2015

An attribution model of trust extension across group boundaries

Shiang-Yi Lin
Lehigh University

An Attribution Model of Trust Extension Across Group Boundaries

by

Shiang-Yi Lin

A Thesis

Presented to the Graduate and Research Committee

of Lehigh University

in Candidacy for the Degree of

Master of Sciences

in

Psychology Department

Lehigh University

12/18/2014

Thesis is accepted and approved in partial fulfillment of the requirements for the Master of Science in Psychology Department.

An Attribution Model of Trust Extension Across Group Boundaries
Shiang-Yi Lin

_____
12/17/2014

_____
Dominic Packer
Associate Professor
Advisor

_____
Michael Gill
Associate Professor
Committee member

_____
Christopher Burke
Assistant Professor
Committee member

_____
Gordon Moskowitz
Professor
Department chair

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Abstract

We reasoned that observing high levels of cooperation among outgroup members might be threatening, causing perceivers to expect little cooperation across group boundaries. Alternately, cooperation among outgroup members might be interpreted dispositionally, causing perceivers to expect cooperation to extend across group boundaries. Across two studies, participants were assigned to a minimal group and observed a series of players – either outgroup-pairs, ingroup-pairs or intergroup-pairs – play prisoner's dilemma games and make overwhelmingly cooperative decisions (90%). Results were consistent with the dispositional rather than the threat hypothesis. Positive cooperative expectations and dispositional inferences for outgroup targets were greatest in the outgroup-pairs condition, followed by the intergroup condition, followed by the ingroup-pairs condition. Effects were not moderated by a possible situational attribution (presence of a third party punisher). Without stereotypes or intergroup conflict, perception of outgroup targets was based on individual-level behavioral evidence – more instances of cooperation translated into stronger dispositional inferences.

Introduction

> "If you want to make peace with your enemy, you have to work with your enemy. Then he becomes your partner." – Nelson Mandela

These are insightful words by Nelson Mandela, who waged a multiple decades' fight against the racial segregation policy ('apartheid)' of South Africa, and they are reflective of a process via which foes can become allies and hatred can turn into rapport. Nelson Mandela highlights that a solution to intergroup conflicts is to initiate cooperation across group boundaries, and psychologists have investigated the conditions under which trust can be built between members from two competing groups (Allport, 1954; Pettigrew, 1991). For example, one common phenomenon in intergroup conflict is that members distribute more resources to their own group than to members of an outgroup; this differential treatment can be mitigated if group members who themselves have no contact with the competing group are made aware of the existence of a cross-group friendship (Wright, Aron, McLaughlin-Volpe, & Ropp, 1997). In other words, instances of successful intergroup interactions can potentially increase an observer's positive expectations for future cross-group cooperation. However, the existing literature rarely discusses what people infer from their observations of intergroup interactions, and whether this inferential process predicts the expectations for future cross-group cooperation. In the present studies, I will focus on the role of social attribution and investigate its influence on perceived

cooperative expectations (inferential aspect) and subsequent intergroup cooperation (behavioral aspect).

Intergroup Trust

According to Good (1988), "*trust is based on an individual's theory as to how another person will perform on some future occasion, as a function of that target person's current and previous claims, either implicit or explicit, as to how they will behave*." The particular theory a person develops about how another individual is likely to behave in the future may be based on a number of factors, including prior experiences with the individual and reputational information about him or her (Deutsch, 1958; Luhmann, 1979). However, trust is not necessarily contingent on prior interpersonal contact (Brewer, 2008); in fact, it can be established in the absence of direct experiences with a target if we obtain other information that helps to establish their trustworthiness (Tanis & Postmes, 2005). For example, learning simply about their group membership can be a powerful source of information. Even an arbitrarily assigned common group membership can foster trust among strangers (Brewer & Kramer, 1986; Tajfel & Turner, 1979).

Social dilemmas have been used to investigate how group identity induces cooperative behaviors. In social dilemmas, participants are presented with decisions in which personal goals and collective goals are in conflict. For example, in a

prisoners' dilemma, the best joint outcome requires cooperation by both parties; although defection can possibly result in an outcome maximizing one's own interest, it can also lead to the greatest collective losses if both parties simultaneously decide to defect. Considerable research has found that when people share a group membership, individuals make decisions that favor the group as a whole at the expense of their own interests (Brewer & Kramer, 1986; De Cremer & Van Vugt, 1999; Tanis & Postmes, 2005). This results in, among other things, more cooperative behaviors among ingroup members in social dilemmas.

Packer and Kugler (2013) proposed the Cooperative Contingencies Model (CCM) to articulate conditions when people are more or less likely to rely on common group identities to facilitate cooperation and trust. According to CCM, choices to selectively coordinate with ingroup members are prevalent in social interactions because group memberships provide predictable incentives. Individuals often need no prior knowledge about ingroup partners to trust them, believing that ingroup partners will act altruistically based on the group norm (Brewer, 2008; Tanis & Postmes, 2005). With respect to ingroup affiliation observed in social dilemmas, a debate has emerged over whether individuals treat their own group as "*an interchangeable component*" of themselves and thus show a genuine care for the group members due to a common group identity (Brewer, 2008; De Cremer & Van

Vugt, 1999; Tanis & Postmes, 2005; Tajfel & Turner, 1979) or demonstrate ingroup preferences only because mutual reciprocity is forthcoming (Foddy, Platow, & Yamagishi, 2009; Yamagishi & Kiyonari, 2000). As is emphasized by social identity theory, loyalty and commitment to the group contribute to ingroup favoritism as a means of social coordination, especially in cases where there is competition or dislike between groups or when people are highly identified with their group (De Cremer & Van Vugt, 1999; Hewstone, Rubin, & Willis, 2002). As such, it is likely that shifts in self-concept and expectations of reciprocity both contribute to heightened cooperation within group boundaries.

Importantly, shared group memberships are not the only means to help achieve successful coordination (i.e., they are only one reason a person might have a "theory" that someone else will behave in a trustworthy fashion). Other factors or mechanisms can also facilitate cooperation – and if they are perceived as more effective than ingroup affiliation they might decrease intergroup biases. For example, when people believe that a mechanism or social institution that facilitates cooperation is effective enough to ensure social coordination, individuals may no longer prefer ingroup members over outgroup members. These are exactly our findings in a series of studies. We have found that when a cooperation-facilitating mechanism exists (e.g., the presence of a third party which either punishes unfair

transactions or rewards fair transactions in a trust game), individuals extend similar levels of trust toward ingroup and outgroup members, thereby decreasing intergroup biases that are prevalent in the absence of a cooperation-facilitating mechanism (Packer & Kugler, under review). These effects are strongest among weakly identified group members because they tend to freely adopt any effective strategy that enables successful coordination instead of being driven by motives or ideologies based on group identities (as highly identified group members tend to be). In a related study, we also found that individuals who were informed about the presence of a cooperation-facilitating mechanism displayed lower implicit racial bias toward outgroup members than did individuals who were not informed about the existence of the mechanism (Lin & Packer, 2014). (I will discuss this series of studies in depth in the section on sanctioning systems.)

Taken together, people often perceive shared group memberships as an effective means to accomplish a positive social outcome. In situations that require cooperation, they may preferentially cooperate with ingroup members so as to achieve a desirable result if other cooperative affordances do not exist. Importantly for the current research, the extent to which groups are perceived as promising cooperative affordances is also likely to vary depending on how cohesive they are perceived to be – i.e., the extent to which members of that group are perceived to

like each other and work well together (Campbell, 1958). The more cohesive a group is perceived to be, the more its members are expected to contribute their efforts to accomplish a group goal or to perform collectively normative behaviors (Abelson, Dasgupta, Park, & Banaji, 1998). Therefore, a highly cohesive group suggests a promising cooperative affordance if group goals are perceived as being compatible with the goals of an individual. This should increase the likelihood that the individual will affiliate with the ingroup in order to achieve a favorable outcome. Conversely, if the goals of a group are perceived as incompatible with the goals of an individual, cohesiveness renders the group a less promising cooperative affordance; this may be particular likely when an outgroup is perceived as cohesive (Packer & Kugler, 2013).

Perception of Intragroup Interactions

Our research to date has focused on how people use group membership information to develop theories or expectations about the likely trustworthiness of other people – in the absence of other knowledge. These situations involve interacting with novel partners who participants have never interacted with or observed before.

There are several approaches to investigate how group membership information influences interpretations of an individual's behavior. Scholars who

portray ingroup cooperation as heuristics declare that a heightened group identity can activate corresponding behaviors under different task structures (see review by Weber, Kopelman, & Messick, 2004). Brewer (2008) also remarks on this propensity to align one's behaviors with a group cue as "a cooperative script" (p. 221). Researchers on stereotype activation and application approach this question by examining how social categories shape social perception (Blair, 2002; Devine, 1989). They suggest that use of group memberships to interpret others' behavior seems almost inevitable (Devine, 1989). In doing so, people do not perceive a social target as a unique individual, but instead interpret his or her behavior in the light of group templates (e.g., stereotypes). In other words, when a group membership is perceived and applied to interpret a social target's behavior, the behavior has its meaning changed in line with the group identity.

In the present studies, I am interested in *how people interpret direct observations of cooperation among people belonging to different groups*. Specifically, I investigate the attributions (and subsequently the trust decisions) that people make when they observe ingroup and outgroup members cooperating - both within and across group boundaries.

In order to generate specific predictions, I will review the attribution literature and formulate a theoretical framework to understanding how cooperative

expectations may be inferred from observed intragroup/intergroup interactions, and how these inferences may affect subsequent cooperative behaviors. We aim to combine two lines of research— social attribution (e.g., Pettigrew, 1979) and strategic cooperation in economic games (e.g., Packer & Kugler, 2013). By doing so, we hope that this framework will help us to better understand: (1) how cooperative behaviors demonstrated by ingroup or outgroup partners engaging in intragroup/intergroup interactions may be interpreted differently, and (2) how these interpretations may influence perceived cooperative expectations and, in turn, alter the likelihood of engaging in subsequent cross-group cooperation. Broadly, we predict that because cooperative expectations may be asymmetrically inferred when observing cooperative acts among ingroup members vs. outgroup members, it may lead decision-makers to exhibit intergroup biases in subsequent interactions (see Figure 1).

Classic attribution theories illuminate the ways in which people give explanations for others' behaviors (Gilbert, 1998; Heider, 1958; Jones & Davis, 1965). Heider (1958) describes attribution processes in terms of a causal model, such that behaviors are teased apart into the components driven by an actor's enduring qualities versus the components driven by their environments and situational causes. Close to our main interest in this project—perceptions of

cooperative dispositions of targets playing economic games—Jones and Davis (1965) and Kelley (1967) also address how people infer the dispositions of others from socially desirable behaviors. According to the *covariation principles* (Jones & Davis, 1965), the dispositions or intentions of an actor are inferred on the basis of the effects correlating or covarying with his or her behavior; the dispositions of an actor can be revealed if other people in the same situation would not behave in the same way (e.g., Ivy is cooperative because she sacrificed her sleep to work on a group project, whereas most of the people did not), or if an actor's choice brings about an effect that cannot be attained by another choice (e.g., Ivan is cooperative because he could have took a day off from the group project, but he decided not to; Gilbert, 1998; Jones & Davis, 1965; Kelley, 1967).

The key question for a decision-maker who observes other people behaving cooperatively is: to what extent is their behavior driven by stable individual cooperative dispositions (in which case they are likely to cooperate in the future) vs. driven by some external non-stable cause (in which case they may not cooperate going forward). In group contexts, people may also make inferences about the groups to which the observed actors belong (see Figure 1)—discussed in more depth below.

How do groups influence attributions? Research has shown that the connotations of a behavior change along with group membership of the actor (Devine, 1989; Duncan, 1976). For example, Duncan (1976) showed that an ambiguous behavior was rated by Caucasian college students as more aggressive if it was carried out by African American targets than by Caucasian American targets. Stereotyping can legitimize the status quo among groups (Hamilton & Sherman, 1994), and attribution processes also play a crucial role in maintaining group boundaries by reproducing or confirming existing views of ingroups and outgroups (Pettigrew, 1979; Sherman, Stroessner, Conrey, & Azam, 2005). The ultimate attribution error (a group-serving bias) refers to a bias to explain groups' behaviors in a way that enhances ingroup superiority or confirms existing prejudices (Hewstone, 1990; Pettigrew, 1979). Specifically, negative behaviors performed by outgroup members are more likely to be attributed to internal causes (i.e., dispositions or intentions), compared to the same behaviors performed by ingroup members. In contrast, positive behaviors performed by outgroup members are more likely to be ascribed to external causes (i.e., situational factors), compared to the same behaviors performed by ingroup members.

Following this logic, we predict that cooperation—a positive behavior—by outgroup members is more likely to be inferred as driven by causes other than

dispositions. Cooperation by ingroup members, on the other hand, may be attributed to stable aspects of their personality. In contrast, defection by outgroup members would more likely be attributed to malevolent intentions or dispositions, whereas equivalent acts of defection by ingroup members may be explained away situationally. Concretely, this implies that when decision-makers perceive exactly the same cooperative acts by ingroup and outgroup members, they are likely to infer that these behaviors reflect stable prosocial dispositions among ingroup members but are due to other external causes among outgroup members.

What cause might perceivers infer drives cooperative behavior among outgroup members? Following from the Cooperative Contingencies approach, one key factor might be the fact that the target individuals share a *common group membership*. That is, perceivers may infer that outgroup members were only cooperative – with one another – because they share a common group membership. However, because they are not perceived as cooperative at an individual dispositional level, it may be assumed that they are unlikely to be cooperative outside their group boundaries (i.e., with people like the perceivers themselves) (see Figure 1).

These type of attribution biases – if they exists – may then contribute to differential perception of cooperative opportunities afforded by ingroup and

outgroup members, and thereby influence subsequent cooperative decisions. According to CCM, decision-makers take all available cooperative affordances into account to predict the possibility of successful social coordination (Packer & Kugler, 2013). As such, observing cooperation by ingroup members may evoke greater cooperative expectations (i.e., greater expectations that they will cooperate with participants themselves) than cooperation by outgroup members, due to the fact that cooperation by ingroup members is more likely to be attributed to internal causes such as genuine concern or cooperative dispositions than cooperation by outgroup members (Hewstone, 1990; Pettigrew, 1979). In sum, an intergroup attribution bias may lead to asymmetric inferences about the cooperative opportunities afforded by ingroup vs. outgroup members, resulting in cooperative decisions in favor of ingroup affiliation (see Figure 1).

Further, perceptual features of intragroup interactions may influence the inferences made not only about the meaning of individuals' behavior, but also characteristics of the groups they belong to (see review by Yzerbyt, Corneille, & Estrada, 2001). For an observer, coordinated interactions among a group of social others may sometimes cause them to be perceived as part of a unified and cohesive group, a construct termed 'entitativity' (Campbell, 1958). Entitativity increases with factors such as perceptual proximity or similarity, as well as interdependence or a

common fate among targets. Research has shown that entitativity facilitates the use of group attributes in interpreting group behaviors (Yzerbyt & Rogier, 2001). For example, in a quiz game paradigm (Ross, Amabile, & Steinmetz, 1977), questioners, who were participants randomly assigned to ask quizzes to others, were consistently rated by an observer as more intelligent than were participants randomly assigned to the role of answerers. This overattribution bias was magnified when the questioner or answerer group was described as highly entitative (i.e., coming from the same school) vs. weakly entitative (e.g., coming from different schools) (Yzerbyt, Rogier, & Fiske, 1998). Further, members of a highly entitative group are perceived as more likely to carry out negative behaviors (Dasgupta, Banaji, & Abelson, 1999), adopt malicious intentions (Abelson, Dasgupta, Park, & Banaji, 1998; Insko & Schopler, & Sedikides, 1998) and are believed to be more accountable for wrongdoings than members of a weakly entitative group (Newheiser, Sawaoka, & Dovidio, 2012).

