THE CONTAGION OF SOCIAL BEHAVIOR

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Milena Venelinova Tsvetkova

May 2015

© 2015 Milena Venelinova Tsvetkova

ALL RIGHTS RESERVED

THE CONTAGION OF SOCIAL BEHAVIOR

Milena Venelinova Tsvetkova, Ph.D.

Cornell University 2015

The dissertation investigates the contagion of prosocial and antisocial behavior among strangers. We distinguish between two contagion mechanisms: generalized reciprocity (a recipient of social behavior is more likely to pay it forward) and third-party influence (an observer of social behavior is more likely to emulate it). Using two large-scale online experiments, we find that individuals who have benefitted from generosity and suffered from mean behavior are more likely to help and, respectively, harm others. Individuals who observe many acts of kindness are more generous towards others, while individuals who observe few mean acts are less likely to be mean towards others. We then conduct computational experiments with an agent-based model to investigate when the spread of prosocial behavior can become self-sustaining. The results offer explanation for the fact that cultures of kindness are rare for anonymous face-to-face interactions but common online, for example in the form of user-generated content communities.

BIOGRAPHICAL SKETCH

Milena Tsvetkova holds a Bachelor of Science in Architecture Design from the Massachusetts Institute of Technology, a Master of Arts in the Social Sciences from the University of Chicago, and a Master of Science in Sociology and Social Research from Utrecht University. Her research interests lie in the fields of computational and experimental social science. She uses online experiments, agent-based models, network analysis, and game theory to identify causal mechanisms and account for complex dynamics in large social groups.

На баба, за заразната ѝ доброта. На семейството ми, за безусловната им помощ. На Даня и Дьома, за това, че са.

ACKNOWLEDGEMENTS

This work was supported by grants SES-1260348 and SES-1303526 from the National Science Foundation. I would like to thank Michael Macy for his guidance, invaluable insights, and endless editorial improvements. I am also thankful to George Berry, Chris Cameron, Daniel DellaPosta, Scott Golder, Patrick Park, and Yongren Shi for their comments and suggestions. I am additionally indebted to John H. Miller, Scott E. Page, and the participants at the 2013 Graduate Workshop in Computational Social Science at the Santa Fe Institute for valuable feedback.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

People exert influence on each other in numerous ways. Friends and peers, as well as complete strangers, to a large degree impact our health and well-being (Christakis and Fowler, 2009; Centola, 2010), cultural tastes (Salganik et al., 2006; Muchnik et al., 2013), and political behavior (Bond et al., 2012). Do other people also affect us in how we behave towards members of our community?

This dissertation investigates the contagion of social behavior. In particular, we study the spread of other-oriented behavior, prosocial as well as antisocial, among anonymous individuals. To do this, we use insights from two of the largest literatures in the social sciences: on social contagion and on the emergence of cooperation. The study of social contagion tries to answer why and how certain innovations, ideas, attitudes, and behavioral patterns spread while others do not (Rogers, 2003; Christakis and Fowler, 2009). The study of the emergence of cooperation is concerned with why people behave prosocially when individuals have the incentive to be self-interested (Axelrod, 1984; Taylor, 1987). We adopt the relational perspective common in the contagion literature to the study of social behavior, a research line that has generally focused on individuals responding independently or in aggregates (Fehr and Gintis, 2007). At the same time, we use insights on other-directed behavior to investigate contagion that occurs both by direct interaction and by observation, two types of spread dynamics that are usually studied separately (Young, 2009).

Previous research has suggested that generosity, as well as minor antisocial infractions may spread between individuals (Cialdini, 2008; Keizer et al., 2008; Fowler and Christakis, 2010; Suri and Watts, 2011; Jordan et al., 2013;

Keizer et al., 2013). Our contribution is to distinguish between two mechanisms through which such contagion can occur: generalized reciprocity (a recipient of social behavior is more likely to pay it forward) and third-party influence (an observer of social behavior is more likely to emulate it). We provide empirical evidence that prosocial and antisocial behavior spread differently through the two mechanisms. We do this with the help of an innovative experimental design and a new online platform for studying the diffusion of behavior in large social groups under controlled conditions. Our findings have implications for strategies for encouraging group-oriented contributions and controlling aggressive and self-serving actions.

1.1 THEORETICAL APPROACH

Following the tradition of analytical sociology (Elster, 1989; Coleman, 1990; Hedström, 2005), the dissertation focuses on identifying mechanisms, establishing causality, and accounting for complex dynamics. We set out to explain how global information, interaction structure, and interaction frequency lead to the spread of prosocial and antisocial behavior in social groups. Although our explanans and explananda are at the macro-level, or the level of the social aggregate, we trace the explanatory chain through the micro-level, or the level of the individual (Hedström and Ylikoski, 2010). Two of the dissertation chapters focus on action-formation mechanisms: how experiencing and observing certain behavior increases the likelihood to behave similarly. A third chapter concentrates on the transformational mechanisms: how individual actions and interactions lead to the spread of the behavior in the population.

By focusing on mechanisms, we aim to outline the causal process. In order to abstract from less relevant details, we assume, rather than investigate, any mechanisms at lower levels. For example, we do not disentangle the mechanisms behind generalized reciprocity. Generalized reciprocity could be driven by feelings of positive affect, gratitude, indebtedness, or obligation (Bartlett and DeSteno, 2006). In turn, these feelings can be genetically hardwired or socially acquired. Establishing the social-psychological mechanisms behind generalized reciprocity is a subject for another dissertation by itself. Instead, this dissertation focuses on distinguishing generalized reciprocity and third-party influence and identifying their contribution to the contagion of prosocial and antisocial behavior among anonymous individuals.

1.2 METHODOLOGY

We rely on the experimental method to isolate and study the two causal mechanisms. Specifically, we took advantage of the research opportunities opened by new digital technologies and conducted two controlled experiments with human subjects online and a computational experiment with an agent-based model.

Controlled experiments present the most robust method to isolate mechanisms and establish causality. Experiments manage to control for confounding mechanisms and contextual factors through random assignment to carefully designed manipulations. Further, they allow identifying general social interaction mechanisms by abstracting the decision situation from any specific social context.

1.2.1 Online experiments

Laboratory experiments have become established as the mainstream method for studying behavioral mechanisms. However, studying social contagion requires large groups to observe the occurrence of cascades and gathering a large number of participants over extended periods of time in a laboratory poses a challenge. We resolved this problem by designing, developing, and conducting experiments with human subjects interacting online.

Compared to traditional laboratory experiments, online experiments have a number of advantages. They involve a more demographically diverse population than the usual undergraduate students in American universities (Paolacci et al., 2010; Ross et al., 2010). They entail simpler participant recruitment, lower costs, and faster completion times (Mason and Suri, 2012). This in turn helps them scale up easily. Unsurprisingly, online experiments have been gaining an increasing number of followers recently (Lawson et al., 2010; Bohannon, 2011; Dodds et al., 2011; Suri and Watts, 2011).

Designing and conducting online experiments also present certain challenges: online subjects tend to have high drop-out rates, low motivation, and low attention spans; in addition, they sometimes use automated responses or participate multiple times (Horton et al., 2011; Mason and Suri, 2012). Yet, the aspect that we found most challenging with an online experimental platform was the fact that it presents weak stimuli. In laboratory settings, social-psychology experimenters often rely on confederates and the power of face-to-face interaction to create decision situations that would be difficult to generate on a computer screen. Unfortunately, this technique is not straightforward to implement online.

Rather than relying on deception, we instead attempted to address the problem of weak stimuli with a believable experimental design. We deliberately designed our experiments as games and made a point of explaining the game logic to participants. It was important to reveal the game design and avoid deception because online participants tend to share their experiences on forums. Any publicly expressed suspicion of deception could thus influence future participants and bias the current experiment, as well as subsequent experiments. Finally, carefully designed interaction experiments that avoid deception are also essential if the analysis needs to be done at the group level. Although we did not employ group-level analyses here, the experiments we designed can be adapted for such studies in the future.

1.2.2 Computational experiments

In situations where experiments with human subjects are unviable or prohibitively costly, computational experiments with agent-based models could be used instead (Hedström and Ylikoski, 2010). Agent-based models are thought experiments that investigate the macro-implications from a set of micro-level behavioral assumptions (Macy and Willer, 2002; Macy and Flache, 2009). In an agent-based model, the modeler specifies the properties and behavior of heterogeneous agents, the interaction rules, and the interaction structure. The properties of the population emerge out of the agents' interactions under the specified constraints. If the agents are interdependent, their actions and interactions often give rise to complex phenomena than cannot be easily intuited from the model's constituting elements. Numerical simulation experiments are often used to study the dynamics in such complex systems.

A major challenge for building agent-based models is to keep them as simple as possible. The goal is to identify the minimal assumptions that give rise to the aggregate pattern of interest. Although the models are most useful when they generate unintuitive results, ultimately, their purpose is to illuminate the causal chain. In other words, although the researcher might not be able to predict the results of a generative model, she should be able to explain them retroactively.

We employed simulation experiments with an agent-based model to study the population patterns generated when the two contagion mechanisms act together under different interaction conditions. We focused on prosocial behavior only. To keep the model simple, we built upon previous threshold models of contagion and incorporated our empirical findings on how recipients and observers of helping behavior react.

1.3 DISSERTATION OUTLINE

Chapter 2 and Chapter 3 are experimental studies that investigate the importance of generalized reciprocity and third-party influence for the contagion of prosocial and antisocial behavior, respectively. The two experiments use similar experimental design, software platform, and online subject pool. Chapter 4 is a simulation study that explores the macro-level implications from the empirical results on the spread of prosocial behavior (Chapter 2). Since the three chapters are each based on stand-alone articles, they can be read independently. Further, a summary of Chapter 1 is included in the first half of Chapter 4. [1]

[1] A Chapter 5 on the macro-level implications from the mechanisms for the spread of antisocial behavior is missing. The reason is that the empirical results in Chapter 3 are relatively weak. More research is needed to establish individuals' behavior when they experience or observe antisocial acts.

Chapter 2 investigates the mechanisms behind the contagion of prosocial behavior. In our first experiment, we presented participants with the opportunity to donate anonymously part of their payment to another participant. We found that receiving or witnessing a small number of donations increases generosity. Witnessing many acts of generosity actually decreases the proclivity to help but only among non-recipients.

Chapter 3 focuses on the mechanisms behind the contagion of antisocial behavior. In our second experiment, participants could anonymously take a portion of another participant's payment. We found that people are more likely to harm others if they have been harmed and they are less likely to do so if they observe that others do not harm.

The results from Chapter 2 suggested that receiving acts of kindness could make the spread of prosocial behavior self-sustaining. Chapter 4 uses computational experiments with an agent-based model to investigate this proposition. The results indicate two possible pathways for the emergence of successful voluntary-contribution communities among anonymous individuals. The two pathways offer explanation for the fact that cultures of kindness are rare for anonymous face-to-face interactions but common online, for example in the form of user-generated content communities.

1.4 ABBREVIATIONS USED

Throughout the dissertation, we use abbreviations to reduce verbosity. These abbreviations are not standard in the literature, so we introduce them here, as well as in every chapter that utilizes them:

- GR: generalized reciprocity
- TPI: third-party influence
- AMT: Amazon Mechanical Turk
- HIT: human intelligence task

CHAPTER 2

THE CONTAGION OF PROSOCIAL BEHAVIOR

"In the order of nature we cannot render benefits to those from whom we receive them, or only seldom. But the benefit we receive must be rendered again, line for line, deed for deed, cent for cent, to somebody.'' — Ralph Waldo Emerson

"...[W]hen you meet with another honest Man in similar Distress, you must pay me by lending this Sum to him; enjoining him to discharge the Debt by a like operation, when he shall be able, and shall meet with another opportunity. I hope it may thus go thro' many hands, before it meets with a Knave that will stop its Progress." — Benjamin Franklin, to a stranger whom he had given money

ABSTRACT

Why do people help strangers when there is a low probability that help will be directly reciprocated or socially rewarded? A possible explanation is that these acts are contagious: those who receive or observe help from a stranger become more likely to help others. We test two mechanisms for the social contagion of generosity among strangers: generalized reciprocity (a recipient of generosity is more likely to pay it forward) and third-party influence (an observer of generous behavior is more likely to emulate it). We use an online experiment with randomized trials to test the two hypothesized mechanisms and their interaction by manipulating the extent to which participants receive and observe

help. Results show that receiving help can increase the willingness to be generous towards others, but observing help can have the opposite effect, especially among those who have not received help. These results suggest that observing widespread generosity may attenuate the belief that one's own efforts are needed. [1]

2.1 INTRODUCTION

On a cold December morning in 2012, in the drive-through of the Tim Hortons in Winnipeg, Canada, a stranger generously picked up the tab for the coffee order of the next customer waiting in line. That person paid the bill of the next stranger in line. And so did the following 226 customers (Mallough, 2013). The practice of "paying it forward" spread not only to other customers of the restaurant but to other restaurants — the Chick-fil-A drive-through off Highway 46 in New Braunfels, Texas, a Dunkin' Donuts drive-through in Detroit, and a McDonald's drive-through in Fargo, North Dakota (Memmott, 2013; Murphy, 2013). "Serial pay-it-forward incidents involving between 4 and 24 cars have been reported at Wendy's, McDonald's, Starbucks, Del Taco, Taco Bell, KFC and Dunkin' Donuts locations in Maryland, Florida, California, Texas, Louisiana, Pennsylvania, Oklahoma, Georgia, Alabama, North Dakota, Michigan, North Carolina and Washington."

