, happens to other Americans.
8. **Interdependence 8:** My well-being is tied to the well-being of every other American.
9. **Identity Fusion 1:** My country is me.

10. **Identity Fusion 2:** I feel immersed in my country.
11. **Identity Fusion 3:** I make my country strong.
12. **Identity Fusion 4:** I am nothing without my country.
13. **Identity Fusion 5:** I am strong because of my country.
14. **Identity Fusion 6:** I am one with my country.
15. **Identity Fusion 7:** I'll do for my country more than any other group member would do.
16. **Identity Fusion 8:** I have a deep emotional bond with my country.
17. **Social Justice 1:** It's our responsibility, not just a matter of personal preference, to provide for those worst off in American society.
18. **Social Justice 2:** Future generations should judge our society by how we treat those Americans most desperate for help.
19. **Social Justice 3:** It is important for those who are better off to share with the most vulnerable members of society.
20. **Social Justice 4:** Americans should feel like it is their duty to care for each other no matter what.
21. **Social Justice 5:** If we want to be a just society, Americans at the top must feel an obligation to improve the well-being of those at the bottom.
22. **Social Justice 6:** As Americans, we have a responsibility to look out for one another.
23. **Social Order 1:** Americans should show respect for a group member who has authority.
24. **Social Order 2:** Americans should be loyal no matter what the cost.

25. **Social Order 3:** If you live here, you should try to fit in with our values and practices.
26. **Social Order 4:** Americans should honor our group's heritage and history.
27. **Social Order 5:** There are good reasons why our traditions have lasted for so long.
28. **Social Order 6:** American is made stronger when its rules and leaders are faithfully followed.
29. **Equality 1:** Increasing equality is ultimately beneficial to everyone in society.
30. **Equality 2:** The gap between the "haves" and "have nots" is too big.
31. **Equality 3:** American should be a fairer place for the 99%.
32. **Equality 4:** Inequality has a harmful effect on our country.

Study 3

1. **Norm Manipulation:** I want you to imagine that you are all in a group with norms or informal rules about how players should behave in these economic games. The norm or informal rule in your group is that players should (pick the number 4/choose to contribute all of their endowment) while playing the (Matching game/Common Pool game). There is no punishment for violating this rule, but it tells you what most people should do in the game.
2. **Interdependence Manipulation:** I want you to imagine a scenario in which you have no idea how well you are going to do in this game, or how many points you are going to have, relative to the other players; how would you want points to be distributed among the (3/4/5) players in the group, so that you have relatively similar outcomes or dissimilar outcomes? You should keep this question in mind, because there is an unusual twist at the end of the game; I'm going to randomly shuffle your payoffs at

the end of the experiment so that you are just as likely to get someone else's payoffs as your own. In other words, there is only a (33%/25%/20%) chance that you will keep your own payoff rather than taking another player's payoff.

3. **Social Justice 1:** It is our responsibility, not just a matter of personal preference, to provide for groups worse off in society.
4. **Social Justice 2:** It is important in a just society for those who are better of to share what resources they can with the most vulnerable members.
5. **Social Justice 3:** In the healthiest societies, those at the top should feel responsible for improving the well-being of those at the bottom.
6. **Social Justice 4:** Looking out for groups worst off in society will not make those groups too dependent on help.
7. **Social Justice 5:** Increased economic equality is ultimately beneficial to everyone in society.
8. **Social Order 1:** In order for society to be stable, there has to be agreement about what behaviors and ways of life are appropriate.
9. **Social Order 2:** It is harmful to society when people choose radically different beliefs and assumptions about the world.
10. **Social Order 3:** There are often good reasons why traditions have lasted for so long, even if people do not fully understand those reasons.
11. **Social Order 4:** In the most decent societies, people should try to fit in and respect the values and practices of the larger community.
12. **Social Order 5:** Society is made stronger when its rules and leaders are faithfully followed.

Study 4

1. **Interdependence Manipulation:** Normally, players just keep the points they earned in the 15 rounds of the Common Pool Game. However, your group will have a different payoff system. At the end of the experiment, we're going to randomly assign each player's payoff to one of the (3/4/5) players in the game. Here's an explanation of how this payoff system works. Imagine that you earned 375 points in the Common Pool Game and then wrote that number on a card. Suppose the other players also wrote how many points they earned on cards. If I shuffled those cards and then randomly handed each player a card, you might get the card with your 375 points, but it is more likely that you will get a card with another player's points written on it. As you can see, there is only a (33%/25%20%) chance that you will keep your original payoff, and a (66%/75%/80%) chance that you will get someone else's payoff! This different payoff system means that how well you actually do in the game depends much more on how many points the other players have, and much less on how many points you personally have.
2. **Prosocial Manipulation:** Normally, players just keep the points they earned in the 15 rounds of the Common Pool Game. However, you personally will have a different payoff system. At the end of the experiment, we're going to increase the size of your payoff based on how much you usually contributed to the common pool. Here's an explanation of how this payoff system works. If you contributed an average of 5 points to the common pool, then we'll increase the size of your payoff by 5%. If you contributed an average of 10 points to the common pool, then we'll increase the size of your payoffs by 10%. And if you contributed an average of 15 points to the

common pool, then we'll increase the size of your payoffs by 15%. As you can see, the more you share with the other players, the more we'll increase the size of your payoff! This different payoff system means that you actually benefit much more than normal from contributing your points to the common pool.