Whereas most prior research in this tradition has focused on inferences following negative outgroup behaviors (Dasgupta et al., 1999; Newheiser et al., 2012), we are interested in the consequences of observing positive interactions among outgroup members. From the evidence above, we propose that observing cooperative intragroup interactions among outgroup members may impede subsequent intergroup cooperation. Based on the CCM and our earlier propositions,

we predict that if an outgroup is perceived as highly cohesive or entitative, its members will be regarded as less promising cooperative partners, which may increase intergroup bias. If perceiving outgroup members cooperating with one another increases perceptions of cohesion, observations of this objectively positive behavior may, ironically, serve to increase bias. Further, outgroup interactions that make its members look coherent and unified may, on occasion, cause the outgroup to be perceived as more hostile and threatening (than if they had not been perceived to engage in a cooperative intragroup interaction). We propose that individuals' levels of identification with their own group may influence the effects of perceptions of outgroup entitativity. As such, we propose that weakly identified group members may tend to situationally attribute outgroup interactions among outgroup members (e.g., due to a shared group membership), thus deeming outgroup members as less promising cooperative partners; in contrast, for highly identified group members, outgroup interactions may evoke threat and directly cause subsequent biased decisions.

Perception of Intergroup Interactions

Returning the example of Nelson Mandela with which we began, another important question has to do with how people interpret cooperative interactions between the members of different groups, extensive research on intergroup contact

has indicated many advantages of contact across group boundaries (Pettigrew & Tropp, 2006): First, ingroup members who engage in a cross-group friendship can provide referent information to other group members about intergroup contact, suggesting that it is an accepted behavior within ingroup norms (Pettigrew, 1991; Wright et al., 1997). Second, observing intergroup contact initiated by ingroup members reduces anxiety that may be elicited due to unfamiliarity with an outgroup (Stroessner & Mackie, 1993). Third, intergroup interactions per se speak to the willingness of outgroup members to foster positive intergroup relationships (Wright et al., 1997).

Given the evidence above, although personal contact with outgroups is not our main focus, it still provides a glimpse about possible mechanisms via which observing intergroup interactions may facilitate cooperative decisions in intergroup contexts. Compared to the mechanisms we elaborated above that might increase intergroup biases when intragroup cooperation among outgroup members is observed, observations of intergroup interactions may result in different inferences – especially about cooperating outgroup members (see Figure 2). In this case, their cooperation cannot be attributed to the external cause of a shared group membership – because they are cooperating across group lines. Further, intergroup cooperation should not increase perceptions of outgroup cohesion or entitativity, which

otherwise may give rise to feelings of threat. As such, perceivers may make equivalently dispositional attributions for both ingroup and outgroup members. However, we believe that observing intergroup interactions may function differently depending on individuals' levels of group identification. Observations of intergroup interactions may facilitate cooperative decisions of both highly and weakly identified individuals, however, via different routes (see Figure 2). We predict that, as is theorized above, weakly identified members may infer that ingroup and outgroup members are equally dispositionally cooperative, and reach similar conclusions about cooperative opportunities afforded by ingroup and outgroup targets, thereby decreasing biased decisions. In contrast, we predict that observing intergroup interactions may function to create a group norm, which may especially encourage unbiased decisions for highly identified members, due to the fact that highly identified members are more strongly influenced by group norms than are weakly identified members (Terry & Hogg, 1996) (see Figure 2).

An Alternate External Attribution: The Role of Sanctions

As evidence for the sort of strategic intergroup biases posited by the CCM, prior research has found that effective sanctioning systems are associated with decreased intergroup biases in lab settings. This conclusion is also supported by nationwide and international survey data, in which trust in social institutions like the

police, government and legal systems predicts lower levels of affiliative bias (Packer & Kugler, under review). Sanctioning systems help to ensure the outcomes of social coordination. Sanctioning systems can facilitate cooperation across group boundaries because people no long rely on group memberships as a means to secure social outcomes or prevent exploitation.

For example, in a prior series of studies, participants engaged in several one-shot trust games (Berg, Dickhaut & McCabe, 1995), in which they could decide to allocate any amount of the points they were initially assigned in each round to a partner and save the rest for themselves. They were informed their partner would receive three times of their decided amount, after which their partner would decide how many points to return. This setting creates outcome dependence on their partner, and accordingly trust in their partner is crucial in determining decisions about how much to invest. In order to examine the role of cooperative-facilitating mechanisms for mitigating intergroup biases, previous investigations in our lab have varied the presence of sanctioning systems. For example, Packer and Kugler (under review) manipulated the presence vs. absence of a third-party punisher or rewarder. Similarly, Lin and Packer (2014), randomly assigned participants to conditions where a third-party punisher was absent or present. In our study 1, the only

difference between conditions was that participants in the punisher condition received an extra instruction:

> "*A third player, Player C, will be watching what happens. Player C won't know anything about Player A or B (e.g., won't see their photos). However, if Player B makes an unfair decision and sends too few points back to Player A, Player C has been instructed to remove most of Player B's points. These points will disappear and nobody will get them*."[1]

Consistent with previous literature on intergroup biases, our findings in conditions without effective cooperation-facilitating mechanisms showed that participants preferentially allocated more to their partners if they belonged to the ingroup instead of the outgroup (arbitrarily assigned novel groups in Packer & Kugler, under review). Critically, Packer and Kugler (under review) found that the presence of a third party punisher or rewarder reduced preferential trust in ingroup (over outgroup partners), especially among low-identified participants. Similarly, faced targets of one's own and other races (i.e., White and Black targets), White participants displayed an implicit racial bias when the punisher was absent: positive

[1] Participants in a second study were instead exposed to the information that varied whether their partners ostensibly knew about the presence of a punisher; half of them received the identical instruction as the punisher condition in the previous study, whereas the rest of them were told that their partner did not know about the existence of a punisher. The pattern of effects was the same.

concepts were more strongly associated with ingroup faces than outgroup faces, whereas negative concepts were more strongly associated with outgroup faces than ingroup faces, suggesting the existence of both ingroup favoritism and outgroup derogation (in Lin & Packer, 2014). However, we found, in the punisher presence condition, implicit racial attitudes were tuned to the cooperative contingencies available in a context: specifically, the responses to positive words following Black faces were significantly facilitated, as opposed to the condition without a punisher. In other words, the presence of an effective cooperation-facilitating mechanism elevated the positive associations with outgroup faces to such an extent that it mirrored the evaluation of ingroup faces.

These findings suggest that rapid evaluative biases and cooperative decisions result, in part, through strategically weighing cooperative affordances and other contingencies salient in a particular context. The efficacy of cooperative affordances determines which coordination strategy people adopt—when group memberships are perceived to be the most effective guarantees of cooperation (e.g., partners share the same membership information), their investment varies as a function of group memberships. However, when another cooperation-facilitating mechanism is an effective guarantee of cooperation and influences the payoff structure (e.g., partners are aware of the sanctions), people forgo group memberships (at least if they are

weakly identified with their ingroup) and are equally cooperative regardless of their partners' group identity.

Because of the strong evidence we obtained about strategic cooperation, an attributional question then arises – *Does this temporary intergroup cooperation enabled by the presence of effective social institutions like sanctioning systems attenuates positive dispositional attributions because it provides an alternate situational attribution*? That is, they may believe that cooperation by their partners simply results from situational requirements instead of anything intrinsically motivated.

Empirical studies on sanctioning systems suggest that they can breed distrust among people and encourage unethical behaviors in the absence of sanctions (e.g., Cialdini, 1996). Similarly, sanctions may propel an economic decision frame instead of an ethical one (Tenbrunsel & Messick, 1999), and evoke greater consideration of defection when avoiding sanctions becomes possible (Mulder, Van Dijk, De Cremer, & Wilke, 2006a). Research also suggests that the presence of sanctions serves as a salient external cause of cooperative behaviors. Such attributions are likely to reduce inferences about internal motivation to cooperate (Cialdini, 1996; Mulder, Van Dijk, De Cremer, & Wilke, 2006b). Further, Mulder et al. (2006b) found that the presence of sanctioning systems dampens trust when sanctions no

longer exist, especially among individuals who were intrinsically motivated to trust or who were initially led to believe their partners were trustworthy.

Based on attribution theories, sanctions covarying with cooperation by an actor make his or her intentions open to explanations other than internal causes, such as prosociality or cooperativeness (Jones & Davis, 1965; Kelley, 1967). As such, we posited that the sanctions may become the main cause that an observer believes to drive an actor's behavior, leaving dispositional or group-based causes as less explanatory (see Figure 3). Study 2 introduced sanctioning systems into our framework, examining whether sanctions altered attributions following observations of cooperative intragroup or intergroup interactions. Specifically, we investigated to what extent the benefits of observing cooperative intergroup interactions (e.g., formation of equivalent cooperative expectations for the ingroup and outgroup as mentioned in the previous section) were eliminated when a sanctioning system existed. Overall, we expected that participants would increase causal inferences related to sanctions (e.g., a fear of being punished) and therefore reduce attributions to both dispositional causes and a shared group membership in the presence of a sanctioning system.

Overview

In the present studies, we investigated how people would infer cooperative dispositions and opportunities from intragroup/intergroup interactions, and to what extent they made cooperative decisions in line with these inferences. We focused on how the inferences about cooperative expectations would be unequally drawn due to biased attributions of ingroup vs. outgroup interactions and in turn cause discriminative cooperative decisions. Study 1 focused on attributions made when individuals observed cooperation within ingroup boundaries, outgroup boundaries and across group lines. These conditions were contrasted to two control conditions, in which participants observed an ingroup or an outgroup member cooperating with a target whose group membership was unknown. In Study 2, following some methodological adjustments based on the results of Study 1, we investigated how an alternate situational force – a sanctioning system – would affect these attribution processes.

STUDY 1

Study 1 examined whether individuals would engage in biased attribution processes, such that cooperation by outgroup members tended to be externally attributed (to the fact that they share a group membership), whereas cooperation by ingroup members tended to be internally attributed. Participants observed a scenario

with two players participating in a repeated Prisoners' Dilemma game and rapidly settling to cooperative choices within 3 rounds (out of a total of 10 rounds). We hoped to demonstrate that due to attribution biases, greater cooperative expectations would be inferred from observing cooperation by ingroup players than by outgroup players despite them displaying what was objectively the same level of cooperation. Furthermore, Study 1 was also designed to investigate whether observations of cooperative intergroup interactions would mitigate intergroup biases. Intergroup interactions may give rise to equally cooperative expectations for both ingroup and outgroup targets for weakly identified members, whereas they may serve as a guiding group norm for highly identified members.

Method

Participants

Three hundred and fifty-six participants were recruited online through Mechanical Turk for a monetary reward of $1. The study had a 5 (scenario: Ingroup–Ingroup vs. Outgroup–Outgroup vs. Intergroup vs. Ingroup–Control vs. Outgroup–Control) between-subjects design.

Procedure

The ostensible purpose of the study was "to investigate people's understanding of behavior in social decision-making tasks." Participants were

introduced to a dot estimation task modified from Howard and Rothbart (1980) and Gerard and Hoyt (1974): "*First of all, you will be participating in a task relating to how people make quantitative judgments. Past studies have shown that, given the task of estimating how many objects they have seen, different people tend to consistently overestimate or underestimate the correct number. The number of overestimators and underestimators in the population seem to be about equal. While psychologists do not place any value judgment on whether it is better to be an overestimator or an underestimator, past studies have shown that whether one is an overestimator or an underestimator tends to reveal something fundamental about one's psychological characteristics and personality. As such, we will use this information to classify participants in this study into groups."* Participants were asked to estimate the number of dots on three paintings: "*How many dots do you think there were?"* Each painting was displayed on the screen for 3 seconds; after the painting disappeared, participants were asked to enter their answer in a blank. Right after they finished the task, the program randomly generated their testing results as an overestimator or an underestimator so that half of the participants were assigned to the team of "underestimators," whereas the other half were assigned to "overestimators."

They were told that "*for the rest of the study, you will be part of a group with several other [underestimators/ overestimators],*" and that their next mission was to observe interactions between two players who previously attended a related study and to respond to a series of questions about these interactions. The instruction read: "*you will observe the behavior of two people playing a series of games together. These individuals were prior participants in a related study, and we are interested in your impressions of their behaviors. Please try to discern the motives behind the players' behaviors, and attempt to be as accurate as you can because later on we will present you several questions regarding their behaviors in these games.*"

The introduction about the Prisoners' Dilemma game was subsequently presented to familiarize participants with the rules: "*We will first introduce the rules of this game. Two people play this game for points. On every round, each player chooses whether to COOPERATE with the other player or to DEFECT (i.e., not cooperate). Each player makes his or her decision without knowing the other player's choice. However, the outcome (points) that each player earns in the game depends on their choice in combination with what the other player chose.*" The possible outcomes of this game were shown in a matrix (see Appendix 1) on the screen, followed by instructions: "*If both of the players choose to cooperate, they*

both earn 40 points; if one player defects and the other cooperates, the player who defects earns 60 points, whereas the person who cooperates earns 0 (zero) points; if both of the players defect, they both earn 20 points. The best thing for both players as a collective is to cooperate with each other. However, the best thing for each individual is to defect while his/her partner cooperates. These incentives typically make cooperation difficult, though not impossible. "

After fully understanding the rules, participants saw the photographs of two players posed in the upper-right corner and the upper-left corner of the screen. Before proceeding to the pages that contained the players' decisions, they were told: "*Some prior participants were overestimators and some were underestimators. If their group membership is known, it will be displayed with their decisions/outcomes. (Note: some participants did not complete the dot estimation task - in these cases, the display will read 'unknown membership').*" The decisions of both players were presented in the center of the screen once per round; the result stayed on the screen till participants pressed the 'continue' button. We controlled that each player defected only once in either the 2nd or 3rd round and both of them cooperated after the 3rd round until the end (see Appendix 2). In the intergroup scenario, we randomized either an ingroup player or an outgroup player to defect

first in the 2nd round of the Prisoners' Dilemma game. Overall, participants observed high levels of cooperation from both players.

Our key manipulation *scenarios* in this study were five conditions constructed by varying the team memberships (underestimator vs. overestimator) of the two players (see an example in Appendix 3). In two control conditions (i.e., *Ingroup–Control* and *Outgroup–Control*), only one player's team membership was revealed whereas the other's was shown as "unknown membership." In this way, they served as a baseline in comparison with other experimental conditions. The experimental conditions included intragroup interactions either among ingroup targets (*Ingroup-Ingroup scenario*), or among outgroup targets (*Outgroup-Outgroup scenario*), and intergroup interactions (*Intergroup scenario*).

Dependent Variables

Our crucial dependent variable was **dispositional attribution** of each player (e.g., "*to what extent do you believe this player is trustworthy?"* see Appendix 4). In order to probe into participants' inferences about fundamental personalities of each player, they were encouraged to think about what each player was like "deep down in terms of their fundamental personality traits." Traits included friendly, trustworthy, cooperative and six other traits tapping into perceived cooperativeness of each player (α = .949 for Player 1; α = .951 for Player 2), along with three filler

traits egalitarian, competent, and intelligent. For this and following measures, the photograph of the relevant player was posed in the center of the screen. Complementary to the dispositional attribution measure, three critical items in the behavioral inference scale probed into the extent to which each player's **cooperativeness disposition was accountable for his behavior** (including "*this player took the other player's needs into account when making decisions"*; "*this player genuinely cared about the other player in these games*"; "*this player would have behaved in the same way no matter who he was playing with.*" see Appendix 5)[2]. Items in the behavioral inference scale were modified from Van Hiel, Vanneste, and De Cremer's (2008) attribution scale.

In addition to the items regarding dispositional attributions, the behavioral inference scale also consisted of situational causes accountable for the players' behaviors (see Appendix 5). Participants were asked to rate how possible each cause resulted in each player's behavior. In particular, three items were designed to assess the extent to which they thought each target's behavior was **attributable to their group membership** (including "*this player's decisions were influenced by the group membership of his partner*"; "*this player would have made different decisions*

[2] In order to construct the dimensions of behavioral causes, we will further conduct an exploratory factor analysis on the behavioral inference scale in the data analysis section.

if the group membership of his partner had been different"; "*the fact that this player shared or did not share a group membership with his partner is important for understanding his behavior in these games*"). Other situational causes (such as "*this player wanted to maximize his own payoff*"; "*this player wanted to avoid negative evaluations from the other player*") were listed in Appendix 5.