"Pay it forward" is not limited to restaurant drive-ins. Acts of generosity occur commonly in daily life, ranging from anonymous blood donations to stopping to help a stranded motorist. In online communities, voluntary con-

[1] This chapter was co-authored with Michael W. Macy and published under the title "The Social Contagion of Generosity" in *PLoS ONE* (Tsvetkova and Macy, 2014b).

tributions are pervasive: every day, millions of people write restaurant reviews, leave product ratings, provide answers to unknown users' questions, or contribute lines of code to open-source software, all without any direct reward or recognition. Why, in the absence of external sanctions and opportunities for reciprocation, do people help strangers?

One possible explanation is that helping is driven by receiving or observing help. In other words, generosity towards strangers may be socially contagious. In a ground-breaking study, Fowler and Christakis (2010) found evidence that generous behavior can indeed ripple through social networks. In particular, the authors showed that the "three degrees of influence" rule observed for other contagions, such as the spread of happiness and obesity (Christakis and Fowler, 2009), applies as well to generous behavior. If you help someone, you not only increase the likelihood that they help others, but that those they help will also help others, and so on, out to three steps. Using similar experimental designs, Suri and Watts (2011) and Jordan et al. (2013) also found that generous behavior was contagious, but that it does not spread beyond the direct interaction.

The contagiousness of generosity may depend on the mechanism by which it spreads. Fowler and Christakis (2010) and Suri and Watts (2011) tested the spread of generosity on networks but their studies were not designed to identify the underlying mechanisms. They used a public goods experiment in which multiple individuals donate to a common pool and then share the investment equally. Contagion occurs when an individual who has interacted with generous partners in one group donates more in the next group. Although useful in demonstrating contagion, the public-goods experimental design, including the N-person Prisoner's Dilemma (Rand et al., 2011; Jordan et al., 2013), does not

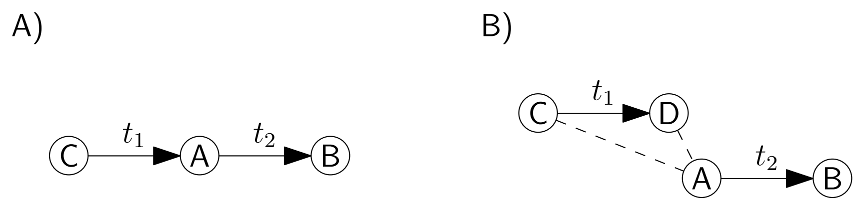


Figure 2.1: **Two mechanisms for the contagion of generosity.** (A) Generalized reciprocity: *A* helps *B* because *C* has helped *A*. (B) Third-party influence: *A* helps *B* because *A* has observed *C* help *D*. Arrows indicate helping or giving, dashed lines indicate observing.

distinguish between receiving and observing generosity since group members also benefit from the generous acts they observe. The present research uses an innovative experimental design to distinguish between the two processes and to measure their contribution to the contagion of generosity.

2.1.1 Generalized reciprocity and third-party influence

Previous research suggests that there are two distinct mechanisms for the social contagion of generosity among strangers: generalized reciprocity and third-party influence. Generalized reciprocity (GR) refers to cases in which those who benefit from the kindness of strangers become more generous towards others in the future. As diagramed in Figure 2.1, *A* helps *B* because *C* has helped *A* (Pfeiffer et al., 2005; Stanca, 2009). Third-party influence (TPI) refers to cases in which those who observe kindness between strangers become more generous towards a stranger: *A* helps *B* because *A* has observed *C* help *D*. GR characterizes "pay it forward" behavior triggered by normative or expressive responses to being helped (Bartlett and DeSteno, 2006), while TPI characterizes social learning through imitation of others' behavior.

The difference in the two mechanisms parallels Deutsch and Gerard's (1955) distinction between normative and informational influence. GR is driven by an "injunctive norm" (Cialdini et al., 1990; Cialdini, 2008) — a normative obligation to express one's gratitude at being helped not by repaying the helper but by acting as the helper acted. TPI is driven by a "descriptive norm" — to follow the example of others' behavior when unsure about how one is expected to act.

GR and TPI also differ in the pattern of transmission. GR transmits the contagion from person to person through direct contact and hence its contagious effect is limited to the one person who was previously helped. In contrast, TPI has the potential to broadcast the contagion from one person to any number of observers. For example, when a stranger stops to help a stranded motorist, only one person receives help but thousands of passersby might observe helping behavior.

This multiplier effect of TPI means that we are far more likely to observe generosity than to receive it. If widespread observation establishes a descriptive norm that in turn makes each individual more likely to be generous, then TPI could generate a powerful self-reinforcing dynamic (Weber and Murnighan, 2008). However, previous research on threshold models of social contagion (Granovetter and Soong, 1983, 1986; Valente, 1996; López-Pintado and Watts, 2008), the "free rider" problem in collective action (Oliver et al., 1985), social loafing in groups (Karau and Williams, 1993), the Volunteer's Dilemma (Diekmann, 1985), the "bystander effect," and the diffusion of responsibility (Darley and Latané, 1968) all point to a very different possibility: that an individual is more likely to help or contribute when confronted with the stark reality that "if you don't do it, nobody else will" (Oliver et al., 1985). Once a descriptive norm

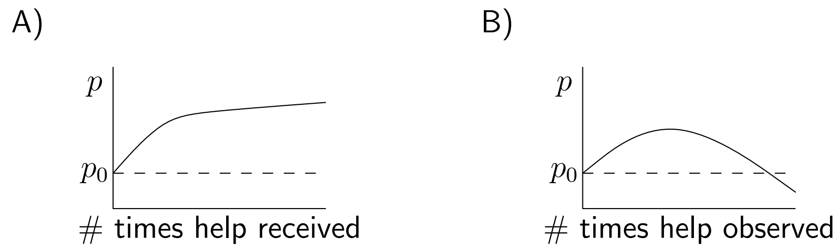


Figure 2.2: **Monotonic and non-monotonic changes in the probability to help.** Both (A) generalized reciprocity and (B) third-party influence are expected to increase the probability to help (p) above the baseline level of "unconditional generosity" (p_0) but the effects from repeatedly receiving and observing help are expected to differ.

has been established and people take for granted that someone else is likely to help, one's own contribution appears less essential. In short, once the observed level of generosity is sufficient to safely assume that one's own contribution is not needed, the positive effect of the descriptive norm can be expected to reverse, such that third-party influence becomes negative (i.e. the observer does the opposite of the observed behavior; see Figure 2.2).

Although conceptually distinct, GR and TPI are not proposed as alternative explanations for the contagion of generosity among strangers. Rather, the two mechanisms are likely to interact, due to the greater likelihood to both receive and observe generosity from strangers in populations where this behavior is normative. When people observe helping behavior after previously receiving help from a stranger, the normative influence from GR is expected to mitigate the negative effects of observing widespread acts of helping.

The present study aims to test GR and TPI as possible mechanisms in the social contagion of generosity. This requires an experimental design in which receiving and observing generosity are not confounded by each other or by the effects of closely related mechanisms. In particular, GR can be confounded by indirect reciprocity and TPI by peer pressure, and both GR and TPI may be confounded by unconditional generosity. In the sections that follow, we elaborate the distinctions, both theoretically and operationally.

2.1.2 Generalized reciprocity vs. indirect reciprocity

Generalized reciprocity should not be confused with indirect reciprocity. Both involve the pattern depicted in Figure 1 in which *A* helps *B* and *C* helps *A*, but they differ in sequencing, and the difference in temporal ordering implies different motivations. With GR, *C* helps *A* before *A* helps *B*, while with indirect reciprocity, *C* helps *A* after *A* helps *B*. GR is more plausibly motivated by feelings of obligation and/or gratitude in response to receiving help, while indirect reciprocity is generally assumed to be instrumentally motivated as a reputational strategy for obtaining help (Nowak and Sigmund, 2005; Bartlett and DeSteno, 2006; Nowak and Roch, 2007).

Generalized reciprocity also differs from generalized exchange (Ekeh, 1974). The latter refers to a pattern of exchange between two members of a group, both of whom give and receive from a group member but not necessarily one another. By that definition, both GR and indirect reciprocity can be classified as two different forms of generalized exchange.

Prosocial behavior could increase when reciprocity is generalized as well as when it is indirect, but only the former leads to social contagion through transmission upon contact. With GR, the helping behavior is backward-looking — a response to the helping behavior of others. In contrast, when reciprocity is indirect, the helping behavior is forward-looking, in anticipation of the receipt of help. Indirect reciprocity could increase generous behavior because it changes the interaction situation by modifying the incentives. GR could increase generous behavior because generosity generates more generosity.

Unfortunately, observational studies of generalized exchange cannot distinguish between GR and indirect reciprocity. For example, the three best documented cases of generalized exchange in naturally occurring environments — the Kula trading ring among South Pacific islanders (Malinowski, 1920), the kinship relations among aboriginal tribes (Bearman, 1997), and the support networks of low-income black women (Uehara, 1990) — involve very small communities, in which helping behavior could be motivated by anticipated rewards rather than as a response to being helped. Similarly, generalized-exchange experiments cannot distinguish GR and indirect reciprocity if interactions are repeated in fixed network structures and/or with full information about others' behavior (Yamagishi and Cook, 1993; Greiner and Levati, 2005; Molm et al., 2007; Tsvetkova and Buskens, 2013).

The effects of GR can be isolated from possible confounding effects of indirect reciprocity by keeping interactions anonymous and by preventing anyone else from knowing about an actor's past behavior. For example, Ben-Ner et al. (2004) and Stanca (2009) isolate GR from indirect reciprocity by using anonymous one-shot interactions that remove opportunities for reputation-based rewards.

2.1.3 Third-party influence vs. peer pressure

Like GR, TPI can also be confused with other types of third-party effects. The TPI we refer to corresponds to what Deutsch and Gerard (1955) call "informational influence," in which an actor models an observed behavior. Deutsch and Gerard distinguish this from "normative influence," in which an actor engages in a behavior that is socially approved. When influence is normative, one conforms to others' behavior in order to be liked and accepted. When influence is informational, the actor conforms to a descriptive rather than prescriptive norm. For example, in "rational herding" (Bikhchandani et al., 1992), conformity occurs because one assumes that others know better what the appropriate behavior should be. The two types of influence are associated with different types of social relationships. Normative social influence (or "peer pressure") depends on the desire for social approval from significant others, which in turn is likely to be greater when there is a pre-existing and on-going social relationship, such as that between family members, friends, or colleagues. In contrast, when relationships are novel and/or transient, as when interacting with strangers, dependence on cues from network neighbors may be more important than dependence on social approval.

This distinction between normative and informational influence is therefore important for the study of generosity among strangers. Normative influence is more relevant for the enforcement of pro-social behavior in tight-knit social groups whose members depend on one another for social approval, while informational influence is more relevant for the contagion of generosity among strangers. Most previous studies of social contagion have been observations of cascades passing through pre-existing social relationships between people who already knew one another (Christakis and Fowler, 2009; Bond et al., 2012). These situations are not well-suited for the study of informational influence, which is likely to be obscured and confounded by normative pressures. The effects of informational influence can be isolated from possible confounding effects of peer pressure by keeping all actors anonymous and precluding repeated local interactions. For example, Salganik et al. (2006) succeed in detecting informational influence in a cultural market by letting participants interact a single time and by revealing to them only the aggregated behavior of others.

2.1.4 Unconditional generosity as baseline

In addition to distinguishing GR from indirect reciprocity and TPI from normative influence, it is crucial to also distinguish both GR and TPI from another important and possibly confounding mechanism — unconditional generosity. Unlike GR and TPI, unconditional generosity occurs when *A* helps *B* even though *A* has not received help from *C* nor observed *C* helping *D*. Thus, when *A* helps *B* after receiving help from *C*, it is possible that *A* would have helped *B* anyway.

This possibility was overlooked by two previous studies of GR (Ben-Ner et al., 2004; Stanca, 2009). These studies offer evidence that individuals who have been recent recipients of generosity are likely to be similarly generous to a third party, even if they know that they cannot benefit from this in the future. However, it is unclear whether participants would have made a similar donation even if they had not received a donation from a stranger. In other words, the observed generosity could have been due to unconditional generosity, rather than the result of contagion through GR.

The effects of GR and TPI can be isolated from possible confounding effects of unconditional generosity by measuring the effect of receiving and observing generosity above and beyond a baseline tendency to help under an otherwise identical decision situation but in which help is neither received nor observed. Similarly, the effects of GR can be isolated from TPI by measuring the effect of receiving help among those who are unable to observe helping behavior more generally. These conditions rarely obtain in natural settings, which limits the ability to identify the underlying mechanisms in observational studies of helping behavior. We therefore designed and conducted an experiment with human participants.

2.2 MATERIALS AND METHODS

2.2.1 Procedure

Subjects were recruited from and paid through the online crowdsourcing platform Amazon Mechanical Turk (see Chapter A) but interacted on a website hosted on our webserver. The study was designed as a sequential two-player investment/gift-exchange game in groups of 150 with random partner selection. In the game, a participant could choose to return part of their payment so that another anonymous participant could benefit (similarly to Greiner and Levati (2005) and Stanca (2009)).

The study was conducted in March–April, 2013. We first recruited a pool of potential participants by posting a task on the online crowdsourcing platform Amazon Mechanical Turk (AMT). The task was called "Sign up to participate in the Invitation Game" and paid $0.20 when submitted. The task invited AMT users to sign up for a study that offered the chance to earn up to $14–21 for doing the same $2–3 ten-minute task multiple times. To sign up, an AMT user simply needed to read and agree to the terms of the study and provide standard demographic information (gender, age, ethnicity, nationality, education, religious affiliation, and income). The instructions emphasized that the demographic information would not be used for selecting the participants. The AMT users were informed that they could only participate in the task and earn the promised amount if they were randomly selected from the pool of potential participants. Participants were eligible to be selected multiple times but there was no guarantee that they would be selected even once. If selected, the participant was to receive an e-mail notification with further instructions. (See the recruitment instructions in section B.1.)