3. **Public Goods Game Directions:** The Common Pool Game is an example of a generic situation in which a group of people want to achieve a common goal, but everyone in the group has the option of free-riding on the others' efforts. You are going to play 15 rounds of this game for points; there will not be any practice rounds. At the beginning of each round, you will receive an endowment of 20 points. You can think of the endowment as an allowance you receive every round. Your task in each round is to decide how much of your endowment, if any, you would like to contribute to the common pool, and how much of your remaining endowment, if any, you would like to keep for yourself. The other players will not know how much you personally contributed to the common pool, but they will know how much was contributed to the common pool in general. At the end of each round, we will increase the number of points in the common pool by 60%. In contrast, the points you kept for yourself will not be increased at all. Finally, the common pool will be divided evenly between all of the players in the game. In other words, each player in the game gets an equal share of the common pool no matter how much or how little they contributed. This means that the group as a whole can earn the most points when everyone contributes to the common pool, but individuals can earn the most points when they free-ride on others' contributions. Your payoff from each round will be equal to the number of

points you kept for yourself plus your share of the points in the common pool. Your payoff is what will determine your likelihood of winning a gift card in the raffle.

4. **Equality:** The points should have been divided pretty equally among the players.
5. **Justice:** The players who got the fewest points should have gotten more points than they did.
6. **Equity 1:** The players who contributed the least to the common pool should have gotten fewer points.
7. **Equity 2:** The players who contributed the most to the common pool should have gotten more points.

Study 5

1. **Punishment Manipulation:** Normally, players just keep the points they earned in each round of the Common Pool Game. However, the players in your group will be able to sanction each other at the end of each round. At the end of each round, you may assign sanction points to the other players if you want to (but you do not have to). Here's an explanation of how this sanctioning system works. It costs you one point from your payoff to assign a sanction point to another player, but that player will then lose three points from their payoff. You may assign as many sanction points as you have points in your initial endowment. For example, if you have a ten point endowment, you may assign up to ten sanction points per round. It is possible to receive negative points because of this sanctioning. Players will find out who sanctioned them (although you will still only be identified by a player number). You can assign sanction points for any reason you want, just remember that other players can sanction you back (and you can sanction them back).

2. **Equality 3:** The players who got the most points should have gotten fewer points than they did.
3. **Equity 3:** It would be fairer if inputs to the common pool were proportional to outputs.
4. **Victim Sensitivity 1:** It makes me angry when others are undeservingly better off than me.
5. **Victim Sensitivity 2:** It worries me when I have to work hard for things that come easily to others.
6. **Beneficiary Sensitivity 1:** I feel guilty when I am better off than others for no reason.
7. **Beneficiary Sensitivity 2:** It bothers me when things come easily to me that others have to work hard for.
8. **Observer Sensitivity 1:** I am upset when someone is undeservingly worse off than others.
9. **Observer Sensitivity 2:** It worries me when someone has to work hard for things that come easily to others.
10. **Perpetrator Sensitivity 1:** I feel guilty when I enrich myself at the cost of others.
11. **Perpetrator Sensitivity 2:** It bothers me when I use tricks to achieve something while others have to struggle for it.

APPENDIX B

TABLES

Table 1: Model Comparisons (Excluding Non-Significant Alternative Models)

Model	χ^2	*df*	$\Delta\chi^2$	Δdf	*p-value*
Fully Constrained Baseline Model	879.26	333	-	-	-
Added Paths from Group Dynamics to Morality	840.19	331	39.07	2	< .001
Partially Variant Measurement	778.50	314	61.69	17	< .001
Variant Paths from Group Dynamics to Morality	748.37	306	30.13	8	< .001

Note. Each model is compared to the prior model in the table.

Table 2: Factor Parameter Estimates for Final Model

Parameter	Strong Solidarity Condition				
	b	*SE*	*t*	*p*	*β*
Interdependence → Justice	.876	.098	8.91	< .001	.674
Fusion → Order	.795	.058	13.80	< .001	.771
Interdependence ↔ Fusion	.627	.078	8.02	< .001	.646
Justice ↔ Order	-.286	.055	-5.18	< .001	-.220
Justice Disturbance	.777	.082	9.51	< .001	.723
Order Disturbance	.653	.077	8.50	< .001	.416
	Weak Solidarity Condition				
	b	*SE*	*t*	*p*	*β*
Interdependence → Justice	.428	.074	5.77	< .001	.425
Fusion → Order	.686	.055	12.40	< .001	.732
Interdependence ↔ Fusion	.780	.099	7.92	< .001	.599
Justice ↔ Order	-.277	.058	-4.80	< .001	-.225
Justice Disturbance	.872	.088	9.92	< .001	.872
Order Disturbance	.719	.080	8.94	< .001	.476
	Both Conditions				
	b	*SE*	*t*	*p*	*β*
Fusion → Justice	-.276	.045	-6.16	< .001	-.323
Interdependence → Order	-.017	.062	-.28	.779	-.011
Justice → Equality	.682	.039	17.30	< .001	.698
Justice → Inclusion	.472	.043	10.87	< .001	.609
Order → Equality	-.291	.031	-9.45	< .001	-.359
Order → Inclusion	-.120	.030	-4.05	< .001	-.188
Equality ↔ Inclusion	.024	.029	.83	.407	.029
Equality Disturbance	.323	.057	5.71	< .001	.303
Inclusion Disturbance	.363	.054	6.77	< .001	.558

Note. The → indicates a path; the ↔ indicates an exogenous or disturbance covariance.