Participants further rated the **perceived entitativity** of each group (e.g., *"they were well coordinated"*; α =.910 for the overestimators; α =.904 for the underestimators; see Appendix 6), **perceived outgroup threat** (e.g., "*I think that the [underestimator] group is a strong rival to my group*"; α =.841 for rated threat from the overestimators; α =.793 for rated threat from the underestimators; see Appendix 7), **perceived group norms** (e.g., "*to what extent do you think watching the players provided me with useful information about how people in general should play these games?*"; α = .777 for the overestimators; α = .791 for the underestimators; see Appendix 8). At about this point, participants completed an ingroup identification scale (e.g, "*To what extent do you feel a bond with [overestimators/ underestimators]*"; α = .955 for the overestimators; α = .957 for the underestimators; *see Appendix 10).*

Participants also indicated their **cooperative expectations** for both players (e.g., "*Please imagine if you were the other player in the game with your picture and*

group membership revealed to this player…to what extent do you believe this player would cooperate with you? ", "*to what extent do you believe this player would be motivated to maximize your outcome?*"; α = .902 for Player 1; α =.906 for Player 2; see Appendix 9). Finally, they subsequently made their own **cooperative decisions** with regard to new players from each team. Participants were told that they would play a different form of economic games with other participants recruited online. The rule was that participants were first given 500 points in each round and they could decide any amount to send to their partners in this round and saved the rest for themselves; the amount sent would be multiplied by 3 and then given to their partners, and their partners could decide any amount to send back to them. Participants then played 8 rounds of trust games, 4 rounds with ingroup players and 4 rounds with outgroup players; these players were novel and never participated in the earlier phase. In each round, a player's photograph was posed in the center of the screen. After the trust games, participants then filled out questions about their political orientation and demographic information. Upon accomplishment of the experiment, participants were thanked and fully debriefed.

Predictions

Prediction 1. We predicted that cooperation by an ingroup player would lead to more positive dispositional attributions than cooperation by an outgroup player.

By comparing the effects of observing the scenarios involving only ingroup or outgroup players (i.e., the homogeneous and control conditions) on our dependent variables, we examined if an ultimate attribution bias would contribute to lowered cooperative expectations and cooperative decisions toward outgroup partners (vs. ingroup partners) in the trust games. That is, we expected that an ingroup player would be perceived as more dispositionally cooperative than an outgroup player, whereas observed cooperation by an outgroup player would be perceived as externally/situationally induced (i.e., due to a shared group membership), thus leading to subsequently biased cooperative decisions.

Prediction 2. Observing cooperative intragroup interactions among outgroup players would reduce cooperative expectations and lead to biased decisions toward outgroup partners, compared to observing an outgroup actor cooperating with an unknown-identity player. The underlying mechanism was hypothesized to differ depending on individuals' identification levels: weakly identified individuals were expected to attribute outgroup players' cooperation with each other to causes other than dispositions (e.g., a shared group membership), whereas highly identified individuals were expected to perceive more threat in the observations of outgroup cooperation (vs. observations of an outgroup actor interacting with an unknown-identity player). By comparing the effects of *Outgroup-Control* and *Outgroup-*

Outgroup on our key dependent variables, we first tested whether observing intragroup cooperation among outgroup members did lower cooperative expectations and induce biased decisions toward outgroup partners in the trust games. The predicted mediation routes for highly and weakly identified individuals were also examined. For weakly identified individuals, the effects of scenario on cooperative expectations may be mediated by reduced dispositional or heightened situational inferences. In contrast, for highly identified individuals, the relationship between observing outgroup-outgroup cooperation and cooperative expectations may be mediated by heightened perceptions of cohesiveness, leading to greater perceptions of intergroup threat (see Figure 1).

Prediction 3. Observing intergroup interactions was expected to boost cooperative expectations, and encourage less biased decisions toward outgroup partners in the trust games, compared to observing intragroup interactions among ingroup players or among outgroup players. The underlying mechanism was hypothesized to differ depending on individuals' identification levels: observing cooperative intergroup interactions, weakly identified individuals may make more dispositional (and less situational) attributions of outgroup players (i.e., perceiving them as more dispositionally cooperative), whereas highly identified individuals

may follow a salient group norm which highlighted the normativity of intergroup cooperation (see Figure 2).

Data Analysis

Sixteen participants were removed from the analysis for failing to correctly report their assigned team, as measured in a manipulation check. Twelve participants with completion times lower than half of the median of all the participants (Median = 14 minutes) were also excluded. Applying both criteria resulted in a sample size of 328 (Ingroup-Ingroup = 64; Outgroup-Outgroup = 70; Intergroup = 66; Ingroup-Ctrl = 60; Outgroup-Ctrl = 68).

The behavioral inference scale contained 12 items. An exploratory factor analysis was conducted with oblique rotation ("Promax" method in SPSS) to extract the factors from attributions about behaviors of Player 1 and Player 2 respectively. By using oblique rotation, we assumed the factors were correlated. The criteria to determine the factors were based on the scree plot and interpretability of the classification consistent with attribution theories. The results based on the behaviors of Player 1 and Player 2 were similar[3] (see the loadings on each factor in Table 1): four factors emerged and explained around 66% of the variance. However, the

[3] There were two inconsistent results with respect to the behaviors of Player 1 and Player 2 in the factor analysis. For the convenience of further analysis, we only kept and interpreted the factor with a higher loading for both of the items.

result was not entirely consistent with our classification of items as mentioned in the method section. The first factor included items 1, 2, 3 and 6, referring to attributions related to a shared group membership. The second factor was composed of items 4, 5, and 8, indicating genuine concern for others and mutual reciprocity. The third factor included items 10, 11, and 12, linked to risk-prevention causes. The forth factor consisted of items 7 and 9, suggesting causes related to strategic cooperation – i.e., players would only cooperate if the situation was in their favor. Scores within the same factors were averaged, thus representing to what degree the behavior of each player was ascribed to four types of causes.

However, due to the high correlations (*r*s = .549 ~ .647) among the key dependent variables (i.e., dispositional attributions, the concern-for-other cause, and cooperative expectations), a single standardized score was created for each participant indicating dispositional attributions about each player. The steps were as follows. We first standardized the scores on the 9 items of the dispositional attribution measure, 3 items (i.e., items 4, 5 and 8) of the concern-for-other cause, and 4 items of the cooperative expectation scale, and we then averaged these standardized scores together[4]. The higher this averaged score was, the more they

[4] This averaged score consists of all the responses we measured about dispositional attributions, and weights them equally. From our perspective, the computing method is relatively open to debate. An alternate computing method would be to first average the scores of each subscale, standardize these

attributed a player's cooperation to his disposition. We used this averaged score instead of the scores on individual subscales for the following analyses.

Results

Ultimate Attribution Error

The first hypothesis in Study 1 was that cooperation by ingroup players was more likely to be attributed to dispositional cooperativeness than cooperation by outgroup players; further, cooperation by outgroup players was more likely to lead to situational attributions (i.e., a shared group membership) in comparison. To test this (Prediction 1), in this section we first analyzed whether behaviors of ingroup players engaging in *intragroup cooperation* or *cooperation with unknown-identity players* would receive greater dispositional attributions than behaviors of outgroup players. Further, we examined whether behaviors of ingroup players involved in *intergroup cooperation* would also be attributed differently from behaviors of outgroup players.

---

means and take their average. In this way, the new score would consist of three components from each subscale with equal weights. We further examined the results based on the both methods, but it did not yield any difference in our important findings, suggesting that our manipulation seemed to cause a stable change and fluctuation among participants' responses on disposition-related items regardless of the composition of the computing score.

For the first analysis, the dependent variables: (1) dispositional attributions; (2) situational attributions to shared group membership[5] were examined between the conditions with only ingroup players involved (i.e., Ingroup-Ingroup and Ingroup-Ctrl scenarios) and the conditions with only outgroup players involved (i.e., Outgroup-Outgroup and Outgroup-Ctrl scenarios). Responses to the two players in the Ingroup-Ingroup and Outgroup-Outgroup scenarios were averaged for participants in the homogeneous conditions, whereas for the control conditions, scores were based only on the player with his team membership revealed (ingroup or outgroup). Dispositional attributions and shared-group-membership attributions were analyzed with 2 (scenario type: Homogeneous vs. Control) x 2 (player membership: Ingroup vs. Outgroup) between-subject factorial ANOVAs. If participants displayed an ultimate attribution error, ingroup players should be rated as more dispositionally cooperative than outgroup players. However, the main effect of player membership was not significant: participants did not perceive ingroup players (M = .081) as significantly more cooperative than outgroup players (M = -.065) across the two scenario types, $F(1, 258) = 2.627$, p = .106. None of other

[5] All the analyses were also conducted on other dependent variables such as risk-prevention cause, strategic-cooperation cause, and group perception (i.e., perceived group norms, perceived entitativity, and outgroup threat). However, since we did not find evidence that these behavioral causes functioned as external attributions for cooperation, nor did these findings contribute to interpretation of the attribution model, we report the findings in Appendix 13.

effects was significant, Fs < 1. We also anticipated that cooperation by outgroup players should be attributed to a shared group membership more than cooperation by ingroup players. However, the main effect of player membership on shared-group-membership attributions was not significant, and neither was the scenario type x player membership interaction, Fs < 1, indicating that participants did not make situational attributions differently in the face of ingroup or outgroup cooperation. We found only a significant main effect of scenario type on shared-group-membership attributions, $F(1, 258) = 7.073$, $p = .008$. Within-group cooperation led to greater attributions related to a shared group membership ($M = 4.013$) than cooperation with an unknown-identity player ($M = 3.588$). To this point, our findings suggest that observing highly cooperative behaviors by ingroup vs. outgroup players did not cause a difference in dispositional or situational attributions in favor of ingroup members.

In a second analysis to examine the ultimate attribution error, we tested whether ingroup players were perceived differently from outgroup players in the intergroup scenario. If participants displayed an ultimate attribution error, they would judge ingroup players in a cross-group interaction as more cooperative than outgroup players. We first examined dispositional attributions and shared-group-membership attributions for participants being assigned to observe the intergroup

scenario with 2 (player membership: Ingroup vs. Outgroup) within-subject factorial ANOVAs. Similar to our results for the homogeneous conditions, the main effect of player membership was only marginally significant: ingroup players (M = .067) in the intergroup interaction seemed to be perceived as only slightly more cooperative than outgroup players (M = -.105), $F(1, 65) = 3.635$, p = .06. As for attributions to a shared group membership, the main effect of player membership did not reach a significant level, $F(1, 65) < 1$. Taken together, in the scenario that members cooperated across group boundaries, ingroup players were not perceived as more dispositionally cooperative than outgroup players, neither were outgroup players perceived as more situationally cooperative than ingroup players.

Consequences of Outgroup Cooperation

In Prediction 2, we hypothesized that outgroup players might be perceived as less cooperative when they engaged in intragroup cooperation than when they engaged in cooperation with a player with an unknown identity. Furthermore, we expected that this effect may be different as a function of group identification. Participants' identification scores were first standardized. Hierarchical multiple regression analyses were conducted with scenario type (effect coded: 1 = Outgroup-Outgroup; -1 = Outgroup-Control), identification and the interaction term predicting dispositional attributions and attributions to a shared group membership. We

anticipated a significant scenario type × identification interaction. The interaction item significantly predicted attributions to the shared-group-membership cause (β = .275, p = .001) but not dispositional attributions (β = .052, p = .53). Specifically, strong identifiers perceived the behaviors of observed players in the Outgroup-Outgroup scenario as more likely to be driven by a common group identity than behaviors of those in the Outgroup-Ctrl scenario (β = .435, p < .001), whereas weak identifiers observing these two scenarios made a similar degree of shared-group-membership attributions (β = -.113, p = .284)[6]. Taken together, observations of outgroup cooperation indeed increased attributions to a shared group membership, especially among highly identified individuals; however, our second hypothesis was

[6] It is worth noting that although outgroup cooperation did not elicit more perceived outgroup entitativity or outgroup threat compared to the control condition (β = .092, p = .211; β = -.019, p = .797), we found that overall strong vs. weak identifiers perceived the outgroup differently. Hierarchical regression analyses were implemented with identification, perceived outgroup entitativity and the interaction term predicting perceived outgroup threat across these two scenarios. We found that perceived outgroup threat was positively correlated with group identification (β = .561, p < .001) after controlling for outgroup entitativity, and the degree of threat elicited by outgroup entitativity depended on group identification (β = -.112, p = .01): for weak identifiers, the more cohesive the outgroup was perceived, the higher the level of threat it triggered (β = .158, p = .009), whereas this correlation was absent for strong identifiers (β = -.027, p = .67): outgroup threat was readily perceived by strong identifiers regardless of outgroup cohesiveness (strong identifiers–whose group identification were 1 S.D. higher than averaged–rated outgroup threat as 4.939 out of 7 when their perceived outgroup entitativity was low, and rated outgroup threat as 4.617 when their perceived outgroup entitativity was high; weak identifiers rated outgroup threat as 1.414 with low perceived outgroup entitativity and as 2.088 with high perceived outgroup entitativity).

not supported that exposures to outgroup cooperation would reduce perceptions of dispositional cooperativeness.

Intergroup Interactions vs. Intragroup Interactions

We then examined whether observations of intergroup cooperation would give rise to different dispositional or situational inferences compared to observations of intragroup cooperation. According to Prediction 3, intergroup cooperation should reduce attributions to a shared group membership, and thus boost the perceived cooperativeness of outgroup players, compared to intragroup cooperation among outgroup players. We expected that weakly identified individuals would be the most likely to follow this pattern that cooperative behaviors of outgroup players in intergroup interactions would elicit more dispositional attributions than those in intragroup interactions. In contrast, intergroup interactions per se would be considered as group norms by highly identified individuals and would directly encourage them to cooperate across group lines. As such, we first examined dispositional attributions and shared-group-membership attributions about outgroup players with scenario (effect coded: 1 = Intergroup scenario; -1 = Outgroup-Outgroup scenario), identification and the interaction term. We anticipated a significant interaction effect, indicating the effect of observing intergroup cooperation varied as individuals' level of group identification changed. However,

the interaction effect was not significant in predicting either dispositional attributions (β = -.083, p = .307) or shared-group-membership attributions about outgroup players (β = -.053, p = .500). We only found that individuals' level of group identification positively predicted both dispositional attributions (β = .358, p < .001) and shared-group-membership attributions about outgroup players (β = .439, p < .001). Although observations of intergroup cooperation itself did not influence attributions about outgroup players, the more individuals identified with their group, the more they tend to attribute cooperative behaviors of an outgroup to both their disposition and group identity.

Inferences about ingroup players were also examined. Hierarchical multiple regression analyses were conducted with scenario (effect coded: 1 = Intergroup scenario; -1 = Ingroup-Ingroup scenario), identification and the interaction term predicting dispositional or behavioral inferences about ingroup players. Neither the main effect of scenario nor interaction effect was significant (ts ≤ 1), indicating that observations of intergroup cooperation itself did not alter attributions about ingroup players' behaviors relative to observations of ingroup cooperation, and this relationship did not vary as a function of group identification. However, individuals' level of group identification positively predicted both dispositional attributions (β = .359, p < .001) and shared-group-membership attributions about ingroup players (β

= .300, p = .001). In conclusion, composition of team membership of the players did not influence cooperativeness inferences about either ingroup or outgroup players. However, unexpectedly, regardless of the observed scenarios, strong identifiers made more dispositional and shared-group-membership attributions about cooperative behaviors of members from both groups, compared to weak identifiers. One limitation in Study 1 is that we focused on participants' inferences about the observed players, and thus their responses to potential players left unexamined. For instance, we did not measure attributions about outgroup members among participants observing the Ingroup-Ingroup scenario since they did not observe any outgroup target. This design was modified in Study 2 so that we could directly compare participants' inferences about both ingroup and outgroup members across all the scenarios.

Ingroup Bias – Allocations in Trust Games

Mere team assignment should lead to ingroup bias - i.e., a preference to affiliate with ingroup partners over outgroup partners (Brewer & Kramer, 1986; Tajfel & Turner, 1979). This hypothesis was examined by testing whether the allocations in the trust games would vary with partner membership: Participants demonstrated ingroup bias if they sent more points to ingroup partners than to outgroup partners. Allocations in the trust games were examined with a partner

membership (within-subject: ingroup vs. outgroup) x scenario (between-subject: Ingroup-Ingroup vs. Outgroup-Outgroup vs. Intergroup vs. Ingroup-Ctrl vs. Outgroup-Ctrl) factorial ANOVA. The main effect of partner membership was significant, $F(1, 322) = 18.198$, $p < .001$: the average amount allocated to ingroup partners ($M = 230.883$) was more than the average amount allocated to outgroup partners ($M = 209.067$). However, neither the interaction of scenario and partner membership, $F(4, 322) < 1$, nor the main effect of scenario reached a significant level, $F(4, 322) = 1.907$, $p = .109$. Statistically, only participants observing Intergroup or Ingroup-Ingroup scenarios showed ingroup bias: they allocated around 30.402 points and 34.313 points more to ingroup partners than to outgroup partners respectively, $t(65) = 2.675$, $p = .008$; $t(63) = 2.973$, $p = .003$[7], whereas participants observing the other scenarios did not show this bias, ts < 1.617, ps > .107 (see Figure 4). As previous studies have repeatedly replicated, team assignment elicited preferential trust in ingroup partners over outgroup partners, but the extent of ingroup preference did not differ with the scenarios participants observed.