The email invitation informed recipients that they were randomly chosen to participate in the Invitation Game, which they had to complete within 24 hours. Participants were given their AMT worker ID and a unique randomly generated Invitation ID to log into our website. On the website, participants read a description of the Invitation Game, answered five multiple-choice questions testing their understanding of the game rules, wrote a short summary of the decision situation they were facing, and made a single decision about whether to donate money to benefit a stranger (see section B.1).

The game description explained to each participant that they would be paid the amount promised in the original solicitation, which included a "base" payment plus a "bonus" payment. Participants were also told that they were part of a group of 150 AMT users and that only members of this group who received an invitation could actually participate and receive the promised payment. The instructions further informed participants that the study had allocated a limited number of invitations to be distributed to randomly selected participants, whom we will here call "seeds." The seeds were invited to participate by the experimenters. In addition to these invitations created by the experimenters, each participant who received and accepted an invitation had the option to create a new invitation and allow one more person than otherwise to participate. However, in order to create a new invitation, the participant had to be willing to donate his or her bonus, even though this would reduce the participant's earnings which would then be limited to just the base payment. If the participant chose to donate his or her bonus, a recipient of the new invitation (called "invitee") would then be randomly selected from the other 149 AMT users in the group. The instructions explained further that when a participant donated his or her bonus, we supplemented the bonus amount so that the next invited participant received the same base payment and bonus and had the same options: to keep his or her bonus or donate it and create a new invitation for one more participant.

The instructions were identical for seeds and invitees, with one exception. Unlike those invited by the experimenters (i.e. the seeds), the recipients of participant-generated invitations (i.e. the invitees) were informed that they were given the opportunity to complete the task because another participant had donated his or her bonus (referred to hereafter as "donated invitations"). This one sentence is the only difference in the treatment received by seeds and invitees and provides a very conservative test of the effects of receiving help, given that participants in both treatment conditions received invitations, with the only difference being the source of the invitation and no difference in the size of the bonus that accompanied the invitation. Information is all that was manipulated; there was no difference in the amount of money received.

All participants knew that the person who received the donated invitation would not know the identity of the participant who made the donation. Thus, anyone receiving a donated invitation was unable to directly reciprocate or to pass along a favorable reputation. We referred to participants by their AMT worker ID, randomly anonymized in a way that precluded the possibility to identify the same individual and be influenced by reputation. We used anonymized identifiers to refer to the other participants in order to dispel any suspicion of deception and to make the information more prominent and compelling. (The detailed instructions used in the study are included in Chapter B, Section B.1.)

2.2.2 Treatments

The experiment involved five manipulations:

- Whether the participant received a donated invitation created by another participant. Some participants were only selected as seeds while others were only selected as invitees. Still other participants were selected as invitees after having been previously selected as seeds. (Previous invitees were ineligible to be selected as seeds since this violated the concept of a seed as the first mover in a sequential decision process.) Invitees were explicitly informed that they were given the opportunity to complete the task because another participant had donated his or her bonus and created the invitation they received.

- The number of times the participant was invited to play the game, either as a seed or invitee. Participants were randomly selected to take part in the game (as a seed or invitee) between one and six times.
- Whether the participant was able to observe donated invitations. In the observation condition, both seeds and invitees were informed about the number of donated invitations that had been created by other participants in their group up to that point in time and saw a list of the pairs of givers and recipients. Participants were permanently assigned to either the observation or no-observation condition; otherwise the effects of observation would carry over to affect behavior in the no-observation condition as well.
- The number of donated invitations the participant observed. Participants in the observation condition were randomly selected to observe different numbers of invitations donated by the members of their group, ranging from zero to 223 observed invitations. Since the number of invitations created by other participants could stay the same or increase, a participant who interacted multiple times in the observation treatment could observe only a higher number of donated invitations in subsequent interactions. Participants could see the total number of donated invitations as well as a list of donors and invitees (with the AMT worker IDs modified to preserve anonymity). Alternatively, we could have displayed the number of members who had chosen to donate, but this would understate members' level of effort since it would not reflect multiple donations.
- The payment the participant received. Previous research on prosocial behavior has shown that the willingness to donate depends in part on the resources that are available (Oliver et al., 1985). We manipulated the pay-

ment in order to measure the robustness of the results across different incentives to return the bonus. In the high payment treatment, participants received \$2 base rate and \$1 bonus and in the low payment treatment, they received \$1 base rate and \$1 bonus. Participants were permanently assigned to either the high or low payment condition.

The two between-individual manipulations, observation: yes/no and payment: high/low were crossed to define four between-individual treatment groups to which participants were randomly assigned. The number of invitations received and observed varied within individual. The number of seeds and invitees varied across treatment groups due to differences across treatments in the rate at which participants were willing to donate (Table B.1).

2.3 RESULTS

A total of 573 AMT users participated in the experiment, with a mean number of interactions of 2.1 (ranging from 1 to 6), for a total of 1,196 observations. For the analyses, we removed data from 55 participants (126 observations) who required more than five attempts to answer the five multiple-choice questions correctly or whose written summaries revealed an apparent lack of understanding of the instructions.[2] This left 518 individuals and 1,070 observations, with between 1 and 6 observations per individual (mean of 2.1 and median of 2 observations).

[2]The results do not change qualitatively if we include participants who required fewer attempts to correctly answer the questions. The results are also qualitatively similar if we use all observations. See Table B.2.

Participants had a mean age of 30.0 (ranging from 17 to 70),[3] were 38.8% female, with a median household income of $40,000–49,999. The sample consisted of 91.3% US citizens and 6.0% Indian citizens, the remaining being from other countries. The most common ethnicities were 72.2% White and 13.7% Asian. 29.3% reported being non-religious and 25.5% atheists, while Christianity was the most common religion (10.4% Protestant, 9.9% Roman Catholic, and 12.4% other Christian). 12.9% reported educational attainment of high school or less, 42.3% some college or Associate's degree, 35.5% Bachelor's degree, and 9.3% graduate degree. (For detailed demographics of the sample, see Table B.3.)

In 68.1% of all interactions, participants chose to donate their bonus and thereby create an invitation for a stranger at personal expense (62.0% in the low-payment condition and 74.1% in the high-payment condition). Subjects were also relatively consistent in their behavior — out of the 327 individuals who interacted more than once, only 47 varied their decision.

We used random-intercepts logistic regression models of observations nested in individuals to estimate the change in the odds of donating under the different manipulations. The models allow us to adjust for the non-independence of repeated measures and control for the effect of payment level and two other potential confounders — the time elapsed between subsequent interactions and the number of previous interactions, both of which differed between seeds and invitees since invitees on average interacted with greater frequency compared to seeds. To better isolate the mechanisms, the models pool data only form the relevant treatment conditions: we test GR in the no-observation condition only, we test TPI for seeds only, and we test the inter-

[3]Amazon does not allow minors to create and maintain AMT accounts, so the two individuals who reported age under 18 must have either reported incorrect information or used an adult's AMT account

action of GR and TPI in the observation condition only. We report odds ratios which have a more intuitive interpretation than logistic coefficients. It is important to note that the baseline condition in which participants neither receive nor observe donated invitations does not completely isolate unconditional generosity as a mechanism because returning one's bonus and creating an invitation slightly increases one's chance to be invited again and hence, could be strategically motivated. Future research could address this possible confound by manipulating group size, but the focus in the present study is on isolating the effects of GR and TPI, which are not confounded by strategic motivation since the possibility to be re-invited is exactly the same for seeds and invitees.

We tested GR by manipulating whether participants in the no-observation condition were seeds or invitees and also the number of donated invitations they received. Few participants received more than two invitations; hence we binned these as two or more. The results are limited to the no-observation condition ($N = 516$) to avoid confounding the effects of receiving and observing invitations (since the more invitations that other participants have previously sent, the higher the number of invitations that can be observed as well as received).

Consistent with GR, Table 2.1A reveals a seven-fold increase in the odds of donating ($p = 0.030$) among invitees compared to the baseline odds for seeds. Although statistically significant, the change in behavior was relatively small, as evident in Figure 2.3, which reports the change in the fraction donating (rather than the odds), and only within individuals (Table B.5). This small effect size may reflect the minimal GR stimulus, which consisted of a single short statement informing invitees that their invitation was created by another participant

Table 2.1: **Odds ratios for donating across treatments.**

Manipulation	**A) GR**	**B) TPI+**	**C) TPI−**	**D) GR × TPI**
Invitee	7.006*			0.327
	(0.030)			(0.262)
Has been invitee	0.712			1.021
	(0.686)			(0.982)
Seeds				
Observes 0–75		11.414*	(baseline)	(baseline)
		(0.043)		
Observes 76–150		1.341	0.047	0.136
		(0.787)	(0.215)	(0.101)
Observes 151+		0.219	0.003	0.015*
		(0.280)	(0.198)	(0.022)
Invitees				
Observes 76–150				19.907*
				(0.041)
Observes 151+				89.948*
				(0.026)
High payment	64.103**	2.532	0.858	3.235
	(0.007)	(0.300)	(0.930)	(0.295)
Time waited (in hours)	0.972*	0.992	1.019	0.976
	(0.023)	(0.577)	(0.619)	(0.075)
Previous participations	0.690	0.784	1.347	0.454
	(0.379)	(0.622)	(0.848)	(0.171)
Baseline odds	4.305	5.323	152.785	268.707***
	(0.181)	(0.100)	(0.130)	(0.000)
Number of observations	516	371	175	554
Number of participants	252	277	133	266
Wald χ^2	5 df, 11.93*	6 df, 6.66	5 df, 2.49	9 df, 11.98
	(0.036)	(0.354)	(0.778)	(0.214)

Two-sided tests: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
The table reports odds ratios and p-values (in brackets) from random-intercept logistic regression models for A) seeds and invitees in the no-observation treatment by number of donated invitations received; B) seeds in the observation and no-observation treatments by number of donated invitations observed; C) seeds in the observation treatment by number of donated invitations observed; and D) seeds and invitees in the observation treatment by number of donated invitations observed by invitees compared to seeds. Results show that receiving and observing donations initially increases the willingness to help others, and that invitees are less susceptible to a subsequent decline in helping.

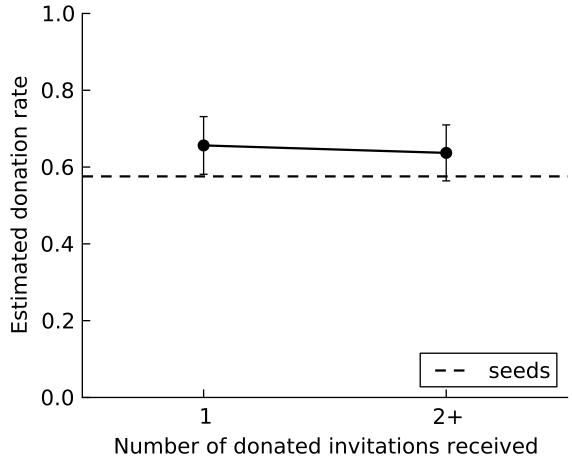


Figure 2.3: **The effect of generalized reciprocity on the willingness to donate in the no-observation condition.** To facilitate interpretation of the odds ratios, the figure shows the estimated donation rate and 95% confidence intervals based on a random-intercept linear regression model with robust standard errors corresponding to the random-intercept logistic model in Table 2.1. The robust standard errors adjust for possible heteroskedasticity with a binary dependent measure. The dashed line shows the baseline donation rate among seeds in the no-observation condition. The donation rate is significantly higher among invitees than among seeds after receiving one donated invitation but does not continue to increase with receipt of additional invitations.

who had donated his or her bonus to make that possible.

In sum, participants were more likely to be generous towards a stranger after experiencing generosity. However, the effect is limited to the first receipt of generosity as the critical event in triggering GR. The odds of donating do not continue to increase but instead slightly decrease with receiving additional donated invitations. A plausible explanation is that participants may feel they fulfilled their normative obligation to "pay it forward" when they donated their bonus after their first donated invitation.

We tested TPI by manipulating whether participants observed invitations created by others and the number of donated invitations they observed. Due to

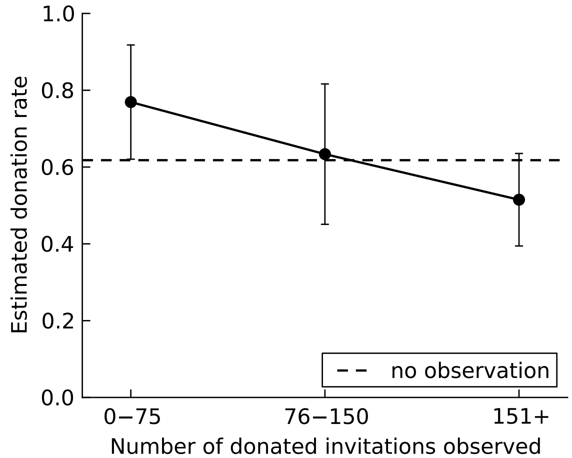


Figure 2.4: **The effect of third-party influence on the willingness to donate among seeds.** To facilitate interpretation of the odds ratios, the figure shows the estimated donation rate and 95% confidence intervals based on a random-intercept linear regression model with robust standard errors corresponding to the random-intercept logistic model in Table 2.1B. The robust standard errors adjust for possible heteroskedasticity with a binary dependent measure. The dashed line shows the baseline donation rate among seeds in the no-observation condition. The donation rate is significantly higher after observing 0–75 donations by other group members but then declines as the level of observed donation increases further. However, the decline is within the confidence intervals of the estimated donation rates, consistent with the results in Table 2.1C (in which the donation rates are compared across levels of observed donation).

the sparsity of data with 223 levels of observed donation and 266 participants, we binned the number of observed donated invitations into three levels: 0–75 (up to about one-third the total number of donations), 76–150 (between one-third and two-thirds), and 151+ (more than two-thirds). Consistent with the expected effects of TPI, Table 2.1B and Figure 2.4 show a statistically significant increase in the odds of donating (OR = 11.41, p = 0.043) among the seeds who had observed between 0 and 75 donated invitations, compared to those who had not observed any. However, the level of donation among those who observed more than 75 invitations was not significantly greater than the baseline level.