Table 3: Indicator Parameter Estimates, Reliability, and Validity for Final Model

Parameter	Strong Solidarity Condition			
	Factor Loading	*Error Variance*	*Reliability*	*Validity*
Interdependence 1	1	.469	.575	.765
Interdependence 2	1.28	.585	.642	.805
Interdependence 3	1.27	.457	.693	.836
Interdependence 4	1.13	.638	.558	.756
Social Justice 1	1	.215	.838	.901
Social Justice 2	1.00	.157	.875	.923
Social Justice 3	1.02	.442	.719	.843
Equality Self-Report	1	.496	.685	.808
Gini Index	-.103	.048	.193	.428
Disparity Rating Task	.212	.165	.223	.466
	Weak Solidarity Condition			
	Factor Loading	*Error Variance*	*Reliability*	*Validity*
Interdependence 1	1	.504	.662	.813
Interdependence 2	1.02	.578	.641	.799
Interdependence 3	1.14	.541	.704	.838
Interdependence 4	.770	.929	.386	.622
Social Justice 1	1	.222	.815	.914
Social Justice 2	1.06	.389	.738	.869
Social Justice 3	1.12	.331	.789	.894
Equality Self-Report	1	.456	.692	.848
Gini Index	-.152	.034	.413	.652
Disparity Rating Task	.212	.162	.224	.479
	Both Conditions			
	Factor Loading	*Error Variance*	*Reliability*	*Validity*
Identity Fusion 1	1	.417	.793	.891
Identity Fusion 2	1.05	.856	.674	.820
Identity Fusion 3	1.02	.488	.773	.881
Identity Fusion 4	1.03	.555	.753	.869
Social Order 1	1	.411	.789	.888
Social Order 2	1.03	.623	.724	.850
Social Order 3	1.09	.706	.723	.847
Moral Expansiveness	1	.762	.460	.678
Identify With All Humanity	.801	.331	.557	.746

Note. All factor loadings and error variances are unstandardized. Reliability is true variance divided by observed variance. Validity is the indicator-factor correlation.

APPENDIX C

EQUATIONS

Study 3

Public Goods Game

$$Contribution_{tij} = \pi_{0ij} + \pi_{1ij}*(Linear\ Round_{tij}) + e_{tij}$$

$$\pi_{0ij} = \beta_{00j} + \beta_{01j}*(Social\ Justice_{ij}) + \beta_{02j}*(Social\ Order_{ij}) + \beta_{03j}*(Familiar_{ij}) + \beta_{04j}*(Understand_{ij}) + r_{0ij}$$

$$\pi_{1ij} = \beta_{10j} + \beta_{11j}*(Social\ Justice_{ij}) + \beta_{12j}*(Social\ Order_{ij}) + \beta_{13j}*(Familiar_{ij}) + \beta_{14j}*(Understand_{ij}) + r_{1ij}$$

$$\pi_{2ij} = \beta_{20j} + \beta_{21j}*(Social\ Justice_{ij}) + \beta_{22j}*(Social\ Order_{ij}) + \beta_{23j}*(Familiar_{ij}) + \beta_{24j}*(Understand_{ij}) + r_{2ij}$$

$$\beta_{00j} = \gamma_{000} + \gamma_{001}(Norm_j) + \gamma_{002}(Interdependence_j) + \gamma_{003}(Size_j) + u_{00j}$$

$$\beta_{01j} = \gamma_{010}$$

$$\beta_{02j} = \gamma_{020}$$

$$\beta_{03j} = \gamma_{030}$$

$$\beta_{04j} = \gamma_{040}$$

$$\beta_{10j} = \gamma_{100} + \gamma_{101}(Norm_j) + \gamma_{102}(Interdependence_j) + \gamma_{103}(Size_j) + u_{10j}$$

$$\beta_{11j} = \gamma_{110}$$

$$\beta_{12j} = \gamma_{120}$$

$$\beta_{13j} = \gamma_{130}$$

$$\beta_{14j} = \gamma_{140}$$

Matching Game

$\text{Prob}(Match_{ijk}=1|\pi_{jk}) = \phi_{ijk}$

$\log[\phi_{ijk}/(1 - \phi_{ijk})] = \eta_{ijk}$

$\eta_{ijk} = \pi_{0jk} + \pi_{1jk}*(Round_{ijk})$

$\pi_{0jk} = \beta_{00k} + \beta_{01k}*(Social\ Justice_{jk}) + \beta_{02k}*(Social\ Order_{jk}) + \beta_{03k}*(\text{Understand}_{jk})$

$+ \beta_{04k}*(Familiar_{jk}) + r_{0jk}$

$\pi_{1jk} = \beta_{10k} + \beta_{11k}*(Social\ Justice_{jk}) + \beta_{12k}*(\text{Social Order}_{jk}) + \beta_{13k}*(\text{Understand}_{jk})$

$+ \beta_{14k}*(Familiar_{jk}) + r_{1jk}$

$\beta_{00k} = \gamma_{000} + \gamma_{001}(Interdependence_k) + \gamma_{002}(Norm_k) + \gamma_{003}(Size_k) + u_{00k}$

$\beta_{01k} = \gamma_{010}$

$\beta_{02k} = \gamma_{020}$

$\beta_{03k} = \gamma_{030}$

$\beta_{04k} = \gamma_{040}$

$\beta_{10k} = \gamma_{100} + \gamma_{101}(Interdependence_k) + \gamma_{102}(Norm_k) + \gamma_{103}(Size_k) + u_{10k}$