[7] When the most stringent correction method—Bonferroni correction—was used to avoid the likelihood of a Type I error for multiple comparisons, the results were still significant (the significant level was alpha divided by comparison number, which yielded a new critical value equal to .01).

Defection Order

During data analysis, we speculated that defection order might matter for inferences about players' cooperativeness even in the homogeneous conditions: Behaviors of the players who defected first may be interpreted differently from behaviors of the other players who defected second, because defections by the second defectors could be perceived as a reasonable response to the first defectors. Specifically, we anticipated that first defectors would be rated as less cooperative than second defectors, and that the extent to which perceived dispositional cooperativeness was discounted by first defection might depend on players' group membership. Dispositional attributions were analyzed with a *player membership* (between-subject: Ingroup vs. Outgroup) × *defection order* (within-subject: First vs. Second) mixed-model ANOVA. If first defectors were perceived differently depending on their membership, a player membership × defection order interaction should be observed. However, the interaction effect was not significant, $F(1, 132) = 2.352$, $p = .146$[8].

[8] A simple main effect of player membership for first defectors almost reached the significant level: Outgroup first defectors ($M = -.135$) were perceived as slightly less cooperative than ingroup first defectors ($M = .160$), $F(1, 132) = 4.605$, $p = .034$ (the criterion for significance was .025 because two simple main effect tests were conducted), whereas second defectors of both group memberships were perceived as equally cooperative (Ms = .036 for ingroup vs. -.051 for outgroup), $F(1, 132) < 1$.

We also examined whether similar biases in inferences occurred in the intergroup scenario. If participants observing the intergroup scenario displayed an ultimate attribution error, outgroup first defectors would be perceived as less cooperative than ingroup first defectors; in particular, outgroup players who defected against an ingroup player first might be perceived as especially uncooperative. We conducted a repeated measures ANOVA with *membership of first defectors* (between-subject: Ingroup vs. Outgroup) × *defection order* (within-subject: First vs. Second) predicting dispositional attributions about the two players in the intergroup scenario. We anticipated membership of first defectors × defection order interactions on attributions, as well as simple main effects of defection order at different levels of first defector membership. However, interaction effect was only marginally significant, $F(1, 64) = 3.589$, $p = .063$. The simple main effect that examining perceived cooperativeness of second defectors by their group membership were not significant, $F(1, 64) = 2.849$, $p = .096$, suggesting that ingroup players who were defected against by outgroup (M = .126) looked like slightly more cooperative than outgroup players who were defected against by ingroup (M = -.178), whereas ingroup first defectors did not look differently from outgroup first defectors, $F(1, 64) < 1$. None of other effects was significant, Fs < 1. Overall, in Study 1, we did not obtain evidence regarding disparate dispositional attributions

about ingroup vs. outgroup players; furthermore, we ruled out the possibility that defection order would affect cooperativeness inferences about ingroup vs. outgroup members differently.

STUDY 2

In study 1, the evidence regarding how observing scenarios with different membership compositions affected cooperativeness inferences was not as strong as we anticipated. One reason may be that the impressions about ingroup vs. outgroup were based on only two target persons, and that although there were nine cooperative acts, observation of only two specific members may not suffice to represent a virtuous group. Engaging in this sort of impression formation process, participants may selectively pick up information that differentiated the two target persons rather than high levels of cooperation of both groups. In order to resolve this issue, we made four changes for the procedure of study 2. First, the instructions about the personality test (team assignment task) were modified so that participants were led to believe the credibility of this test was still in question. In this way, the group boundary created by arbitrary assignment might be less definitive compared to Study 1, and thus participants may be less likely to look for cues that differentiate the two groups. Second, a sanctioning system manipulation was added to test our hypotheses (outlined in the introduction above) concerning how attribution

processes may be altered by the presence of sanctions. Sanctions may provide a more salient explanation for the players' behaviors than other causes, thus allowing us to test whether attributions to dispositional causes would be reduced in the presence (vs. absence) of sanctions. Third, we removed the control scenarios (i.e., Ingroup-Ctrl and Outgroup-Ctrl scenarios) for the reason of simplicity. Fourth, participants watched ten pairs of players cooperating with one another (once for each pair) instead of a single pair of players. As such, participants observed only one out of ten players defecting against their partner, while the rest always cooperated. Finally, instead of rating the behaviors of target persons, participants in study 2 were asked to rate their inferences about both the ingroup and outgroup generally, no matter which scenario they were assigned to observe. The last two changes could make the design a more sensitive test of our theoretical framework by drawing participants' attention to group-based behaviors of the players.

Method

Participants

Four hundred and sixty participants were recruited online through Mechanical Turk for a monetary reward of $1. The design was a 3 (interaction scenario: Ingroup–Ingroup vs. Outgroup–Outgroup vs. Intergroup) × 2 (sanctioning system: Present vs. Absent) between-subjects factorial design.

Procedure

The procedures of study 2 remained as consistent as possible with Study 1 except for the aforementioned changes. Specifically, before conducting the dot estimation task, participants were told that "*more evidence is needed to conclude that whether one is an overestimator or an underestimator reveals something fundamental about one's psychological characteristics and personality*" and "*to investigate this issue, we will use this information to classify participants in this study into groups*." Since the purpose of study 2 was to test how sanctioning systems would influence attributions and cooperative expectations and decision making by providing a possible external attribution for observed cooperation, participants who were assigned to the sanction-present condition were introduced to the existence of a punisher right after they learned the rules of the Prisoners' Dilemma. The instruction read: "*the two players were also told that a third player, Player C, was watching what happened. Player C did not know anything about Player A or B (e.g., didn't see their photos). However, if any of them makes an unfair decision, Player C has been instructed to remove most of the player's points.*" Except for this manipulation, the other procedures remained the same for participants assigned to the sanction-present vs. -absent conditions. Subsequently, they observed a total of 20 players assigned to the roles of Player 1 or Player 2 (i.e., 10 pairs) working with each other

in the Prisoners' Dilemma games with their memberships, decisions and outcomes revealed on the screen. In each scenario, Player 2 always defected on the 2nd round, whereas Player 1 defected on the 3rd round. Importantly, half of the participants in the intergroup scenario observed ingroup players defecting against outgroup players first, whereas the other half observed outgroup players defecting against ingroup players first. After the observation phase, they were asked to rate their inferences about the behaviors of each team as in Study 1, namely dispositional attributions (α = .951 for the overestimators; α = .953 for the underestimators), behavioral inferences (the reliability of subscales will report in the result section), perceived entitativity (α = .906 for the overestimators; α = .900 for the underestimators), perceived outgroup threat (α = .799 for rated threat from the overestimators; α = .745 for rated threat from the underestimators), perceived group norms (α = .855 for the overestimators; α =.826 for the underestimators), and cooperative expectations (α =. 868 for the overestimators; α =.865 for the underestimators) (see Appendices 4-10).

Predictions

Prediction 1. We predicted that when the sanctions were absent as in Study 1, cooperation by an ingroup player would lead to more positive dispositional inferences than cooperation by an outgroup player (based on the ultimate attribution

error). By comparing the effects of observing intragroup cooperation among ingroup players vs. among outgroup players on our dependent variables, we expected that an ultimate attribution bias would contribute to lowered cooperative expectations and cooperative decisions toward outgroup partners (vs. ingroup partners) in the trust games. That is, despite exhibiting equally cooperative behaviors, pairs of ingroup players would be perceived as more dispositionally cooperative than outgroup players, whereas observed cooperation between outgroup players would be perceived as externally induced (i.e., due to a shared group membership), thus leading to subsequently biased cooperative decisions.

However, for the conditions in the presence of sanctions, observing within-group ingroup cooperation might yield less cooperative attributions/expectations for ingroup players compared to when the sanctions were absent. We predicted that the presence of sanctions would reduce dispositional attributions for ingroup cooperation, and thus lead to less cooperative expectations and less biased decisions, compared to when the sanctions were absent.

Prediction 2. As Study 1, based on outgroup entitativity literature, we expected that observing cooperative intragroup interactions among outgroup players would reduce cooperative attributions/expectations and lead to biased decisions toward

outgroup players. Although this hypothesis was not supported in Study 1 when comparing outgroup pairs to the control condition (i.e., interactions between outgroup players and unknown-identity players), Study 2 allowed us to re-test it by comparing the effects of observing the *Outgroup-Outgroup* scenario between the conditions with vs. without sanctions. We expected that observations of within-group outgroup cooperation may yield more cooperative attributions/expectations of outgroup players when the sanctions were present (vs. absent). Specifically, the presence of sanctions might reduce attributions of intragroup cooperation among outgroup members to their shared group membership (by providing an alternate external explanation). To the extent that this would reduce perceived outgroup entitativity, bias in expectations and decisions may also be reduced.

Prediction 3. Observing intergroup interactions was expected to boost cooperative expectations about outgroup partners, and encourage less biased decisions in the trust games, compared to observing intragroup interactions among ingroup players or among outgroup players. The underlying mechanism was hypothesized to differ depending on individuals' levels of group identification: weakly identified individuals may make more dispositional attributions of outgroup players (i.e., perceiving them as more dispositionally cooperative), whereas highly identified individuals may follow a salient group norm which highlighted the

normativity of intergroup cooperation (vs. observing intragroup interactions among ingroup players). Further, we hypothesized that observations of intergroup cooperation would yield less cooperative attributions/expectations for outgroup players when sanctions were present (vs. absent). We predicted the presence of sanctions would reduce dispositional attributions of each player in intergroup cooperation, thus leading to less cooperative expectations. Based on the ultimate attribution error, we expected that this attenuation in dispositional inferences would likely be greater for outgroup than ingroup players, leading to more biased decisions toward outgroup members, compared to when the sanctions were absent.

Results

Eight participants were removed from the analysis because they failed to recognize their assigned team, as measured by the manipulation check. Twenty-one participants with completion time lower than half of the median of all the participants (Median = 14 minutes) were also excluded. Both criteria resulted in a sample size of 431 (Ingroup without sanctions = 82; Ingroup with sanctions = 57; Outgroup without sanctions = 76; Outgroup with sanctions = 74; Intergroup without sanctions = 60; Intergroup with sanctions = 82). As Study 1, scores on items related to dispositional attributions were standardized and averaged in order to create an overall score that represented the degree of dispositional attributions.

Ultimate Attribution Errors

In examining the first prediction, we tested whether cooperation by ingroup players was more likely to be attributed to dispositional cooperativeness than cooperation by outgroup players; further, cooperation by outgroup players was more likely to lead to situational attributions (i.e., a shared group membership) in comparison. Dispositional attributions and attributions to a shared group membership[9] were examined with *scenario* (between-subject: Ingroup-Ingroup vs. Outgroup-Outgroup vs. Intergroup) × *sanction* (between-subject: Present vs. Absent) × *target group membership* (within-subject: Ingroup vs. Outgroup) mixed-model ANOVAs. The main effect of membership on dispositional attributions was significant, $F(1, 425) = 49.622$, $p < .001$. The ultimate attribution error was confirmed in that ingroup players ($M = .135$) were rated as more dispositionally cooperative than outgroup players ($M = -.135$). However, this effect was qualified

[9] An exploratory factor analysis was conducted to extract the factors from attributions about the behaviors of the two target groups. The results showed that while eight items were stably grouped into three factors, namely shared group membership, concern for others, and risk prevention, the other four items were divergently loading on different factors depending on the target groups. Since the nature of the groups were not the focus of the present study, and also the classification itself contained the variability contributed by both ingroup and outgroup, one of our key manipulations, we decided to adopt the four-dimensional classification of behavioral inferences from Study 1. It ended up that except for the strategic-cooperation subscale ($\alpha = .404$ for the overestimators; $\alpha = .283$ for the underestimators), the reliability of other subscales was at level that ranged from acceptable to high ($\alpha = .560$ - $.758$). For the results regarding risk-prevention attributions and strategic-cooperation attributions, please see Appendix 13.

by an interaction of scenario and membership, $F(2, 425) = 70.613$, $p < .001$. Participants observing the ingroup and intergroup scenarios rated ingroup players as more cooperative (Ms = .329 and .216) than outgroup players (Ms = -.536 and -.086), $F(1, 425) = 144.732$, $p < .001$; $F(1, 425) = 18.256$, $p < .001$. Participants observing the outgroup scenario, on the contrary, rated outgroup players as more cooperative ($M = .188$) than ingroup players ($M = -.121$), $F(1, 425) = 20.726$, $p < .001$ (see Figure 5). The main effect of sanctions and its interaction effects were nonsignificant, Fs < 2.441, ps > .119 (See cell means in Table 2).

To answer our key hypothesis – whether observations of different scenarios modified the ultimate attribution errors, we examined the magnitude of ultimate attribution errors (difference in dispositional attributions about ingroup vs. outgroup players) across scenarios and sanction conditions. We found that participants observing the ingroup scenario displayed a stronger ultimate attribution error than participants in the other two conditions, ts(425) ≥ 5.497, ps < .001, while participants observing the intergroup scenario had greater bias than those observing the outgroup scenario, $t(425) > 6.352$, $p < .001$[10]. In conclusion, ingroup members were considered as more cooperative than outgroup members if participants were exposed to information indicative of ingroup members' cooperativeness. In

[10] Tukey post hoc tests were used.

situations when this information was not available, members from a highly cooperative outgroup appeared more cooperative than ingroup members. Furthermore, sanctions did not influence cooperativeness attribution or inference processes.

In addition to dispositional attributions, we hypothesized that observing intergroup cooperation might benefit cross-group affiliation by reducing attributions such as a shared group membership that can otherwise account for within-group cooperation among outgroup members. We conducted *scenario* × *sanction* × *membership* mixed-model ANOVAs on perceived accountability of a shared group membership. A main effect of membership and a *membership* × *sanction* interaction on shared-group-membership attributions were significant, $F(1, 425) = 4.610$, $p = .032$; $F(2, 425) = 5.00$, $p = .026$. However, contrary to our predictions, a shared group membership was perceived as more accountable for behaviors of ingroup players ($M = 3.688$) than behaviors of outgroup players ($M = 3.460$) when the sanctions were absent, $F(1, 425) = 9.75$, $p = .002$, while it was regarded as equally accountable for both groups when the sanctions were present (Ms = 3.607 for ingroup vs. 3.611 for outgroup), $F(1, 425) < 1$.

Moderation by Group Identification

We hypothesized that exposure to different scenarios might influence individuals with high vs. low group identification in different ways. According to our intergroup attribution model (see Figure 2), we proposed that cooperativeness inference processes may occur via disparate routes depending on individuals' levels of group identification. Among weak identifiers, observations of intergroup cooperation (vs. intragroup cooperation) may yield more dispositional attributions, in turn leading to higher cooperative expectations about outgroup members. In contrast, exposure to intergroup cooperation may heighten perceived group norms among strong identifiers. However, we found that the perception of group norms did not seem to play a role in dispositional attributions about outgroup members[11].

Based on the evidence we have presented so far – exposure to different scenarios affected cooperativeness attributions about outgroup members, we instead examined whether group identification moderated dispositional attributions about ingroup or outgroup members. Scenarios were coded as in the analyses above (Dummy 1: 1 = intergroup cooperation, 0 = ingroup cooperation; Dummy 2: 1 =

[11] Hierarchical linear regression analyses on dispositional attributions about outgroup members and decisions toward outgroup partners in the trust games were implemented with perceived group norms, group identification and the interaction term as predictors. Neither group norms nor the interaction of group norms × group identification significantly predicted the outcome variables (ts < 1.860, ps > .07).

outgroup cooperation, 0 = ingroup cooperation), and thus positive (negative) regression coefficients indicated that the alternative scenario increased (reduced) inferences about dispositional cooperativeness relative to the ingroup scenario. Hierarchical linear regression analyses were conducted with dummy variables, identification and their interaction terms predicting dispositional attributions about ingroup vs. outgroup members respectively. In step 1, dummy variables and standardized identification scores were entered into equation. In step 2, their interaction terms were further entered. We were interested in whether the interaction term significantly predicted dispositional attributions about ingroup members. The effect of observing the outgroup scenario relative to the ingroup scenario on dispositional attributions about ingroup players depended on individuals' level of group identification (Dummy 2 × ID: $\beta = .125$, $p = .033$). In contrast, the effect of observing the intergroup scenario on dispositional attributions did not vary with group identification (Dummy 1 × ID: $\beta = .024$, $p = .402$) (see Table 3). Specifically, observing the outgroup scenario (relative to the ingroup scenario) reduced dispositional attributions about ingroup members among weak identifiers (coded as -1; $\beta = -.269$, $p < .001$) but not among strong identifiers (coded as 1; $\beta = -.063$, $p = .340$). This finding indicated that while weak identifiers perceived the ingroup members as less cooperative in situations where positive

information of the ingroup was not available, strong identifiers perceived the ingroup as equally positive regardless of the number of positive instances of ingroup members.