This is also consistent with the results for a model that directly tests for changes in the level of donation among seeds as the number of observed donations increases (reported in Table 2.1C). Here the baseline is the lowest level of observed donation instead of the no-observation condition. Although the direction of the effect is as predicted, the decrease in the probability of donation as the number of observed donations increases is not statistically significant (OR = 0.047, $p = 0.215$ for observing 76–150; OR = 0.003, $p = 0.198$ for observing 151+; χ^2 (1 df) = 1.08, $p = 0.298$ for the difference between observing 76–150 and observing 151+). Similarly to GR, the effect of TPI appears to be non-linear, with most of the effect evident at relatively low levels of observed donation and little subsequent change.[4]

However, as the theory of GR suggests, the effect from observing widespread generosity is significantly different for those who have recently benefited from generosity compared to those who have not. When observing more than 75 donated invitations, the odds of donating decrease for seeds but do not change for invitees (Table 2.1D and Figure 2.5). This difference in the odds-ratios between seeds and invitees is statistically significant (χ^2 (1 df) = 3.88, $p = 0.049$ for observing 76–150; χ^2 (1 df) = 5.55, $p = 0.019$ for observing 151+) and suggests the possibility that seeds eventually succumb to a "bystander" (or "free rider") effect from which invitees are immune due to having been recipients of generosity. This apparent immunity suggests that an injunctive norm to "pay it forward" does not diminish when the level of helping behavior is high, while a descriptive norm to "be generous if that is what others are doing" is less resistant to the temptation to "let George do it" as the opportunity to do so increases.

[4]The conclusion does not change with more fine-grained categories. The rate of donation decreases (albeit not significantly) as the number of observed donations increases from 0–25 to 26–50 to 51–75.

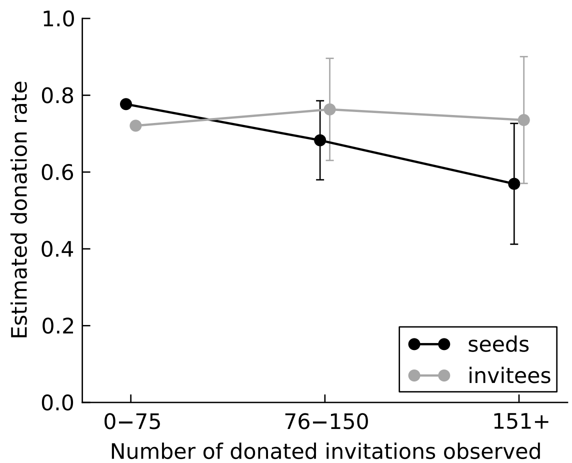


Figure 2.5: **The effect of third-party influence on the willingness to donate among seeds and invitees.** To facilitate interpretation of the odds ratios, the figure shows the estimated donation rate and 95% confidence intervals based on a random-intercept linear regression model with robust standard errors corresponding to the random-intercept logistic model in Table 2.1D. The robust standard errors adjust for possible heteroskedasticity with a binary dependent measure. Relative to the 0–75 baseline, the donation rate declines with the level of observed donation among seeds but not among invitees.

Finally, our analyses also show that the odds of donating are larger in the high-payment condition, especially among seeds in the no-observation condition, as shown in Table 2.1A. Nevertheless, the effects of GR and TPI do not significantly vary by payment (Table B.6 and Table B.7). There was no significant change in the odds of donating with the wait time between invitations or with the number of times one has previously interacted. (We also tested the effect of demographic variables on the odds of donating and apart from a positive effect from age, demographics do not affect generosity, as reported in Table B.4.)

2.4 DISCUSSION

Social contagion offers a compelling theoretical explanation for the emergence and spread of generous behavior, especially when directed towards strangers or in large groups where there is a very low probability that generosity will be directly reciprocated. This study investigated two mechanisms that might explain the contagion of generosity — generalized reciprocity and third-party influence. Causal mechanisms are notoriously difficult to observe in natural settings, and controlled diffusion experiments with large groups are highly impractical in traditional laboratory settings. We therefore designed and conducted a large behavioral experiment online. The experiment used anonymity to isolate the effects of the contagion mechanisms from other cooperation-inducing mechanisms, including direct and indirect reciprocity, as well as peer pressure based on reputation effects. The experiment disentangled the effects of receiving and observing generous behavior by manipulating whether participants benefited from the willingness of others to donate their bonus payment, the number of times they benefited, whether participants were informed of the extent of third party donations, and the number of donations they observed. To ensure the robustness of the results across different incentive levels, we also manipulated participants' payments.

The experimental results show that receiving and observing generosity can significantly increase the likelihood to be generous towards a stranger. However, the results are also consistent with the "bystander" hypothesis that the willingness to contribute can be offset by lower perceived need when the level of helping is sufficiently high. This bystander effect is especially evident among those who have not themselves benefited from generosity, suggesting an important difference between injunctive and descriptive norms: once the level of generosity is sufficient to establish a descriptive norm to be generous towards others, further increases in the level of generosity do not strengthen the norm but instead signal that one's own contribution is not needed. However, an injunctive norm to reciprocate generosity by "paying it forward" does not appear to depend on the belief that one's own contribution is needed. Framed by Cialdini's extensive research (Cialdini, 2008), it seems that the need for help alone is not sufficient to motivate generous behavior unless coupled with either an injunctive or descriptive norm, and norms are not sufficient unless coupled with the need for help, especially if the norm is descriptive.

The study contributes to knowledge about prosocial behavior, altruism, and reciprocity by adopting a relational perspective in a research line that has generally focused on individuals responding independently or in aggregates but rarely as nodes of a social network. We also contribute to knowledge about social contagion by investigating the interaction between transmission through direct contact and transmission through third-party influence, two mechanisms that have been usually studied independently in previous contagion research. We advance social science methodology by developing, demonstrating, and evaluating an online platform for studying the diffusion of behavior in large social groups under controlled conditions, something that is not feasible in a traditional laboratory setting.

In addition to a greater insight into the theoretical puzzle of generosity toward strangers (in the absence of clear opportunities for personal gain), the possibility that generous behavior can trigger cascades has important practical applications, including fund-raising efforts for public broadcasting, contributions to online collaborative projects, and creative participation in online content communities. Our empirical findings could inform strategies for more effectively targeting and structuring interventions intended to promote pro-social behavior, generosity, and cooperative ventures in large groups and organizations, with potential use by philanthropists, activists, policy makers, managers, and administrators.

However, it is important to note that although GR and TPI may be able to increase the level of generosity in a community, they may not be sufficient to jump start the emergence of cooperation. In particular, GR has been shown to be unstable as a strategy for the evolution of cooperation (Boyd and Richerson, 1989). Rather, GR is a behavioral pattern that coevolved with cooperation mechanisms such as direct reciprocity, indirect reciprocity, group selection, and spatial structure (Nowak and Roch, 2007; Rand and Nowak, 2013).

Although the experimental design helps disentangle the effects of GR and TPI, a word of caution is in order. While the AMT participants are much more diverse than the college students used in most previous experiments on prosocial behavior, the sample is nevertheless not perfectly representative of the general population. Future research should replicate the study with other populations with different demographic profiles in order to test whether the findings can be generalized to other populations. Ideally, the external validity of the study should be confirmed in a field experiment with stronger manipulations and more meaningful donations. Such field experiment will be also better suited than the online experiment we conducted to gauge the size of the GR and TPI effects and the practicality of possible interventions. Future research could also extend the present study by testing whether egoistic behavior (e.g. stealing or free riding) can also spread as an "anti-social" contagion, through influence (TPI) or "generalized retaliation" (GR). The online experimental platform that we developed for the current project can be improved and easily adapted to study other populations, with different stimuli, and with participants embedded in large social networks.

Another promising direction for further research is to investigate the macro-level effects of GR and TPI. The effects of GR are limited to the one person who is helped, while the effects of TPI can extend to large numbers of people who observe helping behavior. Thus, TPI may be vital in the early stages of a contagion, by multiplying the number of cascades, while GR could be more beneficial in the later stages, by reinforcing a widely held descriptive norm with an emergent injunctive norm. This reinforcement may be essential in offsetting the growing belief that one's own efforts are not needed as more people are observed to help others. Moreover, these dynamics may depend as well on the structure of social networks that limit the horizons for the observation of helping behavior. The implications of network structure for the dynamics of helping cascades driven by GR and TPI are not intuitively obvious, and we expect agent-based models may prove helpful in generating new hypotheses that can then be tested in a new line of research using online experiments.

CHAPTER 3

THE CONTAGION OF ANTISOCIAL BEHAVIOR

ABSTRACT

Previous research has shown that reciprocity can be contagious when there is no option to repay the benefactor and the recipient instead channels repayment toward strangers. In this study, we test whether retaliation can also be contagious. Extending previous work on "paying it forward," we tested two mechanisms for the social contagion of antisocial behavior: generalized reciprocity (a victim of antisocial behavior is more likely to pay it forward) and third-party influence (an observer of antisocial behavior is more likely to emulate it). We used an online experiment with randomized trials to test the two hypothesized mechanisms and their interaction by manipulating the extent to which participants experienced and observed antisocial behavior. We found that people are more likely to harm others if they have been harmed but they are less likely to do so if they observe that others do not harm. [1]

3.1 INTRODUCTION

The health regime we follow (Centola, 2010), the music we listen to (Salganik et al., 2006), the new technologies we adopt (Rogers, 2003), the news stories we read (Muchnik et al., 2013), and even the likelihood that we vote in an election (Bond et al., 2012) are all to a large degree influenced by our friends and peers.

[1]This chapter was co-authored with Michael W. Macy and published as "The Social Contagion of Antisocial Behavior" in *Sociological Science* (Tsvetkova and Macy, 2015).

Many human behaviors spread through social contact, including some that are often assumed to be acquired independently, such as obesity and fertility (Christakis and Fowler, 2009).

Prosocial behavior has also been shown to be contagious. Fowler and Christakis (2010) found experimental evidence that if you help someone, you not only increase the likelihood that they help others, but that those they help will also help others, and so on, out to three steps. Suri and Watts (2011) and Jordan et al. (2013) similarly found that generous behavior is contagious, but that it does not spread beyond the direct interaction. Tsvetkova and Macy (2014b) identified two mechanisms through which prosocial behavior spreads: both those who benefit from prosocial behavior and those who observe prosocial behavior are more likely to behave prosocially towards strangers. The observation-based mechanism has additionally been confirmed in field experiments: witnessing one kind of prosocial behavior, such as picking up litter, increases the chance for another, such as picking up a fallen letter, picking up a fallen bicycle, or helping a stranger pick up dropped fruit from the ground (Keizer et al., 2013).

Prosocial behavior confers a benefit to others that outweighs any personal cost. Previous research (Tsvetkova and Macy, 2014a,b) has shown that beneficiaries of prosocial behavior become more willing to pay these costs and less vulnerable to the "bystander effect." As a consequence, prosocial behavior can ripple through networks.

In contrast, antisocial behavior imposes costs on others that outweigh any personal benefit. Previous studies have shown that observation of socially irresponsible behaviors like littering (Cialdini, 2008) and graffiti (Keizer et al., 2008) can weaken the protective effects of social norms. If antisocial behavior is also contagious, a single act of misbehavior has the potential to trigger a chain reaction that reaches far beyond the original victims. The present study extends research on antisocial contagion by investigating the underlying causal mechanisms. In parallel with previous research on the diffusion of prosocial behavior (Tsvetkova and Macy, 2014b), we investigate two possible mechanisms: generalized reciprocity and third-party influence. In particular, we focus on a definition of generalized reciprocity as pay-it-forward behavior that is distinct from indirect reciprocity and peer pressure. Using data from an online experiment with human subjects, we find that harming others increases after being harmed but decreases after observing low levels of antisocial behavior. In the discussion, we compare the contagion dynamics of prosocial and antisocial behavior and note important similarities and differences. We also discuss the practical implications of the findings for activists, policy makers, managers, and administrators tasked with developing effective strategies for reducing the incidence and normative acceptance of antisocial behavior.

3.1.1 The contagion of antisocial behavior

Criminologists and scholars of deviance were among the first to argue that antisocial behavior can be contagious, based on observational data showing that violent crime tends to be individually reciprocated, collectively escalating, and spatially clustered (Loftin, 1986). Researchers hypothesized that violent behavior can be transmitted between generations when individuals have been maltreated as children (Widom, 1989) and within generations when criminal behavior is socially learned through communication and interaction (Sutherland and Luckenbill, 1992) or when people who witness or experience violence retaliate or pre-empt it (Loftin, 1986). While theories of social contagion are consistent with observed patterns, the correlational data cannot rule out alternative mechanisms such as a common external cause (Fagan et al., 2007) or heredity (Jones and Jones, 2000). A comprehensive review of observational studies shows that the empirical support for the interpersonal transmission of violent crime is contradictory and problematic (Widom, 1989).