$\beta_{11k} = \gamma_{110}$

$\beta_{12k} = \gamma_{120}$

$\beta_{13k} = \gamma_{130}$

$\beta_{14k} = \gamma_{140}$

Study 4

Public Goods Game

$Contribution_{tij} = \pi_{0ij} + \pi_{1ij}*(Linear\ Round_{tij}) + e_{tij}$

$\pi_{0ij} = \beta_{00j} + \beta_{01j}*(Social\ Justice_{ij}) + \beta_{02j}*(Social\ Order_{ij}) + \beta_{03j}*(Familiar_{ij})$

$+ \beta_{04j}*(Understand_{ij}) + r_{0ij}$

$\pi_{1ij} = \beta_{10j} + \beta_{11j}*(Social\ Justice_{ij}) + \beta_{12j}*(Social\ Order_{ij}) + \beta_{13j}*(Familiar_{ij})$

$+ \beta_{14j}*(Understand_{ij}) + r_{1ij}$

$\beta_{00j} = \gamma_{000} + \gamma_{001}(Prosocial_j) + \gamma_{002}(Interdependence_j) + \gamma_{003}(Size_j) + u_{00j}$

$\beta_{01j} = \gamma_{010}$

$\beta_{02j} = \gamma_{020}$

$\beta_{03j} = \gamma_{030}$

$\beta_{04j} = \gamma_{040}$

$\beta_{10j} = \gamma_{100} + \gamma_{101}(Prosocial_j) + \gamma_{102}(Interdependence_j) + \gamma_{103}(Size_j) + u_{10j}$

$\beta_{11j} = \gamma_{110}$

$\beta_{12j} = \gamma_{120}$

$\beta_{13j} = \gamma_{130}$

$\beta_{14j} = \gamma_{140}$

Post-Game Ratings

$Equality_{ij} = \beta_{0j} + \beta_{1j}*(Social\ Justice_{ij}) + r_{ij}$

$\beta_{0j} = \gamma_{00} + \gamma_{01}*(Interdependence_j) + u_{0j}$

$\beta_{1j} = \gamma_{10} + u_{1j}$

$Justice_{ij} = \beta_{0j} + \beta_{1j}*(Social\ Justice_{ij}) + r_{ij}$

$\beta_{0j} = \gamma_{00} + \gamma_{01}*(Interdependence_j) + u_{0j}$

$\beta_{1j} = \gamma_{10} + u_{1j}$

$Equity_{ij} = \beta_{0j} + \beta_{1j}*(Social\ Justice_{ij}) + r_{ij}$

$\beta_{0j} = \gamma_{00} + \gamma_{01}*(Interdependence_j) + u_{0j}$

$\beta_{1j} = \gamma_{10} + u_{1j}$

Study 5

Public Goods Game

$$Contribution_{tij} = \pi_{0ij} + \pi_{1ij}*(Linear\ Round_{tij}) + e_{tij}$$

$$\pi_{0ij} = \beta_{00j} + \beta_{01j}*(Wealth_{ij}) + r_{0ij}$$

$$\pi_{1ij} = \beta_{10j} + \beta_{11j}*(Wealth_{ij}) + r_{1ij}$$

$$\beta_{00j} = \gamma_{000} + \gamma_{001}(Size_j) + \gamma_{002}(Interdependence_j) + \gamma_{003}(Punishment_j) + u_{00j}$$

$$\beta_{01j} = \gamma_{010} + u_{01j}$$

$$\beta_{10j} = \gamma_{100} + \gamma_{101}(Size_j) + \gamma_{102}(Interdependence_j) + \gamma_{103}(Punishment_j) + u_{10j}$$

$$\beta_{11j} = \gamma_{110} + u_{11j}$$

Post-Game Ratings

$$Interdependent_{ti} = \pi_{0i} + \pi_{1i}*(Wealth_{ti}) + e_{ti}$$

$$\pi_{0i} = \beta_{00} + \beta_{01}*(Interdependence_i) + \beta_{02}*(Punishment_i) + \beta_{03}*(Inequality_i) + r_{0i}$$

$$\pi_{1i} = \beta_{10} + r_{1i}$$

$$Prosociality_{ti} = \pi_{0i} + \pi_{1i}*(Wealth_{ti}) + e_{ti}$$

$$\pi_{0i} = \beta_{00} + \beta_{01}*(Interdependence_i) + \beta_{02}*(Punishment_i) + \beta_{03}*(Inequality_i) + r_{0i}$$

$$\pi_{1i} = \beta_{10} + r_{1i}$$

$$Morality_{ti} = \pi_{0i} + \pi_{1i}*(Wealth_{ti}) + e_{ti}$$

$$\pi_{0i} = \beta_{00} + \beta_{01}*(Interdependence_i) + \beta_{02}*(Punishment_i) + \beta_{03}*(Inequality_i) + r_{0i}$$

$$\pi_{1i} = \beta_{10} + r_{1i}$$

$$Equality_{ti} = \pi_{0i} + \pi_{1i}*(Wealth_{ti}) + \pi_{2i}*(Interdependent_{ti}) + \pi_{3i}*(Justice\ Sensitivity_{ti}) + e_{ti}$$

$$\pi_{0i} = \beta_{00} + \beta_{01}*(Interdependence_i) + \beta_{02}*(Punishment_i) + \beta_{03}*(Inequality_i) + r_{0i}$$

$$\pi_{1i} = \beta_{10} + r_{1i}$$

$\pi_{2i} = \beta_{20} + r_{2i}$

$\pi_{3i} = \beta_{30} + r_{3i}$

$Equality_{ti} = \pi_{0i} + \pi_{1i}*(Wealth_{ti}) + \pi_{2i}*(Interdependent_{ti}) + \pi_{3i}*(Justice\ Sensitivity_{ti}) + e_{ti}$