As for the influence of scenarios on perceived cooperativeness of outgroup members, the magnitude of dispositional attributions inferred from outgroup behaviors in different scenarios also varied with individuals' levels of group identification (Dummy 1 × ID: $\beta = .118$, $p = .004$; Dummy 2 × ID: $\beta = .239$, $p < .001$). For strong identifiers (with identification score 1 SD higher than the average), observing either the intergroup or the outgroup scenario increased perceived dispositional cooperativeness (Dummy 1: $\beta = .398$, $p < .001$; Dummy 2: $\beta = .662$, $p < .001$). For weak identifiers (with identification score 1 SD lower than the average), only observing the outgroup scenario but not the intergroup scenario improved perceived cooperativeness of outgroup members relative to observing the ingroup scenario (Dummy 1: $\beta = .098$, $p = .218$; Dummy 2: $\beta = .268$, $p = .001$) (see Figure 6). These findings showed that group identification determined the extent of dispositional cooperativeness inferred from outgroup behaviors[12]. Opposed to our

[12] Similar to Study 1, group identification significantly predicted perceived outgroup threat ($\beta = .431$, $p < .001$) controlling for outgroup entitativity, and the degree of threat elicited by outgroup entitativity depended on group identification ($\beta = -.157$, $p < .001$). Perception of outgroup threat by weak identifiers increased with their perception of outgroup entitativity ($\beta = .206$, $p < .001$), whereas

predictions, strong identifiers made more dispositional cooperativeness inferences about outgroup members as cooperative behaviors of outgroup exemplars increased (i.e., in both the intergroup and outgroup conditions), whereas this pattern was less pronounced for weak identifiers. However, observing intragroup cooperation among outgroup players (i.e., exposure to overwhelmingly cooperative outgroup exemplars and lack of positive ingroup exemplars) did boost perceived cooperativeness about outgroup members for individuals both highly or weakly identified with their own group.

Ingroup Bias – Allocations in Trust Games

As Study 1, ingroup bias refers to preferential allocations offered to ingroup members over outgroup members in the trust games. The questions of interest were whether ingroup bias in allocations would be exacerbated or attenuated by observations of different scenarios or by the presence of sanctions in the scenarios. Allocations in the trust games were examined with a *scenario* (between-subject: Ingroup-Ingroup vs. Outgroup-Outgroup vs. Intergroup) × *sanction* (between-subject: Present vs. Absent) × *partner membership* (within-subject: Ingroup vs.

---

strong identifiers perceived a similar degree of outgroup threat regardless of perceived outgroup entitativity (β = -.054, p = .359). Weak identifiers rated outgroup threat as 2.721 out of 7 when perceived entitativity was high and as 1.800 when it was low; strong identifiers rated outgroup threat as 4.972 and 4.466 respectively.

Outgroup) mixed-model ANOVA. Ingroup bias in allocations was significant, $F(1, 425) = 27.754$, $p < .001$; allocations to ingroup partners ($M = 245.47$ points) were greater than allocations to outgroup partners ($M = 219.78$ points). This effect was qualified by an interaction of scenario and partner membership $F(2, 425) = 19.343$, $p < .001$. If we further compared the magnitude of ingroup bias across scenarios, only participants observing the ingroup and intergroup scenarios demonstrated significant ingroup bias, $F(1, 425) = 10.342$, $p = .001$; $F(1, 425) = 51.108$, $p < .001$: they allocated 61.772 points and 27.375 points more to ingroup partners (Ms = 266.770 and 250.546) than to outgroup partners (Ms = 204.998 and 223.171) respectively. In contrast, participants observing the outgroup scenario seemed to prefer outgroup partners ($M = 231.156$) over ingroup partners ($M = 219.091$), but this trend was not significant, $F(1, 425) = 2.174$, $p = .141$ (see Figure 7).

General Discussion

The present studies investigated the role of dispositional attributions in decisions regarding cross-group cooperation, especially when individuals witness members of different groups demonstrating prosocial behaviors. In Study 1, we failed to find evidence supporting any of our predictions. The defection order of two observed players was one possibility that caused the nonsignificant results, but it was further ruled out. In Study 2, we modified the procedure and measures in order to more directly examine intergroup processes. Participants observed 10 cooperative dyads composed of ingroup or outgroup members (instead of a pair of individual targets) in the absence or presence of sanctions, and they were asked to infer cooperative dispositions and expectations about the ingroup vs. outgroup as a whole (rather than specific observed targets). We found that observing cooperative behaviors by outgroup players (in either intragroup or intergroup interactions) increased cooperativeness inferences about the outgroup, especially among individuals highly identified with the group.

These findings were inconsistent with two main predictions: First, we predicted that cohesive outgroup interactions would induce attributions to a shared group membership, thus hindering perceptions of dispositional cooperativeness of outgroup players (Prediction 2). Second, we predicted that dispositional attributions

about outgroup members would be elevated only by observations of intergroup cooperation, which was hypothesized to heighten weak identifiers' expectations about cross-group interactions and also to serve as demonstrations of group norms for strong identifiers (Prediction 3). However, outgroup players were rated as more cooperative for strong identifiers when involved in intragroup interactions than in intergroup interactions. These findings suggest that contrary to the literature on outgroup entitativity, cohesive outgroup interactions did not lead to inferences about exclusive within-group reciprocity. Instead, prosocial attributions became greater as the number of cooperative outgroup exemplars observed in the scenarios increased: perceived cooperativeness about outgroup members was highest among participants observing the outgroup scenario, followed by participants observing the intergroup scenario, and then those observing the ingroup scenario. Further, the presence of sanctions did not alter cooperativeness inference processes. In sum, these findings indicate that observing within-group cooperative acts by outgroup exemplars reduced the ultimate attribution error by increasing dispositional attributions about outgroup members. This change in attributions led to attenuation of discriminative decisions.

Exposure to Positive Outgroup Exemplars

We had expected that cooperative behaviors by outgroup members toward an ingroup recipient (in the intergroup scenario) should give rise to greater dispositional attributions about the outgroup than when the recipient was an outgroup member. However, we found that even when no ingroup members were involved in the observed interactions, witnessing cooperation among outgroup players was sufficient to induce dispositional cooperativeness inferences about the outgroup. The implication is that a key to encouraging cross-group cooperation is not necessarily demonstrations of outgroup members' willingness to cooperate across group boundaries (as the mechanism depicted in our intergroup model implies), but the knowledge of dispositional cooperativeness of outgroup members, even though it results from their genuine concern about other outgroup members. A possibility as to why outgroup members were judged as more cooperative following outgroup interactions than intergroup interactions was that exposure to only prosocial outgroup exemplars (i.e., the outgroup scenario) might make one's own group seem less altruistic than the outgroup since outgroup members cooperated with each other so overwhelmingly. According to the covariation principles in attribution theories (Jones & Davis, 1965), dispositions of an actor are constructed by comparing his/her behavior with one's own expectation about the performance of the majority of people (including oneself) in a similar situation. That is, if

participants had generated an idea about how to play in the Prisoners' Dilemma game before their observations, they might find the highly cooperative level of outgroup exemplars exceeding their expectations. Consequently, they might consider the outgroup to be dispositionally cooperative (Gilbert, 1998; Jones & Davis, 1965; Kelley, 1967).

Importantly, because we used the minimal group paradigm, there was no competition or animosity among the two groups (De Cremer & Van Vugt, 1999; Hewstone et al., 2002), and thus participants might have felt free to cooperate with members from different teams, and have anticipated little preferential treatment within group boundaries. In this case, cross-group cooperation might become more or as attractive as ingroup affiliation as the outgroup was perceived as altruistic. In Foddy and colleagues (2009), members from an ingroup (i.e., psychology students) vs. a more stereotypically virtuous outgroup (i.e., nursing students) or a less virtuous outgroup (i.e., economics students) were offered to participants as potential partners in economic games. They found that when the group membership of participants was unknown to their partners, participants chose their allocators based on relative prosociality of the groups; however, when their group membership was disclosed to the allocators (similar to our studies), they greatly favored the ingroup over the outgroups, in line with their expectations that ingroup allocators may reciprocate

more than outgroup allocators. What particularly relates to our studies is that despite expecting greater reciprocity from the ingroup, ingroup-favoring decisions with a more virtuous outgroup as an alternative option was 76%, i.e., 13% (not significantly) less than 89% when the outgroup was a less virtuous outgroup (p. 421). It shows that to some extent, outgroup positivity may make cross-group cooperation look more appealing than ingroup affiliation. Similarly, across two studies, the results of allocations in the trust games also showed that ingroup bias was significant only for participants observing the ingroup or intergroup scenario but not for participants who observed only prosocial outgroup exemplars (although the key interaction was not significant in Study 1).

Interplay of Behavioral Causes

We had hypothesized that a situational cause (i.e., shared group memberships) might hinder inferences about dispositional cooperativeness; however, it was not always the case in the present studies. For instance, outgroup players cooperating with each other were judged as equally cooperative as ingroup players; that is, cohesiveness of the outgroup did not dampen perceived prosocial qualities of outgroup members, as the literature on group entitativity suggests. Additionally, in Study 1, we found that highly identified participants tended to infer both dispositional attributions and shared-group-membership attributions than weak

identifiers. This implies that causes like a shared group membership might not impede dispositional attributions, and instead the two causes can co-exist without cancelling out each other. When we further analyzed the correlations between the shared-group-membership cause and dispositional attributions in different scenarios. Shared-group-membership attributions positively correlated with dispositional attributions (rs = .166 ~ .212, ps < .05) for both the ingroup and outgroup scenarios, but the correlations were nonsignificant for the intergroup scenario (rs = .057 ~ .074, ps > .3). In other words, an outgroup actor can be considered as likely to act based on his group membership and still perceived as dispositionally cooperative.

In contrast, instrumental causes such as strategic cooperation or only cooperating in the presence of sanctions might collide with inferences about dispositional cooperativeness. Indirect evidence by Van Hiel et al. (2008) indicated that participants who were asked to provide possible causes for cooperative (vs. noncooperative) targets in commons dilemmas attributed behaviors of cooperative targets to a higher level of genuine concern and a lower level of greed and fear as opposed to behaviors of noncooperative targets. The strategic-cooperation and risk-prevention causes in our studies seemed to correspond to attributions to greed and fear. In both studies, genuine concern positively correlated with risk prevention (rs = .207 ~ .327), but negatively correlated with strategic cooperation (rs = -.109 ~ -

.267). Compared to Van Hiel et al. (2008), we obtained a much smaller magnitude of correlation between strategic cooperation and genuine concern, perhaps because the responses in present studies consisted of a high proportion of cooperative choices but only few uncooperative choices.

Group Identification

With regard to the effect of group identification, our hypothesis was that highly identified individuals would perceive a higher level of outgroup threat from intragroup interactions among outgroup members, thus leading to biased decisions (as the mechanism depicted in the intragroup interaction model). However, in both studies, we found that although highly identified individuals perceived a higher level of outgroup threat, the perceived threat remained the same level regardless of cohesiveness among the observed players in different scenarios. Importantly, outgroup threat by strong identifiers seemed to be dissociated from their inferences about the outgroup's dispositions. In Study 2, the perceived differences in the outgroup's dispositions inferred from the scenarios were mainly driven by strong identifiers, signifying that they were attentive to cooperative opportunities afforded by outgroup members once the information was available. Weak identifiers were not as responsive to information about outgroup prosociality as strong identifiers. These findings seem contradictory to the CCM at first sight, but if we examine the

preconditions closer, the current approach and CCM emphasize different aspects of intergroup decisions. The CCM posits that weakly identifiers are more likely to accept cross-group cooperation than strong identifiers if a favorable outcome can be guaranteed by cooperative affordances other than a group membership (Packer & Kugler, under review). What drives weak identifiers is a temporary incentive afforded, for example, by the presence of sanctioning systems. However, our current findings indicate that strong identifiers may engage in cross-group cooperation even more assertively than weak identifiers if they find that cooperativeness of outgroup members exceeds that of ingroup members.

Worchel and Coutant (2004) point out that the reference group individuals use to identify themselves may shift during different phases of group development. In qualifying social identity theories, Worchel et al. (2000) propose that comparisons between the ingroup vs. outgroups occur upon formation of an group as a means to develop group identity and group boundaries, whereas comparisons within the ingroup become more prevalent at an later stage so as to evaluate or stabilize one's own status within the group. As such, because the minimal group paradigm was used in the present studies, participants may be at an early group development stage where the main focus was to disambiguate the relations between groups and to construct the representations of the ingroup vs. outgroup. We believe

that individuals highly identified with the ingroup may be also those who attend to instances informative about intergroup relations the most at an early group development stage. That may be why the instances of cooperative outgroup exemplars were more attributed to cooperative dispositions among strong identifiers (vs. weak identifiers) across the two studies.

An alternative explanation underlying strong identifiers' susceptibility to outgroup positivity is a desire to feel socially connected, especially with members from novel groups. It is possible that strong identifiers were simply social connectors and more inclined to affiliate with people they first met and found to be nice. As such, they might be more responsive to positive behaviors of outgroup members than weak identifiers. However, despite their high appraisals of outgroup's positive behaviors, strong identifiers showed a ingroup bias in dispositional attributions in that they perceived ingroup members as highly cooperative across three scenarios where positive instances of ingroup members varied. In contrast, weak identifiers accordingly made lower dispositional attributions about ingroup members when exposed to overwhelmingly positive instances of outgroup members relative to exposed to ingroup cooperation. In sum, strong identifiers in the present studies seemed to demonstrate both ingroup preference and tendency to affiliate with positive outgroup exemplars.

Similar to the desire to feel socially connected, a feeling of secure may also come into play in situations where strong identifiers are willing to cooperate across group lines. Individuals highly identified with the ingroup might also feel secure to search out resources that help achieve their goals. Therefore, strong identifiers might feel freely to recognize cooperative outgroup members as potential resource, and thus have perceived them as more cooperative as their positive behaviors increased. In a related study, Saleem and Anderson (unpublished manuscript) found that individuals with secure attachment primes—they imagined that people surrounding them were willing to help them to solve a problem—displayed less implicit bias toward Arabs compared to those with either neutral or insecure primes. Most importantly, the secure prime attenuated the Arabic implicit bias the most among individuals highly identified with the ingroup (American). In this case, the feeling of secure indeed encouraged highly identified individuals to stray beyond group boundaries.

Limitation and Future Directions

Extended contact theories propose that the knowledge of cross-group friendships reduces intergroup biases via three routes: (1) demonstrating cross-group interactions as a group norm by ingroup exemplars, (2) providing information about outgroup positivity, and (3) in advance encouraging the inclusion of outgroup others

as part of the self (Wright et al., 1997). Wright et al. (1997) emphasized that the preconditions for extended contact theories are that one needs to consider the ingroup exemplar as an interchangeable self, and the cross-group friendship as representative of intergroup relations. Given cooperative behaviors repeatedly demonstrated by members from both groups in Study 2, it was unlikely that intergroup interactions in the observation phase failed to represent intergroup relations. As such, we speculate that the nature of newly-formed groups may dilute the influence of group norms, leaving questions regarding the impact of ingroup role models. Moreover, mere observation itself may also leave some room for participants to have different interpretations as to why the observed players cooperated with each other (e.g., experiment assignment or voluntary decision). In contrast, participants in Wright and colleagues' (1997) studies were randomly assigned to a minimal group and observed two confederates from an ingroup vs. outgroup interacting in a friendly, neutral or hostile way (Study 4). Only participants observing friendly cross-group interactions showed no bias on evaluations on the outgroup target, whereas those observing neutral or hostile cross-group interactions evaluated the outgroup target demonstrating less positive traits and more negative traits. Therefore, the effect of observing different scenarios may be altered if the relation between groups is manipulated. For example, observations of outgroup

cooperation may lead to more perceived outgroup threat and attributions to a shared group membership when there is competition between groups.