The empirical evidence for the contagion of antisocial behavior is more consistent for the spread of non-violent crime, mainly coming from field tests of the "broken windows" theory (Wilson and Kelling, 1982). The hypothesis is that minor infractions such as graffiti, litter, or broken windows can signal the absence of monitoring, enforcement, and public support of laws and social norms, leading to a self-reinforcing downward spiral (Keizer et al., 2008). Thus, if people notice that others litter, they become more likely to litter themselves (Cialdini et al., 1990).

In addition to these field studies, antisocial contagion has also been studied using controlled experiments with monetary incentives. Falk and Fischbacher (2002) used a four-person stealing game to show that if people know that other members of their group steal, they become more likely to steal as well. Jordan et al. (2013) found that "defection" is contagious in a 20-person Prisoner Dilemma's game.

Although these studies have advanced knowledge of the contagious properties of antisocial behavior, they share an important limitation: the inability to distinguish between the effects of victimization and observation. As a result, these studies cannot identify the underlying causal mechanism.

3.1.2 Generalized reciprocity and third-party influence

The present study extends previous research by investigating the relative importance and interaction of victimization and observation as mechanisms for the contagion of antisocial behavior among strangers. These two mechanisms — formalized as generalized reciprocity (GR) and third-party influence (TPI) — have been shown to affect the contagion of prosocial behavior (Tsvetkova and Macy, 2014b).

GR occurs when those who benefit from a stranger's prosocial behavior act more prosocially towards others in the future. TPI occurs when those who observe prosocial behavior by strangers behave more prosocially towards strangers. GR characterizes "pay it forward" behavior triggered by an affective response to receiving help (in contrast to indirect reciprocity and generalized exchange; see below), while TPI characterizes social learning through imitation of others' behavior.

Data from an online experiment (Tsvetkova and Macy, 2014b) has shown that receiving help increases one's likelihood to help but observing help does so only if the frequency of helping is sufficiently low. Indeed, helping behavior decreases as helping becomes more widespread. A possible explanation for this decrease is that once a prosocial norm has been established and people take for granted that someone else is likely to help, one's own contribution appears less essential. This effect is similar to the "free rider" problem in collective action, the problem of social loafing in work teams and small groups, and the "bystander effect" or the diffusion of responsibility in crowds. Nevertheless, the decrease in helping behavior does not occur for those who have benefitted from others' help. In other words, GR can offset the free riding temptation when observing widespread prosocial behavior.

The findings suggest the possibility that similar mechanisms might underlie the contagion of antisocial behavior. If so, then we should expect those who are victims of antisocial behavior to "pay it forward" by "retaliating" not against the perpetrator but against innocent others, and by imitating antisocial behavior by others. However, there is also an important asymmetry between prosocial and antisocial behavior. While widespread prosocial behavior can encourage the belief that one's own contributions are not needed, widespread antisocial behavior does not entail a similar "free rider" opportunity. We therefore do not expect the decreasing effects of third-party observation that were found for prosocial behavior. Instead, we expect that if antisocial behavior is rare, one may be less tempted to behave antisocially, while if antisocial behavior is rampant, one may feel licensed to behave more antisocially.

An analogous difference should apply to the positive interaction between GR and TPI that was reported for prosocial behavior. Although GR was found to temper the tendency to free ride when prosocial behavior was sufficiently widespread, the inability to free ride on the antisocial behavior of others means that there is no bystander effect that GR might then be expected to attenuate. On the contrary, if the personal experience of being victimized weakens the normative acceptance of antisocial behavior, then the interaction may be negative rather than positive. Put differently, the directed (dyadic) antisocial act may no longer be taken personally but attributed instead to a societal pattern. Thus, when antisocial behavior is widely observed, victims may feel less motivated to "do unto others."

Generalized reciprocity, indirect reciprocity, and generalized exchange

It is important to note that GR, as that term is used here, should not be confused or equated with what sociologists and social psychologists have previously referred to as "indirect reciprocity" and "generalized exchange" (Ekeh, 1974; Yamagishi and Cook, 1993; Lawler, 2001; Molm et al., 2007; Molm, 2010). In the evolutionary biology and behavioral economics literature, the term "generalized reciprocity" is used to refer to helping behavior in which there is no opportunity to act on or respond to reputational information (Ben-Ner et al., 2004; Nowak and Sigmund, 2005; Bartlett and DeSteno, 2006; Nowak and Roch, 2007; Stanca, 2009). In contrast, indirect reciprocity and generalized exchange involve reputational mechanisms.

Our study is motivated in part by the need to isolate the effects of GR from reputation effects. Although previous research on indirect reciprocity and generalized exchange has focused exclusively on helping behavior, reputational effects can also influence harmful behavior, as a strategic deterrent and social sanction.

Previous research on indirect reciprocity examined the hypothesized effects of strategic considerations regarding one's reputation or "image-score" (Raub and Weesie, 1990; Wedekind and Milinski, 2000; Seinen and Schram, 2006). Experimental tests of indirect reciprocity as a reputation effect typically use a design in which A interacts with B and then C (knowing how A treated B) interacts with A. In contrast, we use an experimental design in which A interacts with B and then B interacts with C, which removes any possibility for reputational effects as an incentive or reward.

Our design also differs from experimental studies of generalized exchange which typically allow repeated interactions in fixed network structures, where each actor receives from one group member but gives to another (Yamagishi and Cook, 1993; Molm et al., 2007; Molm, 2010) and group members often have full information about all pairwise exchanges (Greiner and Levati, 2005; Tsvetkova and Buskens, 2013). This design allows participants to use generalized exchange not only to express feelings of obligation and gratitude in response to receiving help but also as a reputational strategy for obtaining help in the future and for rewarding those who have helped in the past. Similarly, well-known cases of generalized exchange in naturally occurring environments are likely to include manifestations of reputation effects: the Kula trading ring among South Pacific islanders (Malinowski, 1920), the kinship relations among aboriginal tribes (Bearman, 1997), and the support networks of low-income black women (Uehara, 1990).

Our research design also extends previous knowledge by studying GR in large networks, with the possibility for a "pay it forward" cascade of contagious malevolence. Even if antisocial behaviors are highly contagious, cascades are not possible with dyadic reciprocity, indirect reciprocity, or with generalized exchange in small groups with complete networks. In contrast, GR in open networks allows the possibility for ripple effects that magnify the negative effects of harming others, as the mirror image of the ripple effects from helping (Nowak and Roch, 2007; Fowler and Christakis, 2010).

In sum, motivated by the gaps in the literature on broken windows, indirect reciprocity, and generalized exchange, the present study aims to investigate the conditions under which antisocial behavior emerges and spreads in large groups through processes of victimization and observation, in the absence of reputation-based social controls.

3.2 MATERIALS AND METHODS

To test the mechanisms, we conducted an online experiment. The experiment was called "The Bonus Game" and was conducted with 750 participants in March 2014. We used a power analysis (shown in Figure C.14) to confirm that the number of participants provided sufficient statistical power to test the hypotheses. Participants were recruited from and paid through the online crowdsourcing platform Amazon Mechanical Turk (AMT) but interacted on a website hosted on our webserver. The experiment was designed as a sequential game in which participant i's payment was determined by the amount the previous participant $(i - 1)$ took from i and the amount i decides to take from the next participant $(i + 1)$.

We recruited a large number of participants by posting a human intelligence task (HIT) on AMT. The HIT paid a small amount ($0.25) for a ten-minute task, with the opportunity to earn up to $1.25 as an additional bonus. The HIT asked AMT users to read and agree to the terms of the study and go to our website to complete the task. On our website, AMT users first had to provide their AMT worker ID and standard demographic information (gender, age, ethnicity, nationality, education, religious affiliation, and income). The instructions emphasized that the demographic information is not required and that it does not affect one's participation in the game. (For detailed demographics of the sample, see Table C.4.)

Second, AMT users had to read the instructions for the game and answer a short quiz. The game instructions explained to each participant that they are part of a chain of other participants. Each participant on the chain can decide to transfer either $0.00 or $0.50 from the next participant's earnings to themselves. The amount transferred from the next participant is divided by two and added to the participant's bonus. The instructions further emphasized that all participants on the chain face the same decision.

The AMT users then had to answer a five-question quiz that tested their understanding of the game instructions. If they correctly completed the quiz within three attempts, they earned a $0.50 participation fee and proceeded to play the game with the possibility to earn an additional bonus; otherwise, their participation ended and they were paid the base HIT rate.

The continuing participants were next given information about the amount transferred by the previous player and how this affected their earnings. The participants were then asked to make a decision about the amount to transfer from the next participant. Before they could submit their decision, participants were required to write a short summary to demonstrate that they understood the decision they were asked to make. In the observation condition, participants were also presented with information about the level of transfer by previous participants.

We refer to participants at the beginning of the chain as "seeds" and those who follow the seed as "links." The instructions and the decision situation were identical for seeds and links, with one exception. For seeds, a lottery was used to determine the amount to be transferred from their earnings. For links, the amount was determined by the previous participant. We referred to the previous participants by their randomly anonymized AMT worker ID in order to dispel any suspicion of deception and to make the information more prominent and compelling (as detailed in Chapter C).

In short, we implemented two between-individual manipulations:

- Whether the participant's payment was affected by a transfer by another participant (links) or determined by a lottery (seeds). The experiment did not involve deception, and the instructions explained the source and amount of this transfer differently, depending on whether the participant was a seed or a link.
- Whether the participant was able to observe the level of transfers by others (observation condition) or not (no-observation condition). Since the experiment did not involve deception, the level of transfers the participant observed varied depending on the decisions made by previous participants.

The observation condition provided summary information on the percent others who elected to transfer $0.50. In order to guarantee enough variation in the observed transfer levels, we limited the window of observation to a small but meaningful unit in the population — the chain, defined as a sequence of participants who take from the next in line. However, in order for seeds to observe others, it was necessary that participants observe at least one other chain apart from their own.

We manipulated observation levels by varying which chains the participant was able to observe. Participants were randomly assigned to either of two observation groups, in which they observed the transfer rate in chains that were randomly selected with a probability corresponding to the percent choosing to transfer \$0.50. We needed four chains to ensure sufficient variation in the transfer rates that participants observed.

We did not reveal the number of previous participants over which the transfer level was estimated in order to avoid potentially confounding normative behavior (of the average individual) with the effects of knowledge about the length of chains or the current participant's position. (These would also be potentially important treatment conditions in a chain-level analysis, which we leave for future research). Neither did we inform participants about the possible or actual length of chains or about their position in the chain since this information might confound the effects of the observation of transfers. For example, knowing that one is at the end of a long chain conveys information about the normative behavior of others. Because this study did not use chain length as either a treatment or an outcome measure, we were able to limit the length of chains in order to ensure the necessary distribution of seeds and links. (Without any limit, it would be theoretically possible to end up with a single chain, with only one seed and everyone else a link.) We limited the chains to four links.

The no-observation/low-observation/high-observation manipulations defined three treatment groups. A number of chains were started in each of the groups and participants were randomly assigned to one of these chains. If participants happened to be the first on the chain, they were treated as seeds, otherwise as links.

To isolate the effects of GR and TPI, we tested GR in the no-observation condition (Figure 3.1) and we tested TPI among those with no loss (Figure 3.2). In addition, we conducted separate analyses for seeds and links in order to distinguish between the response to losses imposed by chance and those imposed by

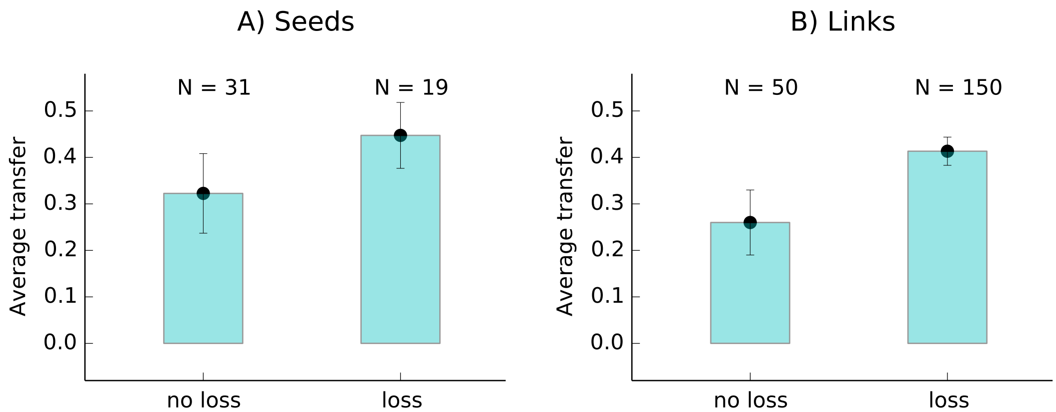


Figure 3.1: **The effect of generalized reciprocity on the amount transferred from the next participant.** The figure shows the mean transfer rate and 95% confidence interval for A) seeds and B) links in the no-observation condition, along with the number of observations in each condition. The figure shows higher transfer rates among those who have experienced a loss but the differences are statistically significant only among links, whose loss was imposed by the actions of another participant.

another participant.