$\pi_{0i} = \beta_{00} + \beta_{01}*(Interdependence_i) + \beta_{02}*(Punishment_i) + \beta_{03}*(Inequality_i) + r_{0i}$

$\pi_{1i} = \beta_{10} + r_{1i}$

$\pi_{2i} = \beta_{20} + r_{2i}$

$\pi_{3i} = \beta_{30} + r_{3i}$

BIBLIOGRAPHY

Allison, P. D. (1978). Measures of inequality. *American sociological review*, 865-880.

Allison, S. T., & Messick, D. M. (1990). Social decision heuristics in the use of shared resources. *Journal of Behavioral Decision Making*, *3*(3), 195-204.

Anderson, L. R., Mellor, J. M., & Milyo, J. (2008). Inequality and public good provision: An experimental analysis. *The Journal of Socio-Economics*, *37*(3), 1010-1028.

Aquino, K., Steisel, V., & Kay, A. (1992). The effects of resource distribution, voice, and decision framing on the provision of public goods. *Journal of Conflict Resolution*, *36*(4), 665-687.

Baumert, A., Beierlein, C., Schmitt, M., Kemper, C. J., Kovaleva, A., Liebig, S., & Rammstedt, B. (2014). Measuring four perspectives of justice sensitivity with two items each. *Journal of Personality Assessment*, *96*(3), 380-390.

Bharadwaj, P., & Janoff-Bulman, R. (in preparation). Fairness and social justice: The politics of individuating versus group-based judgments.

Bianchi, E. C. (2016). American individualism rises and falls with the economy: Cross-temporal evidence that individualism declines when the economy falters. *Journal of Personality and Social Psychology*, *111*(4), 567-584.

Brewer, M. B., Hong, Y., & Li, Q. (2004). Dynamic entitativity: Perceiving groups as actors. In V. Yzerbyt, C. M. Judd, & O. Corneille (Eds.), *The psychology of group perception: Contributions to the study of homogeneity, entitativity and essentialism* (pp. 25-38). Philadelphia, PA: Psychology Press.

Brickman, P., Folger, R., Goode, E., & Schul, Y. (1981). Microjustice and macrojustice. In M. J. Lerner & S. C. Lerner (Eds.), *The justice motive in social behavior* (pp. 173-202). NY: Plenum.

Brosnan, S. F., & De Waal, F. B. (2003). Monkeys reject unequal pay. *Nature*, *425*(6955), 297-299.

Cameron, L. A. (1999). Raising the stakes in the ultimatum game: Experimental evidence from Indonesia. *Economic Inquiry*, *37*(1), 47-59.

Carnes, N. C., & Janoff-Bulman, R. (2016). Restraining self-interest or enabling altruism: Morality and politics. In P. Valdesolo & J. Graham (Eds.), *Bridging ideological divides* (pp. 123-142). NY: Routledge/Psychology Press.

Carnes, N. C., & Lickel, B. (under review). Moral binding: How emotions, convictions, and identity fusion shape progroup behavior.

Carnes, N. C., Lickel, B., & Janoff-Bulman, R. (2015). Shared perceptions: Morality is embedded in social contexts. *Personality and Social Psychology Bulletin*, *41*(3), 351-362.

Cherry, T. L., Kroll, S., & Shogren, J. F. (2005). The impact of endowment heterogeneity and origin on public good contributions: evidence from the lab. *Journal of Economic Behavior & Organization*, *57*(3), 357-365.

Choshen-Hillel, S., Shaw, A., & Caruso, E. M. (2015). Waste management: How reducing partiality can promote efficient resource allocation. *Journal of Personality and Social Psychology*, *109*(2), 210-231.

Clayton, S., & Opotow, S. (2003). Justice and identity: Changing perspectives on what is fair. *Personality and Social Psychology Review*, *7*(4), 298-310.

Colasante, A., & Russo, A. (2016). Voting for the distribution rule in a public good game with heterogeneous endowments. *Journal of Economic Interaction and Coordination*, 1-25.

Crimston, D., Bain, P. G., Hornsey, M. J., & Bastian, B. (2016). Moral expansiveness: Examining variability in the extension of the moral world. *Journal of Personality and Social Psychology, 111*(4), 636-653.

Dawes, C. T., Fowler, J. H., Johnson, T., McElreath, R., & Smirnov, O. (2007). Egalitarian motives in humans. N*ature*, *446*(7137), 794-796.

De Cremer, D., & Van Vugt, M. (1999). Social identification effects in social dilemmas. *European Journal of Social Psychology*, *29*(7), 871-93.

DeScioli, P. (2016). The side-taking hypothesis for moral judgment. *Current Opinion in Psychology*, *7*, 23-27.

DeScioli, P., & Kurzban, R. (2009). Mysteries of morality. *Cognition, 112*(2), 281-299.

DeScioli, P., & Kurzban, R. (2013). A solution to the mysteries of morality. *Psychological Bulletin, 139*(2), 477-496.

Deutsch, M. (1985). Distributive justice: A social-psychological perspective. New Haven, CT: Yale University Press.