In a real-world setting, people have many motives to believe that their own group outperforms an outgroup and that within-group cooperation would benefit them more than cross-group cooperation (Tajfel & Turner, 1979). We expect that our attribution model (see Figure 1) may be more applicable to existing social groups. According to the extensive literature on stereotypes and intergroup bias, interpretations of cooperative behaviors of an outgroup are contingent on its stereotypes (Duncan, 1976). Outgroup behaviors are encoded in a stereotype-consistent manner, which sets a stage for judgments in line with this biased information processing, especially for prejudiced people (Sherman et al., 2005). Put it another way, cooperative behaviors of a stigmatized outgroup would be less attributable to dispositions but more attributable to a shared group membership. Thus, the ultimate attribution error may be especially pronounced for outgroups with negative stereotypes (Hewstone, 1990). Furthermore, cooperation among outgroup members, in this case, might increase perceptions of cohesion, thus paving the way for (malicious) intentionality inferences about outgroup members, as is suggested in the literature on group entitativity (Dasgupta et al., 1997). As a result, we may observe cooperative behaviors demonstrated by outgroup members will be less

appreciated and render them less promising partners compared to the same behaviors by ingroup members.

In conclusion, our work lends support to existing literature that exposure to instances of outgroup positivity promotes future cross-group cooperation; we further extend the researchers' view (e.g., Tanis & Postmes, 2005) to attribution processes underlying further cooperation with outgroups. The findings suggest that dispositional attributions of outgroup behaviors lies at the heart of boosting confidence in cross-group cooperation. The significance of our work is that observations of outgroup cooperative acts can transfer to a stable representation (i.e., dispositions) about the outgroup; furthermore, the level of perceived outgroup cooperativeness increases with the number of observed instances. We also captured that this tendency was mainly contributed to individuals highly identified with the ingroup, suggesting that highly identified individuals may form stable representations/stereotypes about outgroups at a relatively early stage of group development, which may plant a seed for difficulties in changing their stereotypes about outgroups later on.

References

Abelson, R. P., Dasgupta, N., Park, J., & Banaji, M. R. (1998). Perceptions of the collective other. *Personality and Social Psychology Review, 2*, 243–250.

Allport, G. W. (1954).*The nature of prejudice*. New York: Addison-Wesley.

Berg, J., Dickhaut, J., & McCabe, K. (1995). Trust, reciprocity, and social history. *Games and Economic Behavior*, *10*, 122-142.

Blair, I.V. (2002). The malleability of automatic stereotypes and prejudice. *Personality and Social Psychology Review, 6*, 242-261.

Brewer, M. B. (2008). Depersonalized trust and ingroup cooperation. In J. I. Krueger (Ed.), *Rationality and Social Responsibility* (pp. 215–232). New York, NY: Psychology Press.

Brewer, M. B., & Kramer, R. M. (1986). Choice behavior in social dilemmas: Effects of social identity, group size, and decision framing. *Journal of Personality and Social Psychology, 50,* 543-549.

Campbell, D. T. (1958). Common fate, similarity, and other indices of the status of aggregates of persons as social entities. *Behavioral Science, 3*, 14–25.

Carpenter, S. & Radhakrishnan, P. (2002). The Relation between Allocentrism and Perceptions of In-groups. *Personality and Social Psychology Bulletin, 28*, 1528-1537.

Cialdini, R. B. (1996). Social influence and the triple tumor structure of organizational dishonesty. In D. M. Messick & A. E. Tenbrunsel (Eds.), *Codes of conduct: Behavioral research into business ethics.* New York: Russel Sage Foundation.

Dasgupta, N., Banaji, M. R., & Abelson, R. P. (1999). Group entitativity and group perception: Associations between physical features and psychological judgment. *Journal of Personality and Social Psychology, 77*, 991–1003.

Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology, 56*, 680–690.

Deutsch, M. (1958). Trust and suspicion. *Journal of Conflict Resolution, 2*, 265-279.

De Cremer, D., & Van Vugt, M. (1999). Social identification effects in social dilemmas: A transformation of motives. *European Journal of Social Psychology, 29*, 871-893.

Duncan, B. L. (1976). Differential social perception and attribution of intergroup violence: Testing the lower limits of stereotyping of Blacks. *Journal of Personality and Social Psychology, 34*, 590–598.

Foddy, M., Platow, M., & Yamagishi, T. (2009). Group-based trust in strangers: The roles of stereotypes and expectations. *Psychological Science, 20,* 419-422.

Gerard, H. B., & Hoyt, M. E (1974). Distinctiveness of social categorization and attitude toward ingroup members. *Journal of Personality and Social Psychology, 29*, 836-842.

Gilbert, D. T. (1998). Ordinary personology. In D. T. Gilbert, S. T. Fiske, & G. Lindzey, (Eds.), *Handbook of Social Psychology* (4th Ed.). New York: McGraw Hill.

Good, D. (1988). Individuals, interpersonal relations and trust. In D. Gambetta, (Ed.), Trust: Making and Breaking Cooperative Relations (pp. 31–48). New York, N.Y.: Blackwell.

Hamilton, D. L., & Sherman, J. W. (1994). Stereotypes. In R.S. Wyer, Jr., & T.K. Srull (Eds.), *Handbook of Social Cognition* (2nd Ed., Vol. 2, pp. 1-68). Hillsdale, NJ: Erlbaum.

Heider, F. (1958). *The psychology of interpersonal relations*. New York: Wiley.

Hewstone, M. (1990). The 'ultimate attribution error'?: A review of the literature on intergroup causal attribution. *European Journal of Social Psychology, 20*, 311–335.

Hewstone, M., Rubin, M. & Willis, H. (2002). Intergroup bias. *Annual Review of. Psychology, 53*, 575–604.

Howard , J. Rothbart , M. (1980). Social categorization and memory for ingroup and outgroup behavior. *Journal of Personality and Social Psychology, 58*, 301-308.

Insko, C. A., Schopler J., & Sedikides C. (1998). Personal control, entitativity, and evolution. In C. Sedikides, J. Schopler, C. A. Insko (Eds.), *Intergroup Cognition and Intergroup Behavior* (pp. 109–20). Mahwah, NJ: Erlbaum.

Jones, E. E., & Davis, K. E. (1965). From acts to dispositions: the attribution process in social psychology. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 2, pp. 219-266), New York: Academic Press.

Kelley, H. H. (1967). *Attribution theory in social psychology*. In D. Levine (Ed.), Nebraska Symposium on Motivation (Vol. 15, pp. 192-238). Lincoln: University of Nebraska Press.

Lin S. Y., & Packer, D. J. (2014). *Rapid Evaluative Tuning: The Presence of Cooperation-Facilitating Institutions Reduces Implicit Racial Bias*. Poster session presented at the 15th Annual Meeting of the Society of Personality and Social Psychology, Austin, TX.

Luhmann N. (1979). *Trust and Power*. New York, NY: John Wiley and Sons.

McConahay, J. B. (1986). Modern racism, ambivalence, and the modern racism scale. In J. F. Dovidio & S. L. Gaertner (Eds.), *Prejudice, discrimination, and racism.* London: Academic Press.

Mulder, L. B., Van Dijk, E., De Cremer, D., & Wilke, H. A. M. (2006a). When sanctions fail to increase cooperation in social dilemmas: Considering the presence of an alternative option to defect. *Personality and Social Psychology Bulletin, 13*, 1312-1324.

Mulder, L. B., Van Dijk, E., De Cremer, D. & Wilke, H. A. M. (2006b). Undermining trust and cooperation: The paradox of sanctioning systems in social dilemmas. *Journal of Experimental Social Psychology, 42*, 147-162.

Newheiser, A., Sawaoka, T., & Dovidio, J. F. (2012). Why do we punish groups? High entitativity promotes moral suspicion. *Journal of Experimental Social Psychology, 48*, 931-936.

Packer, D. J., & Kugler, M. B. (2013). *Cooperation-enhancing social institutions reduce strategic intergroup biases*. Working paper available at SSRN.

Packer, D. J., & Kugler, M. B. (under review). *Whom can I trust? A cooperative contingencies model of strategic intergroup bias.* Manuscript under review.

Preacher, K. J., & Hayes, A. F. (2004). SPSS and SAS procedures for estimating indirect effects in simple mediation models. Behavior Research Methods, Instruments, and Computers, 36, 717–731.

Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. Behavior Research Methods, 40, 879–891.

Pettigrew, T. F. (1979). The ultimate attribution error: Extending Allport's congnitive analysis of prejudice. *Personality and Social Psychology Bulletin, 5*, 461–476.

Pettigrew, T. F. (1991). Normative theory in intergroup relations: Explaining both harmony and conflict. *Psychology and Developing Societies, 3*, 3-16.

Pettigrew, T. F., & Tropp, L. R. (2006). A meta-analytic test of intergroup contact theory. *Journal of Personality and Social Psychology, 90*, 751–783.

Plant, E. A., & Devine, P. G. (1998). Internal and external motivation to respond without prejudice. *Journal of Personality and Social Psychology, 75*, 811–832

Ross, L., Amabile, T. M., & Steinmetz, J. L. (1977). Social roles, social control, and biases in social-perception processes. *Journal of Personality and Social Psychology, 35*, 485-494.

Saleem, M., Prot, S., Anderson, C.A., (unpublished manuscript). Ingroup identification moderates effects of secure attachment primes on intergroup bias and conflict.

Sherman, J. W., Stroessner, S. J., Conrey, F. R., & Azam, O. (2005). Prejudice and stereotype maintenance processes: Attention, attribution, and individuation. *Journal of Personality and Social Psychology, 89,* 607-622.

Stanley, D., Sokol-Hessner, P., Banaji, M. R., Phelps, E. A. (2011). Implicit race attitudes predict trustworthiness judgments and economic trust decisions. *Proceedings of the National Academy of Science, 108*, 7710-7715.

Stephan, W. G., Boniecki, K. A., Ybarra, O., Bettencourt, A., Ervin, K. S., Jackson, L. A., McNatt, P. S., & Renfro, C. L. (2002). The role of threats in the racial attitudes of Blacks and Whites. *Personality and Social Psychology Bulletin, 28,* 1242-1254.

Tajfel, H., & Turner, J. C. (1979). An intergrative theory of intergroup conflict. In W. G. Austin & S. Worchel (Eds.), *The Social Psychology of Intergroup Relations* (pp. 33–47). Monterey, CA: Brooks/Cole.

Tanis, M., & Postmes, T. (2005). A social identity approach to trust: Interpersonal perception, group membership and trusting behaviour. *European Journal of Social Psychology, 35,* 413-424.

Tenbrunsel, A. E., & Messick, D. M. (1999). Sanctioning systems, decision frames, and cooperation. *Administrative Science Quarterly, 44*, 684–707.

Terry, D. J., & Hogg, M. A. (1996). Group norms and the attitude behavior relationship: A role for group identification. *Personality and Social Psychology Bulletin, 22*, 776–793.

Van Lange , P. A. M. Liebrand , W. B. G. Kuhlman , D. M. (1990). Causal attribution of choice behavior in three N-person prisoner's dilemmas. *Journal of Experimental Social Psychology, 26,* 34-48.

Van Hiel, A., Vanneste, S. & De Cremer, D. (2008). Why did they claim too much? The role of causal attributions in explaining level of cooperation in commons and anticommons dilemmas? *Journal of Applied Social Psychology*, 38, 173-197.

Weber, M., Kopelman, S., & Messick, D. M. (2004). A conceptual review of social dilemmas: Applying a logic of appropriateness. *Personality and Social Psychology Review, 8*, 281–307.

Worchel, S. & Coutant, D. (2005). It takes two to tango: Relating group identity to individual identity within the framework of group development. (reprint). In

M. Brewer & M. Hewstone (Eds.), *Self and social identity*. Malden: Blackwell.

Worchel, S., Iuzzini, J., Coutant, D, & Ivaldi, M. A (2000). Multidimensional model of identity: Relating individual and group identity to intergroupp behavior. In R. Brown & D. Capozza (Eds.), Recent Developments in Social Identity Research. London: Blackwell.

Wright, S. C., Aron, A., McLaughlin-Volpe, T., & Ropp, S. A. (1997). The extended contact effect: Knowledge of cross-group friendships and prejudice. *Journal of Personality and Social Psychology, 73*, 73–90.

Yamagishi, T. & Kiyonari, T. (2000). The group as the container of generalized reciprocity. *Social Psychology Quarterly, 63,* 116-132.

Yzerbyt, V. Y., Corneille, O., & Estrada, C. (2001). The interplay of subjective essentialism and entitativity in the formation of stereotypes. *Personality and Social Psychology Review, 5,* 141-155.

Yzerbyt, V. Y., & Rogier, A. (2001). Blame it on the group: Entitativity, subjective essentialism and social attribution. In J. Jost & B. Major (Eds.), *The*

psychology of legitimacy: Emerging perspectives on ideology, justice, and intergroup relations (pp. 103-134). New York : Cambridge University Press.

Yzerbyt, V. Y., Rogier, A., & Fiske, S. (1998). Group entitativity and social attribution: On translating situational constraints into stereotypes. *Personality and Social Psychology Bulletin, 24*, 1090-1103.