3.3 RESULTS

3.3.1 Generalized reciprocity

GR is evident when the victim recoups a loss by "paying it forward" to someone else. Operationally, we test GR by comparing the responses of those in the loss and no-loss conditions. Figure 3.1 displays the effects of experiencing a loss among links and seeds in the no-observation condition. The results show that the willingness to impose a loss on others is greater among victims of a loss, but the difference is statistically significant only among links. The logistic re-

Table 3.1: **The effect of generalized reciprocity on the amount transferred from the next participant.**

	A) Seeds	B) Links
Loss	1.542	1.482***
	(0.837)	(0.356)
Constant	0.598	0.080
	(0.375)	(0.283)
Number of observations	50	200
LR χ^2	1 df, 4.19*	1 df, 17.36***

Two-sided tests: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
The table reports coefficients and standard errors (in brackets) from logistic regressions for seeds (A) and links (B) in the no-observation treatment. Results show that participants who experience a loss are more likely to impose a loss on others, but the difference is statistically significant only among links.

gressions in Table 3.1 confirm that the effect of loss is marginally non-significant for seeds ($p = 0.065$). However, the size of the effect is very similar for seeds ($b = 1.542$) and links ($b = 1.482$). Thus, the much larger standard error among seeds is likely due to the smaller number of observations.[2] In short, links "did unto others as others had done unto them," but we cannot rule out the possibility that even those whose loss was imposed by chance also become more willing to pass that loss along to others as well.

3.3.2 Third-party influence

We tested the TPI hypothesis by comparing transfers in the no-observation condition (in which participants had no information about others' behavior) with two observation conditions, one with a low level of transfer by others and one with a high. This design distinguishes the effects of uncertainty (no informa-

[2] The similarity in effect sizes for links and seeds is confirmed by the absence of a significant interaction between being a link and experiencing a loss, as reported in Table C.3.

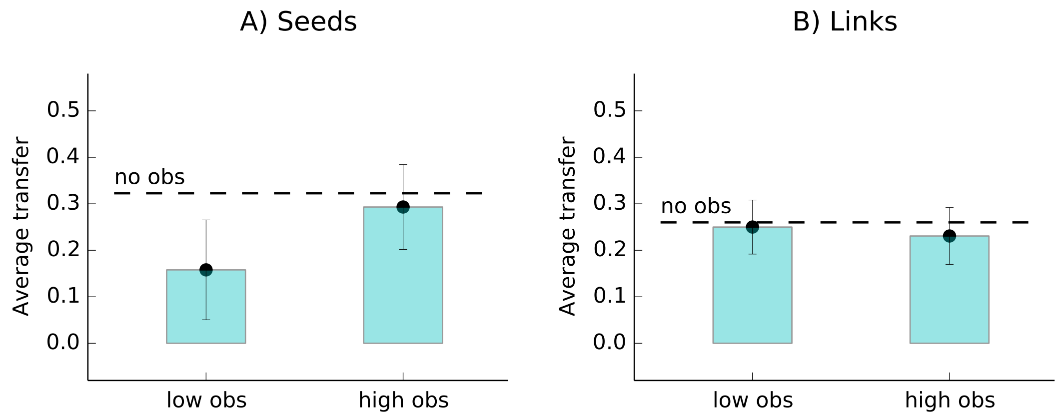


Figure 3.2: **The effect of third-party influence on the amount transferred from the next participant.** The figure shows the mean transfer rate and 95% confidence interval for A) seeds and B) links with no loss. Compared to seeds in the no-observation treatment condition, seeds who observed low transfer levels had a significantly lower transfer rate but seeds who observed high transfer levels did not have higher transfer rates. This effect of observing the behavior of others is not evident among links.

Table 3.2: **The effect of third-party influence on the amount transferred from the next participant.**

	A) No loss		**B) Loss**	
	Seeds	Links	Seeds	Links
Low observation	−1.371*	−0.080	−0.908	−0.378
	(0.620)	(0.368)	(0.862)	(0.300)
High observation	−0.250	−0.234	−1.224	−0.266
	(0.532)	(0.377)	(0.890)	(0.301)
Constant	0.598	0.080	2.140**	1.562***
	(0.375)	(0.283)	(0.748)	(0.216)
Number of observations	79	187	71	413
LR χ^2	5.54	0.42	2.19	1.69

Two-sided tests: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
The table reports coefficients and standard errors (in brackets) from logistic regressions for seeds and links without a loss (A) and with a loss (B). Results show that observing a low level of transfers reduces the amount transferred but only among seeds who experienced no loss.

tion) from the effects of observing different normative behaviors. To isolate TPI from GR, we measured the effects of observation only among participants who did not experience a loss directly. Figure 3.2 shows the level of transfers among seeds and links who did not experience a loss, broken down by the level of transfer observed. The dotted line shows the level of transfers among participants in the no observation condition. Compared to participants in the no-observation condition, those who are aware of the transfer behavior of others are slightly less likely to impose a loss on others, but the difference is significant only among seeds who observe a low level of transfers. This effect is confirmed by the corresponding logistic regressions reported in Table 3.2. The results also show a marginally non-significant difference between the low- and high-observation conditions (χ^2 = 3.26, p = 0.071). The similarity in transfer behavior between the no-observation and high-observation conditions contradicts the "broken windows" hypothesis (Cialdini et al., 1990; Falk and Fischbacher, 2002; Keizer et al., 2008; Jordan et al., 2013), while the low transfer rate in the low-observation condition points instead to the positive effect of "unbroken windows," but only among seeds — those who were not themselves vulnerable to having their own windows broken. [3]

3.4 DISCUSSION

Social contagion offers a compelling theoretical explanation for the spread of prosocial behavior (Tsvetkova and Macy, 2014b). This study extends the investigation to the spread of antisocial behavior, using a controlled online exper-

[3]As with Table 3.1, the separate models for seeds and links who experience low and high loss can be aggregated as a model with two- and three-way interactions, none of which were statistically significant.

iment to test the independent effects of two contagion mechanisms — generalized reciprocity and third-party influence. We found similarities as well as important differences between prosocial and antisocial behavior in the effects of GR and TPI. Recipients of gifts as well as losses were more willing to "pay it forward" to a stranger, suggesting that the rule to "do unto others" does not require that the others are the same people who "did unto you." However, while a gift increases prosocial behavior compared to a chance windfall, a loss due to the malevolence of another person does not increase antisocial behavior compared to a loss due to random chance. Simply put, retaliation is not necessarily the flip side of reciprocity. This asymmetry corresponds to findings from studies on "loss aversion" and "attribution error" (Jones and Harris, 1967; Ross, 1977; Tversky and Kahneman, 1991). An unexpected surplus does not in itself make us more generous, but an unexpected loss can induce efforts at recovery, whether or not the loss was caused by a deliberate choice by another person. It is possible that seeds experienced random gains as good fortune but random losses as transfers to the experimenter. We leave the exploration of these possibilities to future research.

Prosocial and antisocial behavior also differ in the response to observing the behavior of others. Compared to not observing, observing low levels of prosocial behavior increases the willingness to help others but as the observed level increases, the "bystander effect" becomes increasingly evident. In contrast, observing low levels of antisocial behavior decreases the likelihood to harm others compared to not observing.

These differences in findings could also reflect an important difference in the design of the prosocial and antisocial experiments. The prosocial study

(Tsvetkova and Macy, 2014b) used repeated interactions, while the present study did not. The prosocial GR effect was largely within-individuals and may have reflected a self-interested response to the opportunity to benefit from the possibility that "what goes around may come around." The one-shot design in the antisocial study removed this possibility. Future research should explore whether repeated interaction weakens antisocial GR by posing the fear that harming others might increase the probability to be harmed in the future.

Future research is also needed to explore other conditions that might strengthen the contagious properties of both prosocial and antisocial behavior. One is to manipulate the possibility for chains to branch, in order to measure the motivating effect of the opportunity for acts of generosity and harm to reverberate exponentially rather than linearly. For example, in the movie *Pay It Forward*, recipients of help were expected to help three other people. In both the experimental versions, recipients of help and harm could only pay it forward to one.

Future research should also replicate the study with participants with different demographic profiles in order to test whether the findings can be generalized beyond a population that was disproportionately young, well-educated, white, American, and male. The tendency to "pay it forward" in both the prosocial and antisocial situations may differ, for example, between collectivist and individualist cultures.

Further, the external validity of the study should be confirmed in a field experiment with stronger manipulations and more consequential incentives, such as online multiplayer games or user-generated content communities that allow for anonymous interactions. These experiments would produce more socially meaningful measures of the effects of GR and TPI on the spread of antisocial behavior and the practicality of possible interventions.

In contrast to the decision situation in our experiment, our everyday interactions are rarely anonymous and are often embedded in social context. In daily life, people indeed sometimes react to negative experiences by displacing their aggression; the expressions “scapegoating” and “shooting the messenger” describe such behavior. In general, however, “pay it forward” aggression is rarely arbitrary but rather targets role equivalents or subordinates. For example, employees abused by their supervisors are likely to be aggressive towards co-workers and subordinates at work or family members at home (Hoobler and Brass, 2006). Among the indigenous American tribe Kwakiutl, mourners revenge the death of a loved one by killing someone whose rank is equivalent to the rank of the dead relative (Fiske and Rai, 2014). Even among primates, low-ranked males who have lost to high-ranked males tend to attack females (Sapolsky, 2006). Our experiment was conducted on random social interactions among anonymous peers but future research should investigate whether antisocial behavior spreads similarly in existing hierarchical and/or clustered social networks.

With these caveats in mind, we conclude by calling attention to possible implications of our results for intervention strategies. For encouraging prosocial behavior, it has been suggested that TPI may be vital in the early stages of a contagion, by multiplying the number of cascades, while GR could be more beneficial in the later stages, by offsetting the growing belief that one’s own efforts are not needed as more people are observed to help others (Tsvetkova and Macy, 2014b). The current results point to similar strategies for discouraging antisocial behavior. While GR can be self-reinforcing in the early stages of a contagion, TPI may prevent the further spread of antisocial behavior by reinforcing the fact that opportunistic behavior is rare.

CHAPTER 4

THE CONTAGION OF PROSOCIAL BEHAVIOR AND THE EMERGENCE OF VOLUNTARY-CONTRIBUTION COMMUNITIES

ABSTRACT

Every day, millions of people write online restaurant reviews, leave product ratings, provide answers to unknown users' questions, or contribute lines of code to open-source software, all without any direct reward or recognition. People help strangers offline as well, as when they anonymously donate blood or stop to help a stranded motorist, but these behaviors are relatively rare compared to the pervasiveness of online communities based on user-generated content. Why are mutual-help communities far more common online than in traditional offline settings that are not mediated by the Internet? We address this puzzle in two steps. We begin with empirical evidence from an online experiment that tests two mechanisms for the contagion of helping behavior — generalized reciprocity and third-party influence. We then use an empirically-calibrated agent-based model to show how these mechanisms interact with the rivalness of contributions, that is, the extent to which the benefit from a contribution is limited to just one beneficiary (as when helping a stranded motorist) or benefits many people at once (as when contributing a product review online). The results suggest that the non-rivalness of most user-generated content on the Web provides a plausible explanation for the rapid diffusion of helping behavior in online communities. [1]

[1]This chapter was co-authored with Michael W. Macy and is currently under peer review for publication in a book on computational approaches to the study of social phenomena.

4.1 INTRODUCTION

Many human behaviors spread through social contact, including prosocial behavior. In a groundbreaking study, Fowler and Christakis (2010) found experimental evidence that if you help someone, you not only increase the likelihood that they help others, but that those they help will also help others, and so on, out to three steps. This finding has provoked new questions. What are the mechanisms through which prosocial behavior spreads among strangers? How do these mechanisms affect the contagion dynamics? Can they lead to the emergence of cooperation in an initially non-cooperating population?

4.1.1 The puzzle of online generosity

The puzzle of contagious generosity is compounded further by the emergence of online communities with user-generated content, from open-source software development to advice forums to Wikipedia (Kollock, 1999). Why are mutual-help communities far more common online than in traditional offline settings that are not mediated by the internet?

We address this puzzle using an empirically calibrated agent-based model. The results suggest that the answer may lie in the differences in the rivalness of online and offline goods involving anonymous contribution. Many offline goods, like blood donation, charities, and giving up one's seat, are rival, meaning that the contribution transfers resources from the giver to a particular receiver. In contrast, many online goods, especially in communities based on user-generated content, are non-rival — everyone in the community can benefit from a given contribution. The difference is not limited to the effect of non-rival incentives on the independent probability of contribution by a member of the community. Computer simulation shows that this "within individual" difference is amplified by the "between individual" effects of the contagion dynamics. More precisely, we identify two mechanisms of contagion — "generalized reciprocity" and "third-party influence" — and show how these mechanisms interact with differences between rival and non-rival contributions to explain the spread of helping behavior in online communities.

4.1.2 Outline of a theory of prosocial contagion

Previous research has suggested that there are two distinct mechanisms for the contagion of prosocial behavior among strangers: generalized reciprocity and third-party influence. Generalized reciprocity (GR) refers to cases in which those who benefit from a stranger's prosocial behavior behave more prosocially towards another in the future. As diagramed in Figure 4.1, *A* helps *B* because *C* has helped *A* (Pfeiffer et al., 2005; Stanca, 2009). Third-party influence (TPI) refers to cases in which those who observe prosocial behavior by strangers behave more prosocially towards a stranger: *A* helps *B* because *A* has seen *C* help *D*. GR characterizes "pay it forward" behavior triggered by a normative or affective response to being helped (Bartlett and DeSteno, 2006), while TPI characterizes social learning through imitation of others' behavior.