Dovidio, J. F. (1984). Helping behavior and altruism: An empirical and conceptual overview. *Advances in Experimental Social Psychology*, *17*, 361-427.

Dreber, A., Rand, D. G., Fudenberg, D., & Nowak, M. A. (2008). Winners don't punish. *Nature*, *452*(7185), 348-351.

Durkheim, E. (1893/1984). *The division of labor in society*. NY: Free Press.

Ellemers, N., Pagliaro, S., Barreto, M., & Leach, C. W. (2008). Is it better to be moral than smart? The effects of morality and competence norms on the decision to work at group status improvement. *Journal of Personality and Social Psychology*, *95*(6), 1397-1410.

Ellemers, N., & van den Bos, K. (2012). Morality in Groups: On the Social-Regulatory Functions of Right and Wrong. *Social and Personality Psychology Compass*, *6*(12), 878-889.

Enders, C. K., & Tofighi, D. (2007). Centering predictor variables in cross-sectional multilevel models: a new look at an old issue. *Psychological Methods, 12*(2), 121-138.

Fehr, E., & Gächter, S. (2000). Fairness and retaliation: The economics of reciprocity. *The Journal of Economic Perspectives*, *14*(3), 159-181.

Fehr, E., & Gächter, S. (2002). Altruistic punishment in humans. *Nature*, *415*(6868), 137-140.

Fischbacher, U. (2007). z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, *10*(2), 171-178.

Gelfand, M. J., Raver, J. L., Nishii, L., Leslie, L. M., Lun, J., Lim, B. C., ... & Aycan, Z. (2011). Differences between tight and loose cultures: A 33-nation study. S*cience*, *332*(6033), 1100-1104.

Gómez, A., Brooks, M. L., Buhrmester, M. D., Vázquez, A., Jetten, J., & Swann Jr, W. B. (2011). On the nature of identity fusion: insights into the construct and a new measure. *Journal of Personality and Social Psychology*, *100*(5), 918-933.

Graham, J. (2013). Mapping the moral maps: From alternate taxonomies to competing predictions [Commentary on Janoff-Bulman & Carnes]. *Personality and Social Psychology Review, 17,* 237-241.

Graham, J., & Haidt, J. (2010). Beyond beliefs: Religions bind individuals into moral communities. *Personality and Social Psychology Review, 14*(1), 140-150.

Graham, J., Nosek, B. A., Haidt, J., Iyer, R., Koleva, S., & Ditto, P. H. (2011). Mapping the moral domain. *Journal of Personality and Social Psychology, 101*(2), 366-385.

Graham, J., Waytz, A., Meindl, P., Iyer, R., & Young, L. (in press). Centripetal and centrifugal forces in the moral circle: Competing constraints on moral learning. *Cognition*.

Haidt, J. (2007). The new synthesis in moral psychology. *Science, 316*, 998-1002.

Haidt, J. (2008). Morality. *Perspectives on Psychological Science, 3*(1), 65-72.

Haidt, J. (2012). *The righteous mind: Why good people are divided by politics and religion*. NY: Vintage.

Haidt, J., & Graham, J. (2009). Planet of the Durkheimians, Where Community, Authority, and Sacredness are Foundations of Morality. In J. Jost, A. C. Kay, & H. Thorisdottir (Eds.), *Social and Psychological Bases of Ideology and System Justification*. (pp. 371-401). NY: Oxford.

Haidt, J., & Joseph, C. (2004). Intuitive ethics: How innately prepared intuitions generate culturally variable virtues. *Daedalus, 133*(4), 55-66.

Haidt, J., & Kesebir, S. (2010). Morality. In S. Fiske, & D. Gilbert (Eds.) *Handbook of Social Psychology, 5th Edition*. NY: John Wiley & Sons.

Haslam, N., Rothschild, L., & Ernst, D. (2000). Essentialist beliefs about social categories. *British Journal of Social Psychology, 39,* 127-139.

Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. NY: Guilford Press.

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). Most people are not WEIRD. *Nature, 466*(7302), 29-29.

Janoff-Bulman, R., & Carnes, N. C. (2013). Surveying the moral landscape: Moral motives and group-based moralities. *Personality and Social Psychology Review, 17*, 219-236.

Janoff-Bulman, R., & Carnes, N. C. (2016a). Social Justice and Social Order: Binding Moralities across the Political Spectrum. *PLOS ONE, 11*(3), e0152479.

Janoff-Bulman, R., & Carnes, N. C. (2016b). Morality: Hypo-egoic for us but not them. In K. Brown & M. Leary (Eds.), *Oxford handbook of hypo-egoic phenomena* (pp. 175-188). NY: Oxford University Press.

Janoff-Bulman, R., Sheikh, S., & Baldacci, K. (2008). Mapping moral motives: Approach, avoidance, and political orientation. *Journal of Experimental Social Psychology, 44*(4), 1091-1099.

Johnson, T., Dawes, C. T., Fowler, J. H., McElreath, R., & Smirnov, O. (2009). The role of egalitarian motives in altruistic punishment. *Economics Letters, 102*(3), 192-194.

Kelley, H. H., & Thibaut, J. W. (1978). *Interpersonal Relations: A Theory of Interdependence*. NY: Wiley.

Kelley, H. H., Holmes, J. G., Kerr, N. L., Reis, H. T., Rusbult, C. E., & Van Lange, P. A. M. (2003). *An Atlas of Interpersonal Situations*. NY: Cambridge.