Figure 1. A Model of Intragroup Cooperation

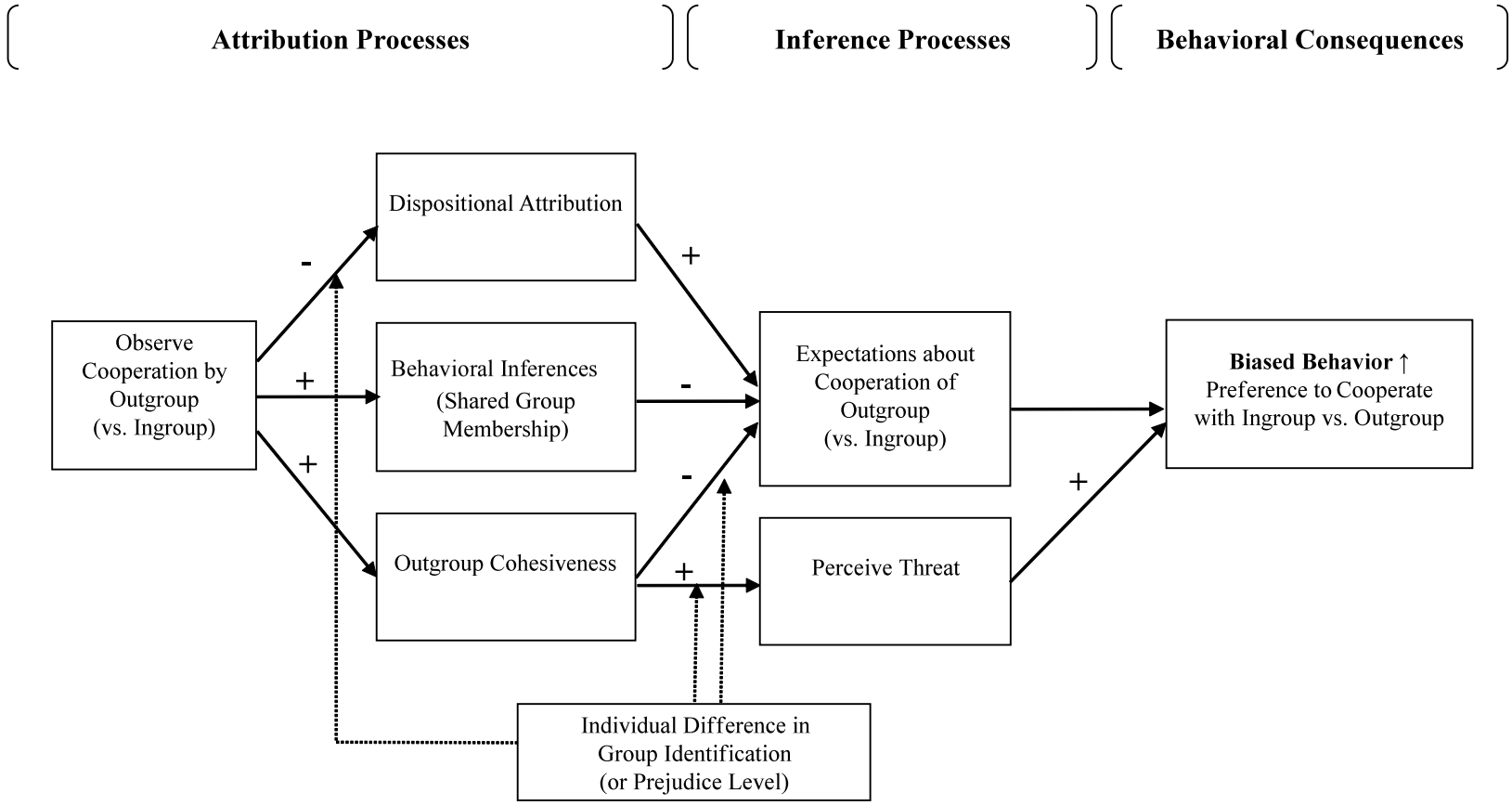


Figure 2. A Model of Intergroup Cooperation

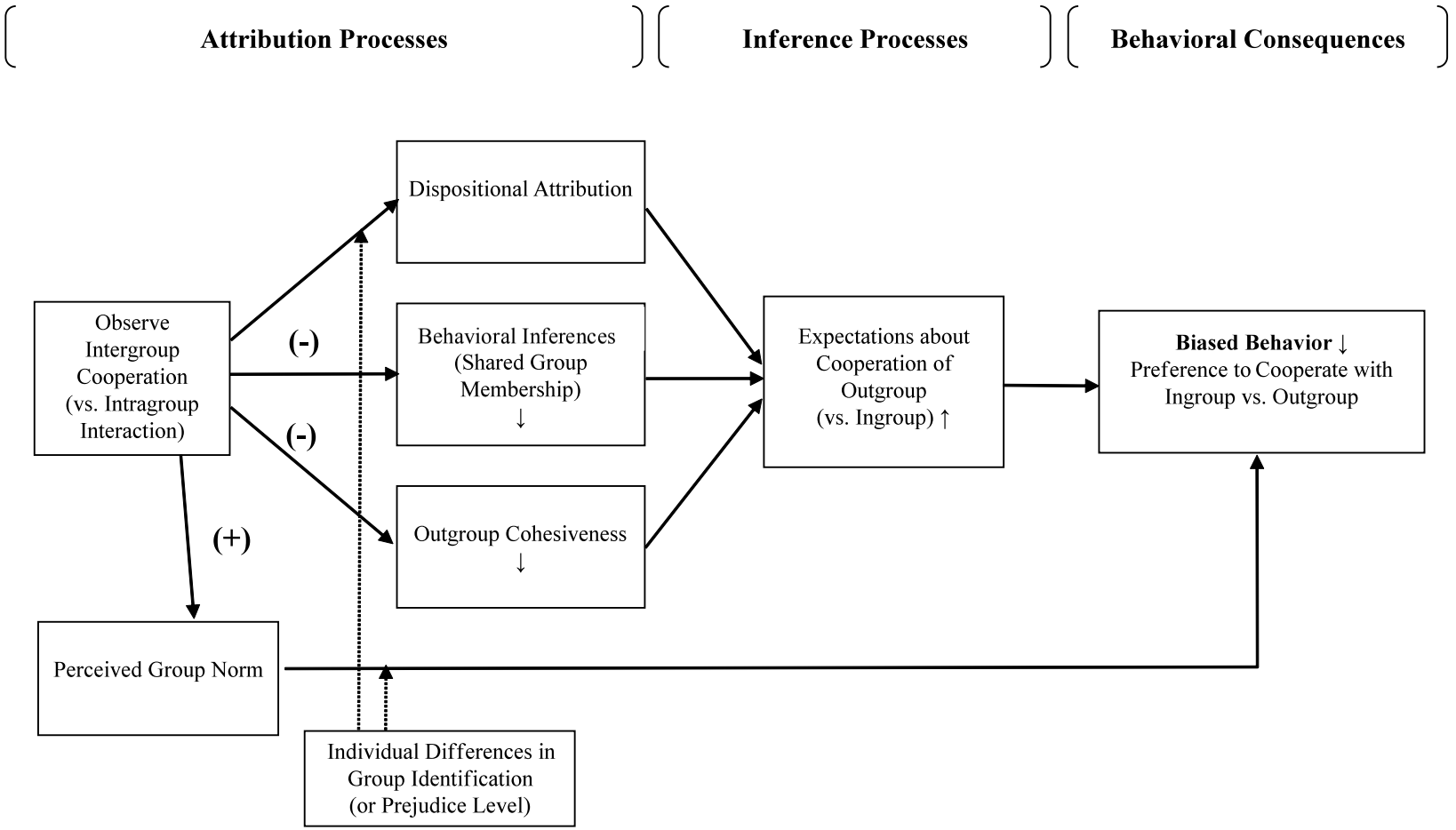


Figure 3. A Model of the Effect of Sanctioning Systems on Attribution

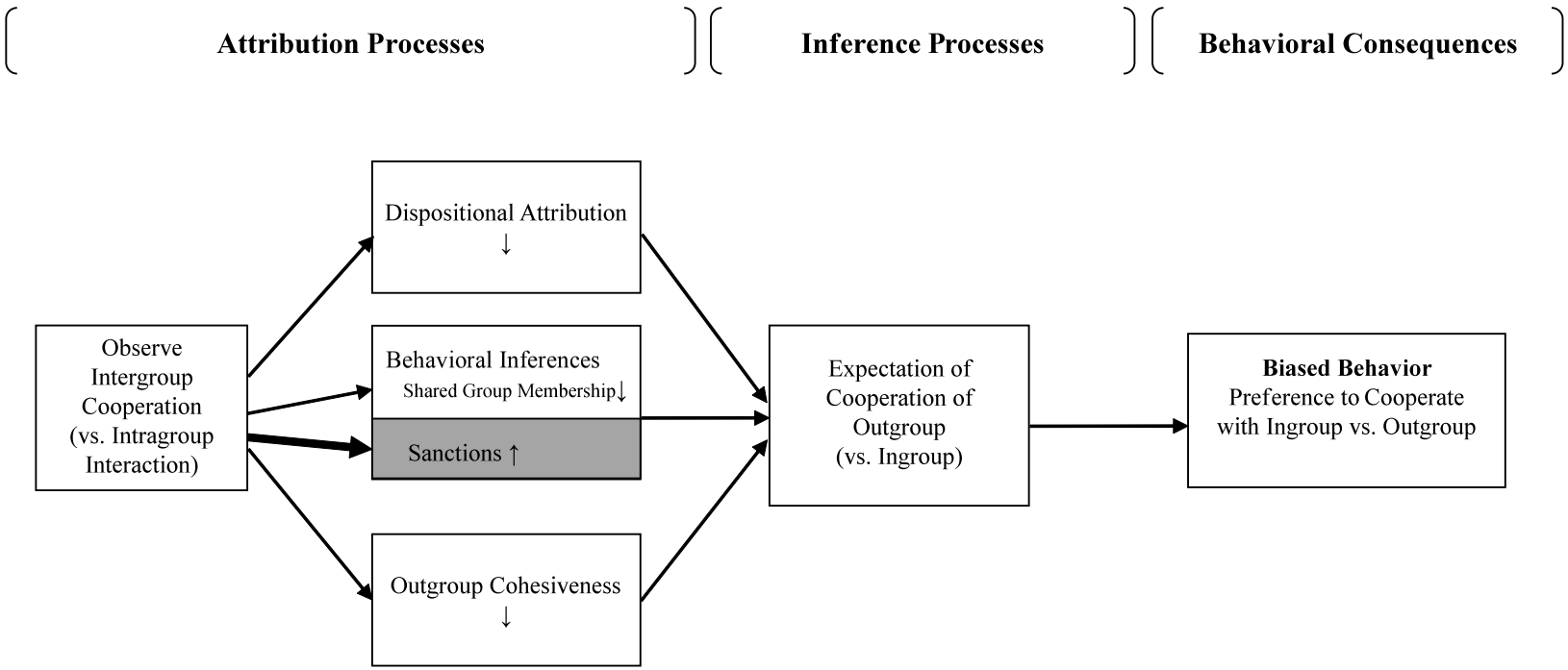


Figure 4. Effects of scenario x partner membership on allocations in the trust games in Study 1.

Figure 5. Effects of scenario x target membership on standardized scores on dispositional cooperativeness in Study 2.

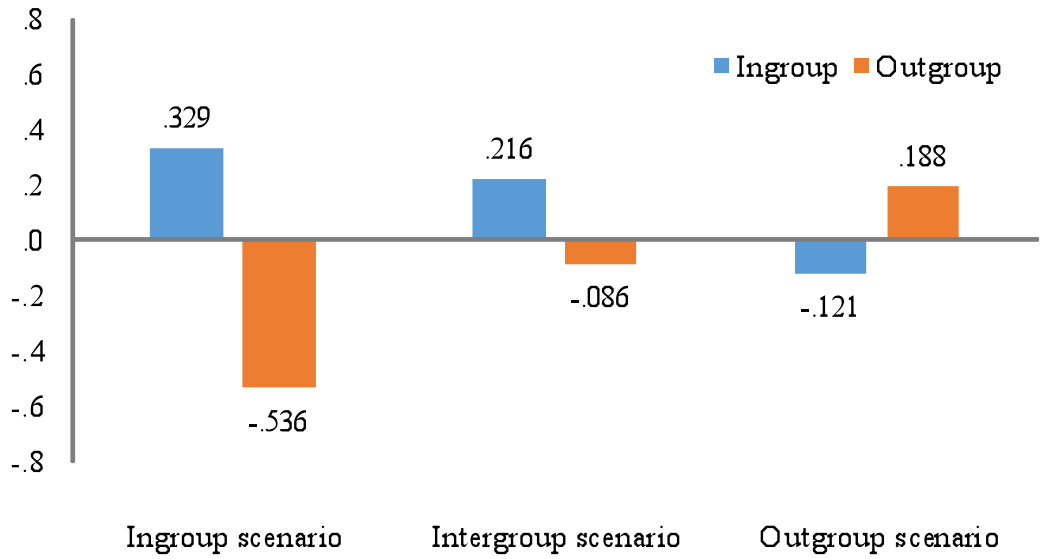


Figure 6. Scenario x group identification predicting dispositional attributions in Study 2.

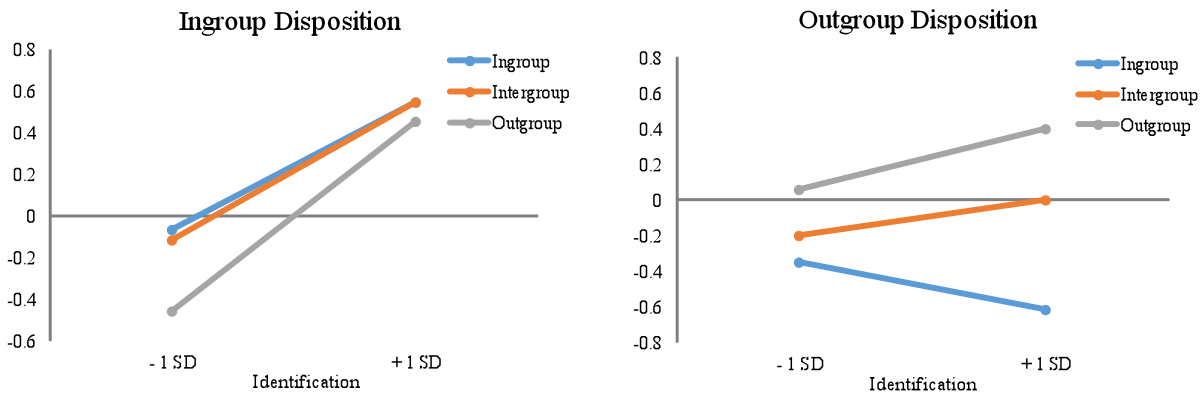


Figure 7. Effects of scenario x partner membership on allocations in the trust games in Study 2.

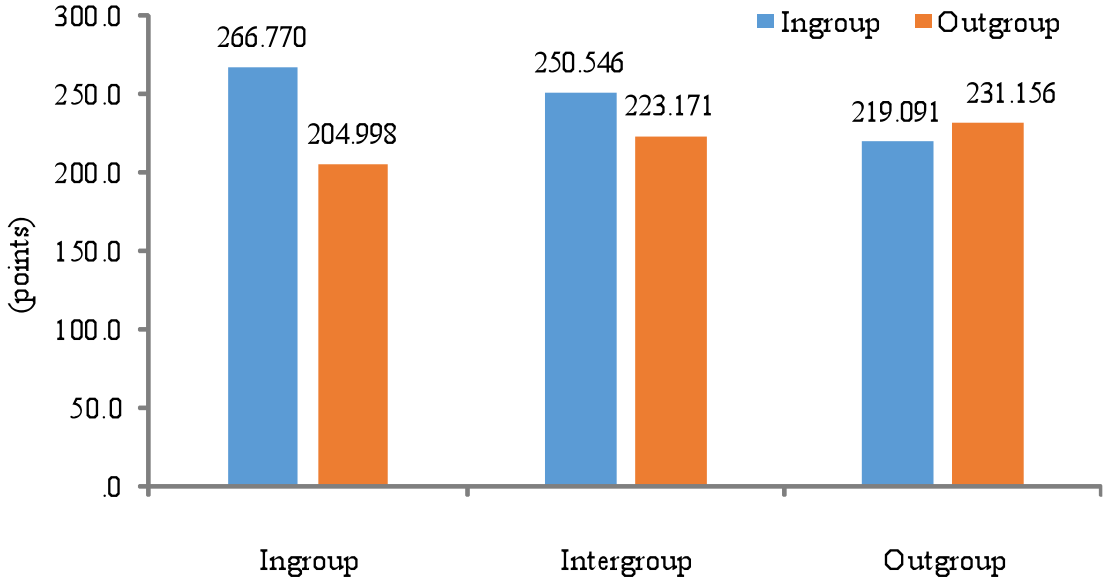


Table 1. Factor analysis of of causal attribution scale

	Player 1 Factor				**Player 2** Factor			
	1	2	3	4	1	2	3	4
1. This players decisions were influenced by the group membership of his partner.	**.827**	.093	.000	.058	**.864**	.002	.005	.041
2. This player would have made different decisions if the group membership of his partner had been different.	**.830**	-.050	.143	.139	**.868**	-.073	.128	.059
3. The fact that this player shared or did not share a group membership with his partner is important for understanding his behavior in these games.	**.720**	.261	-.122	.259	**.664**	.131	.017	.013
4. This player genuinely cared about the other player in these games.	.202	**.724**	-.094	-.252	.065	**.675**	.079	-.318
5. This player took the other player's needs into account when making decisions	.163	**.724**	-.004	-.263	.017	**.703**	.039	-.339
6. This player would have behaved in the same way no matter who he was playing with.	**-.636**	.518	.001	.267	**-.712**	.149	.193	-.012
7. This player would have been less cooperative if his outcomes did not rely on the other player	.225	-.086	.040	**.761**	.187	.212	-.012	**.865**
8. This player expected that the other player would reciprocate his decisions.	-.152	**.734**	.095	.223	-.133	**.946**	-.103	.450
9. This player wanted to maximize his own payoff (points)	.065	.002	-.050	**.848**	-.087	-.063	.142	**.697**
10. This player wanted to avoid negative evaluations from the other player	.068	.306	**.558**	-.131	.063	.182	**.630**	-.097
11. This player was afraid of being exploited by the other player	-.001	-.041	**.872**	.024	-.037	-.136	**.887**	.162
12. This player felt unsure about taking a risk	.004	-.023	**.819**	.018	-.063	-.010	**.783**	.041
Eigenvalue	3.375	2.050	1.412	1.170	3.298	2.126	1.504	1.018
% Explained Variance	28.123	17.083	11.763	9.748	27.487	17.716	12.529	8.482