Further, GR transmits the contagion from person to person through direct contact and hence its contagious effect is constrained to the chain of those who were previously helped. In contrast, TPI has the potential to broadcast the con-

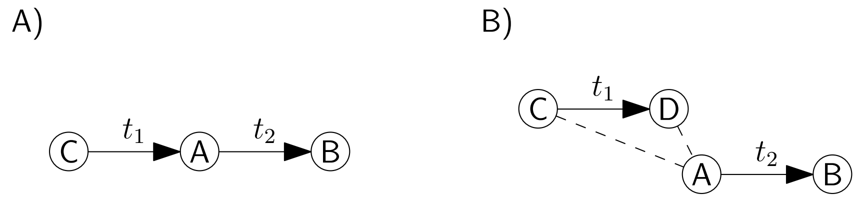


Figure 4.1: **Two mechanisms for the contagion of prosocial behavior.** (A) Generalized reciprocity: A helps B because C has helped A. (B) Third-party influence: A helps B because A has observed C help D.

tagion from one person to any number of observers. The interaction of the two mechanisms could generate a powerful self-reinforcing dynamic that dramatically increases the rate of prosocial behavior in an initially uncooperative population.

In this chapter, we summarize an online experiment that distinguishes between the behavioral effects of the two contagion mechanisms (Chapter 2) and present an agent-based model that investigates the resulting contagion dynamics and population-level outcomes. The empirical results show that receiving help can increase the willingness to be generous towards others, but observing help can have the opposite effect, particularly among those who have not received help. We use a threshold model with dynamic interaction structure and adaptive behavior to simulate a population of agents with this behavior. The computational experiments indicate that the agents can self-organize in communities based on voluntary contributions in two possible ways. On the one hand, when contributions are rival, a handful of altruists can lead to the emergence of small clusters of contributors as long as agents observe contribution beneficiaries in a relatively large radius (for example, via gossip) and unsatisfied agents are not too mobile. On the other hand, when contributions are non-rival, communities are much more likely to emerge and the level of contributions is higher when agents observe contributors rather than recipients. These two pathways roughly correspond to offline and online interactions. They offer explanation for the fact that cultures of kindness are rare for anonymous face-to-face interactions but common on the Web, for example, in the form of communities based on user-generated content.

4.2 TESTING INDIVIDUAL MECHANISMS

Causal mechanisms are notoriously difficult to observe in natural settings, and controlled diffusion experiments with large groups are highly impractical in traditional laboratory settings. To test the two contagion mechanisms, we therefore designed and conducted a large behavioral experiment online. The experiment used anonymity to isolate the effects of GR and TPI from other cooperation-inducing mechanisms, including direct and indirect reciprocity, as well as peer pressure based on reputation effects. To isolate GR from TPI, we manipulated the extent to which participants received and observed help.

4.2.1 Online experiment

The study was designed as a sequential two-player investment/gift-exchange game in groups of 150 with random partner selection. In the game, a participant could choose to return part of their payment so that another anonymous participant could benefit.

We first recruited a pool of potential participants by posting a task on the online crowdsourcing platform Amazon Mechanical Turk (AMT). The task invited

AMT users to sign up for a study that offered the chance to earn up to $14-21 for doing the same $2-3 ten-minute task multiple times. The AMT users were informed that they could only participate in the task and earn the promised amount if they were randomly selected from the pool of potential participants. Participants were eligible to be selected multiple times but there was no guarantee that they would be selected even once. If selected, the participant was to receive an e-mail notification with further instructions.

The email invitation informed recipients that they were randomly chosen to participate in the game, which they had to complete within 24 hours. Participants were then directed to our website, where they read a description of the game and made a single decision about whether to donate money to benefit a stranger. The game description explained to each participant that they would be paid the amount promised in the original solicitation, which included a "base" payment plus a "bonus" payment.

Participants were also told that they were part of a group of 150 AMT users and that only members of this group who received an invitation could actually participate and receive the promised payment. The instructions further informed participants that the study had allocated a limited number of invitations to be distributed to randomly selected participants ("seeds"). The seeds were invited by the experimenters to participate. In addition to these invitations created by the experimenters, each participant who received and accepted an invitation had the option to create a new invitation and allow one more person than otherwise to participate. However, in order to create a new invitation, the participant had to be willing to donate his or her bonus, even though this would reduce the participant's earnings. If the participant chose to donate his or her bonus, a recipient of the new invitation (the "invitee") would then be randomly selected from the other 149 AMT users in the group. The instructions explained further that when a participant donated his or her bonus, we supplemented the bonus amount so that the next invited participant received the same base payment and bonus and had the same options: to keep his or her bonus or donate it and create a new invitation for one more participant.

All participants knew that the person who received the donated invitation would not know the identity of the participant who made the donation. Thus, anyone receiving a donated invitation was unable to directly reciprocate or to pass along a favorable reputation. We referred to participants by their AMT worker ID, randomly anonymized in a way that precluded the possibility to identify the same individual and be influenced by reputation.

The experiment involved five manipulations: whether the participant received a donated invitation created by another participant (i.e. being an "invitee"), the number of times the participant was invited to play the game (ranging from one to six), whether the participant was able to observe donated invitations, the number of donated invitations the participant observed (ranging from zero to 223), and the payment the participant received ($2 base rate and $1 bonus or $1 base rate and $1 bonus).

The observation and payment manipulations were crossed to define four between-individual treatment groups to which participants were randomly assigned. The number of invitations received and observed varied within individuals. Further, some participants were only selected as seeds, others were only selected as invitees, and still others were selected as invitees after having been previously selected as seeds.

4.2.2 Results

After removing data from participants who did not demonstrate an adequate understanding of the instructions, we were left with 518 individuals and 1,070 observations. We used random-intercepts logistic regression models of observations nested in individuals to estimate the change in the odds of donating under the different manipulations. The models allow us to adjust for the non-independence of repeated measures and control for the effect of payment level and two other potential confounders, the time elapsed between subsequent interactions and the number of previous interactions. To better isolate the mechanisms, the models pool data only form the relevant treatment conditions: we test GR in the no-observation condition only, we test TPI for seeds only, and we test the interaction of GR and TPI in the observation condition only.

Consistent with GR, participants were more likely to be generous towards a stranger after experiencing generosity (Table 4.1A). However, the effect is limited to the first receipt of generosity as the critical event in triggering GR. The odds of donating do not continue to increase with receiving additional donated invitations.

Consistent with TPI, there was a statistically significant increase in the odds of donating among seeds who observed between 0 and 75 donated invitations, compared to those who did not observe any (Table 4.1B). However, the level of donation among those who observed more than 75 invitations was not significantly greater than the baseline level. In other words, similarly to GR, the effect of TPI appears to be concave, with most of the effect evident at relatively low levels of observed donation and little subsequent change.

Table 4.1: **Odds ratios for donating across treatments.**

Manipulation	**A) GR**	**B) TPI**	**C) GR × TPI**
Invitee	7.006*		0.327
	(0.030)		(0.262)
Has been invitee	0.712		1.021
	(0.686)		(0.982)
Seeds			
Observes 0–75		11.414*	(baseline)
		(0.043)	
Observes 76–150		1.341	0.136
		(0.787)	(0.101)
Observes 151+		0.219	0.015*
		(0.280)	(0.022)
Invitees			
Observes 76–150			19.907*
			(0.041)
Observes 151+			89.948*
			(0.026)
High payment	64.103**	2.532	3.235
	(0.007)	(0.300)	(0.295)
Time waited (in hours)	0.972*	0.992	0.976
	(0.023)	(0.577)	(0.075)
Previous participations	0.690	0.784	0.454
	(0.379)	(0.622)	(0.171)
Baseline odds	4.305	5.323	268.707***
	(0.181)	(0.100)	(0.000)
Number of observations	516	371	554
Number of participants	252	277	266
Wald χ^2	5 df, 11.93*	6 df, 6.66	9 df, 11.98
	(0.036)	(0.354)	(0.214)

Two-sided tests: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
The table reports odds ratios and p-values (in brackets) from random-intercept logistic regression models for A) seeds and invitees in the no-observation treatment by number of donated invitations received; B) seeds in the observation and no-observation treatments by number of donated invitations observed; and C) seeds and invitees in the observation treatment by number of donated invitations observed by invitees compared to seeds. Results show that receiving and observing donations initially increases the willingness to help others, and that invitees are less susceptible to a subsequent decline in helping.

Less intuitively, the effect from observing widespread generosity is significantly different for those who have recently benefited from generosity compared to those who have not. When observing more than 75 donated invitations, the odds of donating decrease for seeds but do not change for invitees (Table 4.1C). This difference between seeds and invitees is statistically significant ($\chi^2(1df) = 3.88$, $p = 0.049$ for observing 76–150; $\chi^2(1df) = 5.55$, $p = 0.019$ for observing 151+) and suggests the possibility that seeds succumb to a "free riding" effect from which invitees are immune due to having been recipients of generosity. Free riding represents the temptation to refrain from contributions, especially when one becomes aware that others are already contributing. The behavior is common in collective-action situations (Oliver et al., 1985) and is also known as social loafing (Karau and Williams, 1993) and as the "bystander effect" or the "diffusion of responsibility" (Darley and Latané, 1968).

In sum, the experimental results show that receiving and observing generosity can significantly increase the likelihood to be generous towards a stranger. However, the willingness to contribute can be offset by lower perceived need when the level of helping is sufficiently high. This "bystander effect" is especially evident among those who have not themselves benefited from generosity. In other words, the norm to "be generous if that is what others are doing" weakens when the level of helping behavior is high, unless it interacts with the normative obligation to "pay it forward."

The implications of the effects of the two contagion mechanisms for the dynamics of helping cascades are not intuitively obvious. We therefore incorporated the empirical findings in an agent-based model to investigate the macro-level effects of GR and TPI.

4.3 EXTRAPOLATING TO POPULATION OUTCOMES

Our model is a threshold model of collective behavior. Such models have been previously used to study the emergence of collective action and the resolution of social dilemmas (Granovetter, 1978; Macy, 1991; Oliver, 1993). In this literature, a threshold is the critical number or proportion of contributors at which an individual becomes willing to contribute to a collective action or to join a collective behavior. Depending on the distribution of individual thresholds, cascades are possible in which each additional participant triggers participation by others. It has been established that the emergence of widespread participation critically depends on the composition of the population, and in particular, the existence of a critical mass of altruists, or unconditional contributors.

We model diffusion through the dynamics of selection and influence by relaxing two common assumptions in existing threshold models: fixed interaction structure and fixed individual interests in contributing. Our model assumes that agents both move in space (similarly to Schelling (1971)) and adapt their behavior (similarly to Granovetter and Soong (1983), Granovetter and Soong (1986), Macy (1991), and López-Pintado and Watts (2008)). By combining dynamic interaction structure with adaptive behavior, our model is similar to evolutionary-game models on cooperation (Eguíluz et al., 2005; Biely et al., 2007; Hanaki et al., 2007; Helbing and Yu, 2009; Meloni et al., 2009; Fehl et al., 2011). In these models, agents choose an action or a strategy in the Prisoner's Dilemma and play it against each of their interaction neighbors. The agents update their behavior by imitating successful neighbors and find more beneficial interaction partners by moving on a spatial grid or by rewiring their interaction network. In our model, agents play a gift game with a different number of their neighbors, depending on the rivalness of the exchanged gifts. Influence occurs not because agents imitate others but because they condition their behavior on others' behavior.

4.3.1 Agent-based model

Assumptions

The model assumes that agents are heterogeneous with respect to their natural proclivity to condition their contributions on others' behavior and their own outcomes. These proclivities are exogenously predetermined and remain fixed throughout social interactions. In addition to generalized reciprocity, third-party influence, and free riding, the model assumes two other behavioral mechanisms: aspiration and unconditional altruism. Aspiration is the expectation about the extent to which one should benefit from others' contributions. Aspiration is the benchmark against which the agent evaluates outcomes as satisfactory (Macy and Flache, 2002). If outcomes are unsatisfactory, the agent can decide to move to a different community (similarly to Schelling (1971)). We assume that aspiration $\theta_A \sim Uniform\,(0, 0.5)$.

Unconditional altruism captures the extent to which individuals are willing to help strangers regardless of others' behavior or their own outcomes. Following previous research (López-Pintado and Watts, 2008), agents are assigned a level of unconditional altruism that is randomly drawn from a beta distribution: $\theta_{UA} \sim Beta\,(\alpha, \beta)$. The model fixes $\alpha = 5$ and $\beta = 5$. The resulting distribution lacks a critical mass of altruists because the majority of individuals have values close to 0.5. This distribution matches the empirical distribution of behavioral types in the general population, characterized by few unconditional altruists (about 13%) and a majority of conditional contributors (50-63%) (Fischbacher et al., 2001; Kurzban and Houser, 2005). Nevertheless, previous analytical work on deterministic threshold models in fixed populations has shown this type of distribution not to favor the emergence of high-levels of contribution (Granovetter, 1978; López-Pintado and Watts, 2008). Compared to these earlier models, we start from a lower level of unconditional altruism that is more empirically plausible.

The model assumes that generalized reciprocity $GR \sim \mathit{Uniform}\,(0,1)$ and third-party influence $TPI \sim \mathit{Uniform}\,(0,1)$. The higher the value of GR (TPI), the more the agent's contribution behavior is sensitive to benefits received (observed). For consistency, the free riding value is always at least as large as the unconditional-altruism value: $\theta_{FR} = \theta_{UA} + FR \times (1 - \theta_{UA})$, where $FR \sim \mathit{Uniform}\,(0,1)$. The higher the value of FR, the lower the observed level of contribution at which the agent refrains from contributing in order to free ride on others' effort.

The model also assumes that the interaction structure is a square lattice that wraps into a torus. This structure is characterized by a high average clustering, long average path-lengths, and regularity in network positions. The structure is a poor representation for persistent social relations such as friendships and business contacts. However, it is a suitable heuristic for interactions between strangers in geographical space or in a field of topics and interests. Further, we assume that an agent's interaction neighborhood does not entirely coincide with the agent's observation neighborhood. In both cases, the neighborhood is a Moore neighborhood (a square with the focal agent in the center) but the radius of the neighborhood can vary. A larger interaction neighborhood corresponds to a larger community size while a larger observation neighborhood corresponds to a higher degree of gossip or centralized broadcasting.