Kline, R. B. (2015). *Principles and Practice of Structural Equation Modeling*. NY: Guilford publications.

Kohlberg, L. (1981). *Essays on moral development, vol. I: The philosophy of moral development*. San Francisco, CA: Harper & Row.

Kramer, R. M., & Brewer, M. B. (1984). Effects of group identity on resource use in a simulated commons dilemma. *Journal of Personality and Social Psychology*, *46*(5), 1044-1057.

Kroll, S., Cherry, T. L., & Shogren, J. F. (2007). Voting, punishment, and public goods. *Economic Inquiry*, *45*(3), 557-570.

Kurzban, R., & Houser, D. (2005). Experiments investigating cooperative types in humans: A complement to evolutionary theory and simulations. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(5), 1803-1807.

De Kwaadsteniet, E. W., van Dijk, E., Wit, A., & de Cremer, D. (2006). Social dilemmas as strong versus weak situations: Social value orientations and tacit coordination under resource size uncertainty. *Journal of Experimental Social Psychology*, *42*(4), 509-516.

Levine, M., Cassidy, C., Brazier, G., & Reicher, S. (2002). Self-Categorization and Bystander Non-intervention: Two Experimental Studies1. *Journal of Applied Social Psychology*, *32*(7), 1452-1463.

Lind, E. A., & Tyler, T. R. (1988). *The social psychology of procedural justice*. NY: Plenum.

List, J. A., & Cherry, T. L. (2000). Learning to accept in ultimatum games: Evidence from an experimental design that generates low offers. *Experimental Economics*, *3*(1), 11-29.

Litman, L., Robinson, J., & Abberbock, T. (2016). TurkPrime. com: A versatile crowdsourcing data acquisition platform for the behavioral sciences. *Behavior Research Methods*, 1-10.

Little, T. D., Cunningham, W. A., Shahar, G., & Widaman, K. F. (2002). To parcel or not to parcel: Exploring the question, weighing the merits. *Structural Equation Modeling*, *9*(2), 151-173.

Masclet, D., & Villeval, M. C. (2008). Punishment, inequality, and welfare: a public good experiment. *Social Choice and Welfare*, *31*(3), 475-502.

McFarland, S., Brown, D., & Webb, M. (2013). Identification with all humanity as a moral concept and psychological construct. *Current Directions in Psychological Science*, *22*(3), 194-198.

Mullen, E., & Skitka, L. J. (2006). Exploring the psychological underpinnings of the moral mandate effect: motivated reasoning, group differentiation, or anger?. *Journal of Personality and Social Psychology*, *90*(4), 629.

Nikiforakis, N. (2008). Punishment and counter-punishment in public good games: Can we really govern ourselves?. *Journal of Public Economics*, *92*(1), 91-112.

Nishi, A., Shirado, H., Rand, D. G., & Christakis, N. A. (2015). Inequality and visibility of wealth in experimental social networks. *Nature*, *526*(7573), 426-429.

Norton, M. I., & Ariely, D. (2011). Building a better America—One wealth quintile at a time. *Perspectives on Psychological Science, 6*(1), 9-12.

Ostrom, E., Walker, J., & Gardner, R. (1992). Covenants with and without a Sword: Self-governance Is Possible. *American Political Science Review, 86*(02), 404-417.

Pagliaro, S., Ellemers, N., & Barreto, M. (2011). Sharing moral values: Anticipated ingroup respect as a determinant of adherence to morality-based (but not competence-based) group norms. *Personality and Social Psychology Bulletin, 37*(8), 1117-1129.

Piketty, T. (2014). *Capital in the twenty-first century* (A. Goldhammer, Trans.). Cambridge, MA: Harvard University Press.

Piketty, T., Saez, E., & Zucman, G. (2016). *Distributional national accounts: Methods and estimates for the United States*. Washington, DC: Washington Center for Equitable Growth.

Rand, D. G., & Nowak, M. A. (2011). The evolution of anti-social punishment in optional public goods games. *Nature Communications, 2*, 434.

Rand, D. G., & Nowak, M. A. (2013). Human cooperation. *Trends in Cognitive Sciences, 17*(8), 413-425.

Rapoport, A., Bornstein, G., & Erev, I. (1989). Intergroup competition for public goods: Effects of unequal resources and relative group size. *Journal of Personality and Social Psychology, 56*(5), 748-756.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (Vol. 1). Thousand Oaks, CA: Sage.

Raudenbush, S. W., Spybrook, J., Congdon, R., Liu, X., Martinez, A., Bloom, H., & Hill, C. (2011). Optimal design plus empirical evidence. *William T. Grant Foundation, NY*.

Rawls, J. (1971/2009). *A theory of justice*. Cambridge, MA: Harvard University Press.

Richerson, P. J., & Boyd, R. (2005). Not by genes alone: How culture transformed human evolution. Chicago, IL: University of Chicago Press.

Rock, M. S., & Janoff-Bulman, R. (2010). Where do we draw our lines? Politics, rigidity, and the role of self-regulation. *Social Psychological and Personality Science*, *1*(1), 26-33.

Rusbult, C. E., & Van Lange, P. A. (2008). Why we need interdependence theory. *Social and Personality Psychology Compass*, *2*(5), 2049-2070.

Scheve, K., & Stasavage, D. (2010). The conscription of wealth: mass warfare and the demand for progressive taxation. *International Organization*, *64*(04), 529-561.

Scheve, K., & Stasavage, D. (2012). Democracy, war, and wealth: lessons from two centuries of inheritance taxation. *American Political Science Review*, *106*(01), 81-102.