Table 2. Cell means table

Scenario	Sanction Condition	Membership	No. of Participants	Standardized Scores on Disposition		Shared Group Membership		Concern for Others		Risk Prevention		Strategic Cooperation		Allocation in Trust Game (points)	
				Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
Intergroup Scenario	No Sanction	Ingroup	60	0.244	0.645	3.463	1.524	4.556	1.296	4.078	1.491	4.683	1.426	259.117	160.652
		Outgroup	60	-0.017	0.611	3.217	1.242	4.272	1.151	4.006	1.320	4.783	1.406	221.617	150.777
	Sanction	Ingroup	82	0.189	0.673	3.613	1.411	4.557	1.321	4.549	1.211	4.799	1.342	241.976	146.036
		Outgroup	82	-0.155	0.614	3.601	1.284	4.134	1.206	4.265	1.287	5.238	1.332	224.726	147.263
Ingroup Scenario	No Sanction	Ingroup	82	0.348	0.580	3.899	1.329	4.780	1.224	4.122	1.417	4.421	1.320	239.848	154.061
		Outgroup	82	-0.466	0.839	3.512	1.310	3.780	1.398	3.849	1.331	4.787	1.237	183.909	147.372
	Sanction	Ingroup	57	0.310	0.620	3.754	1.412	4.596	1.240	4.515	1.144	4.526	1.163	293.693	145.997
		Outgroup	57	-0.605	0.662	3.662	1.377	3.561	1.147	4.181	1.389	5.000	1.188	226.088	156.688
Outgroup Scenario	No Sanction	Ingroup	76	-0.138	0.791	3.701	1.264	3.943	1.275	4.259	1.261	4.737	1.147	206.750	137.747
		Outgroup	76	0.261	0.640	3.651	1.364	4.798	1.209	4.057	1.280	4.658	1.105	232.697	142.018
	Sanction	Ingroup	74	-0.104	0.672	3.453	1.426	4.067	1.352	4.230	1.190	4.703	1.271	231.432	146.061
		Outgroup	74	0.116	0.633	3.571	1.259	4.510	1.350	4.198	1.167	4.527	1.176	229.615	140.615

Table 3. Regression analysis

Predictors		Ingroup Dispositional Attributions Beta	t-value		Outgroup Dispositional Attributions Beta	t-value	
Step 1		$R^2 = .343^{**}$			$R^2 = .164^{**}$		
	Constant		4.584	**		-9.280	**
	Dummy 1 (Intergroup = 1 vs. Ingroup = 0)	-.003	-.076		.284	5.460	**
	Dummy 2 (Outgroup = 1 vs. Ingroup = 0)	-.170	-3.626	**	.482	9.103	**
	Mo: Identification (standardized)	.529	13.129	**	.076	1.681	
Step 2		$\Delta R^2 = .008^{+}$			$\Delta R^2 = .027^{**}$		
	Dummy 1 × ID	.024	.402		.188	2.880	**
	Dummy 2 × ID	.125	2.134	*	.239	3.649	**

$+ p < .1$, $* p < .05$, $** p < .01$

Appendix

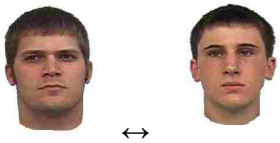
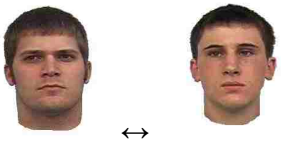
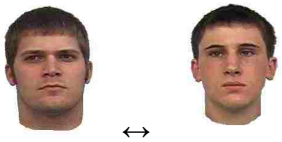
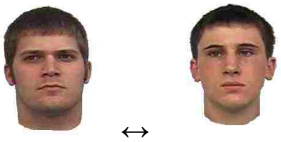
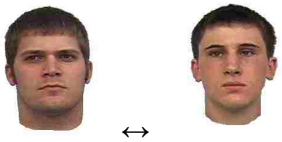
1. Choice matrix of Prisoners' dilemma

	Player 1 Cooperate	**Player 1 Defect**
Player 2 Cooperate	40 points / 40 points	60 points / 0 points
Player 2 Defect	0 points / 60 points	20 points / 20 points

2. Two players' choices observed by participants in all scenarios

Round	Player 1	Player 2
1	C	C
2	C	D
3	D	C
4	C	C
5	C	C
6	C	C
7	C	C
8	C	C
9	C	C
10	C	C

3. An example of five interaction scenarios

<table>
<tr>
<td>Ingroup–Control
Player 1 Player 2
Underestimator Unknown Membership
 
↔</td>
<td>Outgroup–Control
Player 1 Player 2
Unknown Membership Overestimator
 
↔</td>
<td></td>
</tr>
<tr>
<td>Ingroup–Ingroup
Player 1 Player 2
Underestimator Underestimator
 
↔</td>
<td>Outgroup–Outgroup
Player 1 Player 2
Overestimator Overestimator
 
↔</td>
<td>Ingroup–Outgroup
Player 1 Player 2
Underestimator Overestimator
 
↔</td>
</tr>
</table>

4. Dispositional attributions
(1 = not at all, 7 = very much)

Please rate what you think this person is like DEEP DOWN, in terms of their fundamental personality traits. What do you think they are really like?

1. friendly
2. considerate
3. thoughtful
4. sincere
5. kind
6. trustworthy
7. generous
8. likeable
9. cooperative
10. egalitarian
11. competent
12. intelligent

5. **Behavioral inferences** (modified from Van Hiel, Vanneste, & De Cremer, 2008)
(1 = very unlikely, 7 = very likely)

Please indicate the possible causes of the behavior of this player?

1. This player's decisions were influenced by the group membership of his partner.
2. This player would have made different decisions if the group membership of his partner had been different.
3. The fact that this player shared or did not share a group membership with his partner is important for understanding his behavior in these games.
4. This player genuinely cared about the other player in these games.
5. This player took the other player's needs into account when making decisions
6. This player would have behaved in the same way no matter who he was playing with.
7. This player would have been less cooperative if his outcomes did not rely on the other player
8. This player expected that the other player would reciprocate his decisions.
9. This player wanted to maximize his own payoff (points)
10. This player wanted to avoid negative evaluations from the other player
11. This player was afraid of being exploited by the other player
12. This player felt unsure about taking a risk

6. **Entitativity Scales** (modified from Carpenter & Radhakrishnan, 2002)
(1 = not at all, 7 = very much)

Please rate the extent to which you think each statement below likely describes the group of [team name]

1. They are similar in terms of their behaviors.
2. They are similar to each other with respect to their traits.
3. They agree about values.
4. They are well coordinated
5. They are efficient when they work together
6. They are interdependent, depending on each other

7. Outgroup threat (for minimal group paradigm)
(1 = strongly disagree, 7 = strongly agree)

Please rate to what extent you agree or disagree with each statement

1. I think that [group] is a strong rival to my group.
2. I believe that [group] and [group] are highly competitive with each other.
3. Our two groups are in competition for important resources.
4. I feel threatened by [group].
5. Our groups are different from each other in important ways.

8. Perceived group norms
(1 = strongly disagree, 7 = strongly agree)

Please rate to what extent you agree or disagree with each statement

1. Watching the two players earlier gave me a strong sense of how I should behave in these games.
2. I got a good understanding of how members of my group behave in these games.
3. Watching the players provided me with useful information about how people in general should play these games.

9. Trust/cooperative expectations
(1 = not at all, 7 = very much)

Please imagine if you were the other player in the game with your picture and group membership revealed to this player…

1. To what extent do you believe this player would cooperate with you?
2. To what extent do you believe this player would be motivated to maximize your outcome?
3. To what extent would you want to play with this player?
4. To what extent would you trust this partner?

10. **Ingroup identification scale** (Leach et al., 2004)
(1 = strongly disagree, 7 = strongly agree)

1. I feel a bond with [group]
2. I feel solidarity with [group]
3. I feel committed to [group]
4. I think that [group] has a lot to be proud of
5. It is pleasant to be [group]
6. I have a lot in common with the average [group] member
7. I am similar to the average [group] member
8. I identify with other [group] people

11. Political Orientation

1. Politically, do you think you tend to be more Liberal, or Conservative? (7 Point Scale, Very Liberal to Very Conservative)
2. On ECONOMIC issues, do you think you tend to be more Liberal, or Conservative? (7 Point Scale, Very Liberal to Very Conservative)
3. On SOCIAL issues, do you think you tend to be more Liberal, or Conservative? (7 Point Scale, Very Liberal to Very Conservative)

12. Demographics

Age: _____ Date of Birth: _____

Sex: Male ( ) Female ( )

Education (indicate highest level achieved):

( ) Grade 8 ( ) University degree or college diploma

( ) Grade 10 ( ) Master's degree

( ) High school diploma ( ) Doctoral degree

( ) Some university or college

Ethnicity

( ) White ( ) Hawaiian Native & Pacific Islander

( ) Black or African American ( ) Hispanic/Latino(a)

( ) Indian/Alaskan Native ( ) South or East Asian

( ) Other

How long have you lived in the United States? _____

Is English your first language: ( ) Yes ( ) No

If no, at what age did you learn English? _____

13. **Additional analyses**

STUDY 1

Dependent variables: (1) situational attributions including risk prevention and strategic cooperation; (2) group perceptions including perceived group norms, perceived entitativity, and outgroup threat were examined between the conditions with only ingroup players involved (i.e., Ingroup-Ingroup and Ingroup-Ctrl scenarios) and the conditions with only outgroup players involved (i.e., Outgroup-Outgroup and Outgroup-Ctrl scenarios). Responses to the two players in the Ingroup-Ingroup and Outgroup-Outgroup scenarios were averaged for participants in the homogeneous conditions, whereas for the control conditions, scores were based only on the player with his team membership revealed (ingroup or outgroup). Dependent variables were analyzed with 2 (scenario type: Homogeneous vs. Control) x 2 (player membership: Ingroup vs. Outgroup) between-subject factorial ANOVAs.

The results showed that there was a significant main effect of player membership on strategic-cooperation attributions, $F(1, 258) = 4.633$, $p = .032$. Opposed to our predictions, ingroup players were considered to be more strategic ($M = 5.055$) than outgroup players ($M = 4.746$) (we will discuss implications of this later). With respect to perceived group norms, we found a significant main effect of scenario type, $F(1, 258) = 7.779$, $p = .006$, which was qualified by a scenario type × player membership interaction, $F(1, 258) = 5.566$, $p = .019$. Observing the Ingroup-Ingroup scenario ($M = 4.990$) elevated perception of a clear group norm compared to its control condition ($M = 4.444$ for Ingroup-Ctrl scenario), $F(1, 258) = 5.041$, $p = .026$, while observation of the scenarios involving outgroup players did not influence perception of group norms ($M = 4.129$ for Outgroup-Outgroup scenario; $M = 4.373$ for Outgroup-Ctrl scenario), $F(1, 258) = 2.053$, $p = .29$. None of other effects was significant, Fs ≤ 3.817.

STUDY 2

Dependent variables including (1) situational attributions including risk prevention and strategic cooperation; (2) group perceptions including perceived group norms, perceived entitativity and outgroup threat were examined with *scenario* (between-subject: Ingroup-Ingroup vs. Outgroup-Outgroup vs. Intergroup) × *sanction* (between-subject: Present vs. Absent) × *target group membership* (within-subject: Ingroup vs. Outgroup) mixed ANOVAs. With respect to the inferences about risk prevention, participants considered behaviors of ingroup players ($M = 4.292$) to be more risk-preventive than behaviors of outgroup players ($M = 4.093$), $F(1, 425) = 7.72$, $p = .006$. Also, players in the sanction-present condition ($M = 4.323$) were perceived as more risk-preventive than players in the sanction-absent condition ($M = 4.062$), $F(1, 425) = 6.40$, $p = .01$.

The only evidence that confirmed our prediction was a significant interaction of *membership* × *scenario* on strategic cooperation, $F(2, 425) = 6.39$, $p = .002$. Participants observing ingroup or intergroup scenarios thought that strategic cooperation was more accountable for outgroup behaviors (Ms = 4.893 and 5.011) than ingroup behaviors (Ms = 4.474 and 4.741), $F(1, 137) = 11.75$, $p < .001$; $F(1, 140) = 5.67$, $p = .02$; however, participants observing outgroup scenario rated strategic cooperation as equally accountable for behaviors of both groups ($M = 4.720$ for ingroup; $M = 4.593$ for outgroup), $F(1, 148) = 1.50$, $p = .22$. Although we did not

observe the effect of sanctions on cooperativeness inferences, it seemed that the presence of sanctions increased attributions to strategic cooperation for outgroup behaviors.

Updated 4/25/2015

Vita

Shiang-Yi Lin

Email: shl812@lehigh.edu

Address: Department of Psychology
Lehigh University
17 Memorial Drive East
Bethlehem, PA 18015

EDUCATION

2012-present	**Doctoral student in Social Psychology**, Lehigh University
2014	**Master of Science in Psychology**, Lehigh University Thesis title: "Threat or opportunity? Attributions following high levels of cooperation within an outgroup" (Advisor: Dr. Dominic J. Packer)
2006-2007	**Master of Science in Psychology**, National Chung Cheng University Thesis title: "Retrieval-induced forgetting of stereotypical traits: Cross-category inhibition and valence asymmetry" (Advisor: Dr. Chuan-Feng Kao)
2000-2004	**Bachelor of Science in Psychology**, National Chung Cheng University Honored project title: "Role of implicit personality theory of intelligence in the modulation of the dilution effect" (Advisor: Dr. Chuan-Feng Kao)

POSITIONS HELD

2012-present	Doctoral student in Social Psychology, Lehigh University
2007-2011	Research Associate, Trendgo Consultant Co. Ltd., Taiwan
2004-2005	Research Assistant, Insight Marketing Research Consultancy, Taiwan

FELLOWSHIPS/SCHOLARSHIPS

2014-present	Teaching Assistantship, Psychology Department, Lehigh University
2012-2014	Fulbright Graduate Study Grants, Fulbright Foundation
2006	Graduate Research Scholarship, Psychology Department, National Chung Cheng University

HONORS AND AWARDS

2015	Diversity Travel Award, Society for Personality and Social Psychology
2014-2015	Graduate Students Travel Grants, Lehigh University
2007	Outstanding Master's Thesis Award [the Su Shiang-Yeu Award], Taiwanese Psychological Association

CONFERENCE PRESENTATIONS

Packer, D. J., Kugler, M. B., & **Lin S-Y.** (May, 2015). *How social institutions affect decisions to make contact across group boundaries*. Talk presented at the 27th annual convention of the Association for Psychological Science, New York, NY.

Lin S-Y., & Packer, D. J. (February, 2015). *Threat or opportunity? Attributions following high levels of cooperation within an outgroup*. Poster session presented at the 16th annual meeting of the Society for Personality and Social Psychology, Long beach, CA.

Packer, D. J. & **Lin S-Y.** (February, 2015). *Who can I trust?: Rapid evaluative tuning and changes in intergroup trust in response to variable cooperative opportunities*. Talk presented at the 16th annual meeting of the Society for Personality and Social Psychology, Long beach, CA.

Lin S-Y., & Packer, D. J. (February, 2014). *Rapid evaluative tuning: the presence of cooperation-facilitating institutions reduces implicit racial bias*. Poster presented at the 15th annual meeting of the Society for Personality and Social Psychology, Austin, TX.

Chen, H. C., & **Lin, S-Y.** (October, 2011). *Explicit and implicit emotional words in autobiographical event recollections.* Talk presented at the 50th annual meeting of the Taiwanese Psychological Association, Taichung, Taiwan.

Chen, H. C., Cho S.L., & **Lin, S-Y.** (September, 2011). *Emotion-related word counting system: a comparison to LIWC.* Talk presented at the 7th Chinese Psychologist Conference, Taipei, Taiwan.

Lin, S-Y. & Kao, C. F. (October, 2007). *Retrieval-induced forgetting of stereotypical traits: Cross-category inhibition and valence asymmetry.* Poster presented at the 46th annual meeting of Taiwanese Psychological Association, Tainan, Taiwan.

PUBLICATIONS

Lin, S-Y., & Kao, C. F. (2008). Retrieval-induced forgetting of stereotypical traits: Cross-category inhibition and valence asymmetry. *Chinese Journal of Psychology*, *50*, 403-423. [In Chinese] (To read the abstract please follow http://www.airitilibrary.com/searchdetail.aspx?DocIDs=10139656-200812-50-4-403-423-a)

RESEARCH INTERESTS

My primary research interest is to investigate how intergroup relations influence biased information processing (e.g., attributional biases or implicit stereotyping). My second line of research explores the mechanisms underlying the flexible nature of implicit evaluations in response to situational contingencies.

PROFESSIONAL AFFILIATIONS

Association for Psychological Science (APS)

Society for Personality and Social Psychology (SPSP)

TEACHING EXPERIENCES

2014-2015	Teaching Assistant, Statistics, Department of Psychology, Lehigh University
2006-2007	Teaching Assistant, Attitude and Persuasion, Social Psychology, and Stress Adjustment, Department of Psychology, National Chung Cheng University
2003	Lecturer, Social Psychology, Psychology Winter Camp, National Chung Cheng University

TECHNICAL SKILLS

Data Collection: E-Prime, Qualtrics, and Mechanic Turk

Statistics: SPSS, SAS, R, and LISREL

Programming: Delphi, Visual Basic, and HTML