Scheve, K., & Stasavage, D. (2016). *Taxing the Rich: A History of Fiscal Fairness in the United States and Europe*. Princeton, NJ: Princeton University Press.

Schmitt, M., Baumert, A., Gollwitzer, M., & Maes, J. (2010). The justice sensitivity inventory: Factorial validity, location in the personality facet space, demographic pattern, and normative data. *Social Justice Research*, *23*(2-3), 211-238.

Shaw, A. (2013). Beyond "to share or not to share" The impartiality account of fairness. *Current Directions in Psychological Science*, *22*(5), 413-417.

Shaw, A., Choshen-Hillel, S., & Caruso, E. M. (2016). The development of inequity aversion understanding when (and why) people give others the bigger piece of the pie. *Psychological Science*, *27*(10), 1352-1359.

Shaw, A., Montinari, N., Piovesan, M., Olson, K. R., Gino, F., & Norton, M. I. (2014). Children develop a veil of fairness. *Journal of Experimental Psychology: General*, *143*(1), 363-375.

Shaw, A., & Olson, K. R. (2012). Children discard a resource to avoid inequity. *Journal of Experimental Psychology: General*, *141*(2), 382-395.

Shaw, A., & Olson, K. (2014). Fairness as partiality aversion: The development of procedural justice. *Journal of Experimental Child Psychology*, *119*, 40-53.

Shweder, R. A., Much, N. C., Mahapatra, M., & Park, L. (1997). The "big three" of morality (autonomy, community, divinity) and the "big three" explanations of suffering. In A. Brandt & P. Rozin (Eds.), *Morality and Health* (pp. 119-169). NY: Routledge.

Skitka, L. J. (2003). Of different minds: An accessible identity model of justice reasoning. *Personality and Social Psychology Review*, *7*(4), 286-297.

Slonim, R., & Roth, A. E. (1998). Learning in high stakes ultimatum games: An experiment in the Slovak Republic. *Econometrica*, 569-596.

Sonnemans, J., Van Dijk, F., & Van Winden, F. (2006). On the dynamics of social ties structures in groups. *Journal of Economic Psychology*, *27*(2), 187-204.

Swann Jr, W. B., & Buhrmester, M. D. (2015). Identity fusion. *Current Directions in Psychological Science*, *24*(1), 52-57.

Tajfel, H. E. (1978). *Differentiation between social groups: Studies in the social psychology of intergroup relations*. Cambridge, MA: Academic Press.

Tajfel, H. E., Billig, M. G., Bundy, R. P., & Flament, C. (1971). Social categorization and intergroup behaviour. *European Journal of Social Psychology*, *1*(2), 149-178.

Tavoni, A., Dannenberg, A., Kallis, G., & Löschel, A. (2011). Inequality, communication, and the avoidance of disastrous climate change in a public goods game. *Proceedings of the National Academy of Sciences*, *108*(29), 11825-11829.

Thomas, K. A., De Freitas, J., DeScioli, P., & Pinker, S. (2016). Recursive mentalizing and common knowledge in the bystander effect. *Journal of Experimental Psychology: General*, *145*(5), 621-629.

Thomas, K. A., DeScioli, P., Haque, O. S., & Pinker, S. (2014). The psychology of coordination and common knowledge. *Journal of Personality and Social Psychology*, *107*(4), 657-676.

Tricomi, E., Rangel, A., Camerer, C. F., & O'Doherty, J. P. (2010). Neural evidence for inequality-averse social preferences. *Nature*, *463*(7284), 1089-1091.

Turiel, E. (1983). *The development of social knowledge: Morality and convention*. Cambridge, MA: Cambridge University Press.

Turner, J. C., Hogg, M. A., Oakes, P. J., Reicher, S. D., & Wetherell, M. S. (1987). *Rediscovering the social group: A self-categorization theory*. Oxford: Basil Blackwell.

Tyler, T. R., & Blader, S. L. (2000). *Cooperation in groups: Procedural justice, social identity, and behavioral engagement*. NY: Psychology Press.

Van Dijk, F., Sonnemans, J., & van Winden, F. (2002). Social ties in a public good experiment. *Journal of Public Economics, 85*(2), 275-299.

Van Huyck, J. B., Gillette, A. B., & Battalio, R. C. (1992). Credible assignments in coordination games. *Games and Economic Behavior, 4*(4), 606-626.

Van Lange, P. A. (1999). The pursuit of joint outcomes and equality in outcomes: An integrative model of social value orientation. *Journal of Personality and Social Psychology*, *77*(2), 337-349.

Walster, E., Walster, G. W., & Berscheid, E. (1978). *Equity: Theory and research.* Boston, MA: Allyn & Bacon.

Wenzel, M. (2000). Justice and identity: The significance of inclusion for perceptions of entitlement and the justice motive. *Personality and Social Psychology Bulletin, 26*(2), 157-176.

Wenzel, M. (2002). What is social about justice? Inclusive identity and group values as the basis of the justice motive. *Journal of Experimental Social Psychology, 38*(3), 205-218.

Wenzel, M. (2004). An analysis of norm processes in tax compliance. *Journal of Economic Psychology*, *25*(2), 213-228.

Wilkinson, R. G., & Pickett, K. (2009). *The spirit level: Why more equal societies almost always do better*. NY: Allen Lane.

Wilson, E. O. (2012). *The social conquest of earth*. NY: Norton & Company.