Behavioral rules

The five behavioral mechanisms come together in two separate threshold functions that determine whether agents contribute to a neighbor (or multiple neighbors) from their interaction neighborhood and whether agents move to a new location in their observation neighborhood.

The contribution threshold models the combined effect from receiving and observing others' contributions on one's likelihood to contribute. As in the empirical results, benefiting from others' contributions increases one's likelihood to contribute and decreases one's likelihood to free ride, while observing others' contributions could increase both one's likelihood to contribute and one's likelihood to free ride. Following previous models of non-monotonic threshold functions (Granovetter and Soong, 1983, 1986; López-Pintado and Watts, 2008), the function is characterized by two thresholds: an upward threshold $\theta_{(0\to1)}$ and a downward threshold $\theta_{(1\to0)}$. The agent contributes as long as the number of received and observed contributions is within these two thresholds. The upward threshold is pre-determined by the agent's unconditional altruism but decreases if the agent experiences third-party influence. The downward threshold is anchored by the agents' proclivity to free ride but increases if the agent succumbs to generalized reciprocity (Figure 4.2). More specifically:

$$\theta_{(0\to1)}(t) = \theta_{UA} - TPI \times M_o(t) \times \theta_{UA} \tag{4.1}$$

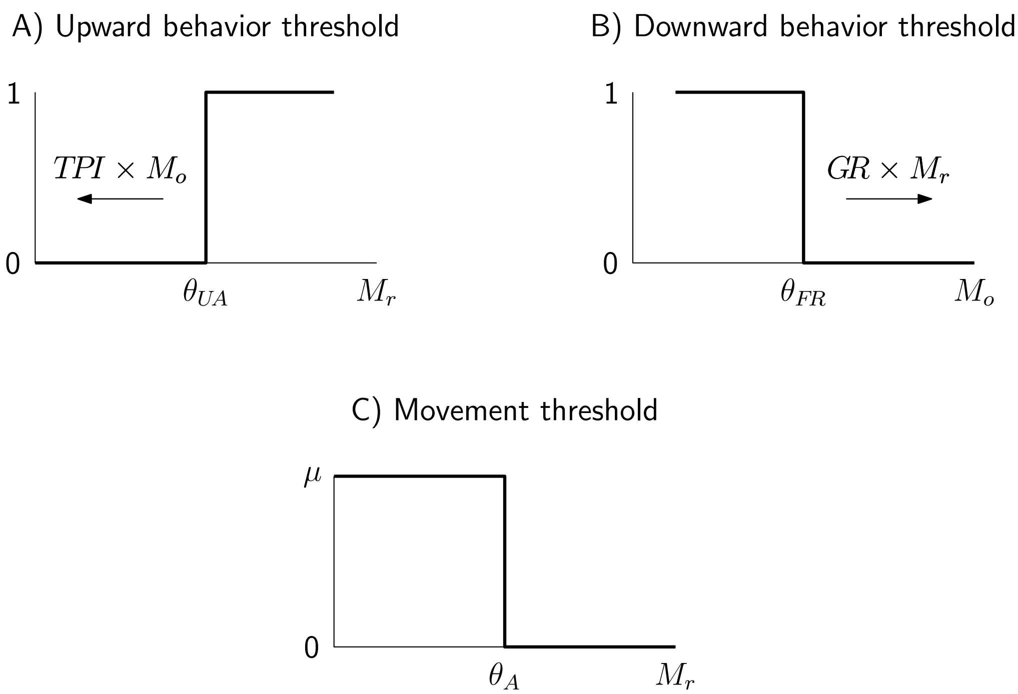


Figure 4.2: **Three thresholds in the simulation model.** (A) The upward behavior threshold depends on unconditional altruism (θ_{UA}) but can decrease due to third-party influence ($TPI \times M_o$). (B) The downward behavior threshold depends on the proclivity to free ride (θ_{FR}) but can increase due to generalized reciprocity ($GR \times M_r$). (C) The movement threshold depends on the aspiration (θ_A). The agent makes a contribution to the benefit of a random neighbor(s) within her interaction neighborhood if the contributions she remembers receiving (M_r) match or surpass her upward threshold but the contributions she remembers observing (M_o) do not exceed her downward threshold. The agent moves to a new empty site within her observation neighborhood if the contributions she remembers receiving (M_r) fall below her aspiration.

and

$$\theta_{(1\to 0)}(t) = \theta_{FR} + GR \times M_r(t) \times (1 - \theta_{FR}), \tag{4.2}$$

where $M_r(t)$ is the number of contributions the agent remembers receiving and $M_o(t)$ is the proportion of contributions the agent remembers observing in her observation neighborhood. The agent makes a contribution to the benefit of a random neighbor(s) within her interaction neighborhood if the contributions she remembers receiving match or surpass her upward threshold but the contributions she remembers observing do not exceed her downward threshold:

- Behavior Rule 1: Contribute if $M_r(t) \geq \theta_{(0\to1)}(t)$ and $M_o(t) \leq \theta_{(1\to0)}(t)$.

Similarly, the agent moves with probability μ (mobility) to a random empty site within her observation neighbourhood if the contributions she remembers receiving do not match her aspiration:

- Behavior Rule 2: Move with probability μ if $M_r \leq \theta_A$.

Thus, agents who are satisfied with their outcomes tend to stick to the community they have found but unhappy agents tend to move to new communities until they find one with a high level of contributions.

$M_r(t)$ and $M_o(t)$ are simply the number of contributions the agent received and the proportion of local contributions the agent observed in the previous m time periods, where m is the length of memory. More formally, $M_r(t) = \frac{\sum_{t-m}^{t-1} r_t}{m}$ and $M_o(t) = \frac{\sum_{t-m}^{t-1} \frac{o_t}{n_t}}{m}$, where r_t is the number of times the agent benefited from a contribution at time t, o_t is the number of contributions the agent observed at time t, and n_t is the size of the agent's neighborhood at time t. For the model, $m = 5$ was chosen because this value produced high variability in the results. Increasing m constricts the conditions for emergence of contributions since more random events become necessary in an agent's neighborhood in order to convert that agent into a contributor.

Updating is synchronous for both the decision to contribute and to move. At each time period, agents are drawn in random order to decide whether to contribute, given the contributions they observed and the amount of contributions they received up until the last period. Once all agents have had the chance to update their behavior, the agents decide whether to move, given the amount of contributions they have received until the end of the current period. Thus, the model assumes that agents observe and receive contributions within each time period and then decide whether to contribute (Behavior Rule 1) and whether to leave a community (Behavior Rule 2). Since threshold models have been shown not to be robust to noise (Macy and Tsvetkova, 2013), the model assumes that there is a small probability $\epsilon = 0.001$ that an agent's contribution or movement decision is reversed.

Parameter space

To preclude sensitivity to initial conditions and synchronous updating, the model used behavioral and movement noise, the simulations were run for a sizeable agent population, and the results were averaged over multiple repetitions. The fixed parameters in the model (the shape and the range of the distributions and the length of memory) were chosen with the goal to keep them as simple as possible while producing the highest variation in results along the variable parameters.

We ran the computational experiments for a population of 1000 agents on a 50×50 torus (40% occupied locations). The experiments investigated the average contribution level (i.e. the proportion of contributors) for two different levels of rivalness: we assumed that rival contributions benefit one recipient, while non-rival contributions benefit 3 recipients. We explored the effects of four parameters:

- The mobility $\mu \in [0, 0.05, 0.5]$. This is the probability to move if the agent is unhappy with what she receives from the current community. This parameter represents community turnover. (Turnover could also be adjusted by varying the average aspiration θ_A.)
- The radius of the interaction neighborhood ∈ [1, 2, 3, 4, 5, 7, 10, 15]. Since the model uses Moore neighborhoods, this is equivalent to a maximum of [8, 24, 48, 80, 120, 224, 440, 960] potential interaction partners for each agent. This parameter corresponds to community size.
- The radius of the observation neighborhood ∈ [0, 1, 2, 3, 4, 5, 7, 10, 15]. Since the model uses Moore neighborhoods, this is equivalent to a maximum of [0, 8, 24, 48, 80, 120, 224, 440, 960] observed neighbors for each agent. This parameter is related to gossip and centralized broadcasting.
- The observation targets ∈ [recipients, contributors]. Agents observe either the proportion of contributors or the proportion of beneficiaries within their observation neighborhood.

The simulations were run for 5000 periods which was sufficient for convergence to an equilibrium. The equilibrium proportion of contributors was then estimated by averaging the proportion of contributors over the last 1000 periods. The resulting equilibrium proportion of contributors was then averaged over 25 replications for each parameter combination.

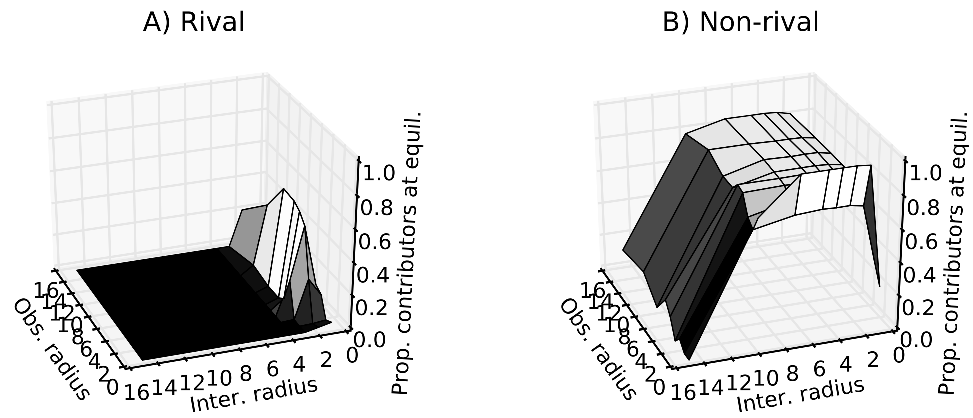


Figure 4.3: **The equilibrium proportion of contributors by observation radius and interaction radius for contributors as observation target and no mobility (μ = 0).** Results are shown for (A) rival and (B) non-rival contributions.

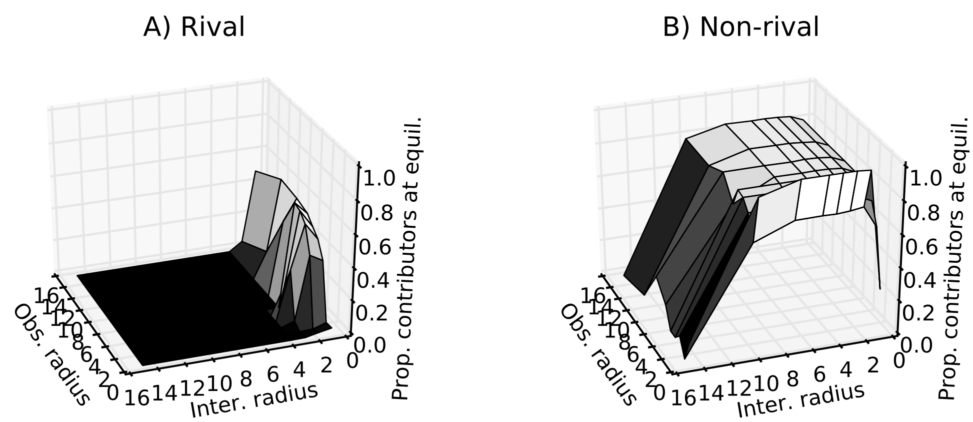


Figure 4.4: **The equilibrium proportion of contributors by observation radius and interaction radius for contributors as observation target and little mobility (μ = 0.05).** Results are shown for (A) rival and (B) non-rival contributions.

4.3.2 Results

Figures 4.3, 4.4, and 4.5 show that for non-rival contributions, the equilibrium level of contributing is visibly higher than for rival contributions. Further, for non-rival contributions, the conditions for the emergence of contribution-based communities are significantly less restricted.

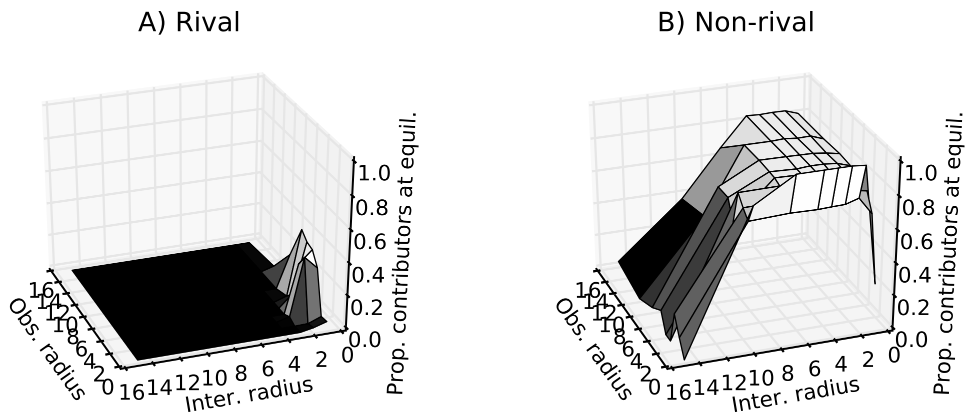


Figure 4.5: **The equilibrium proportion of contributors by observation radius and interaction radius for contributors as observation target and high mobility (μ = 0.5).** Results are shown for (A) rival and (B) non-rival contributions.

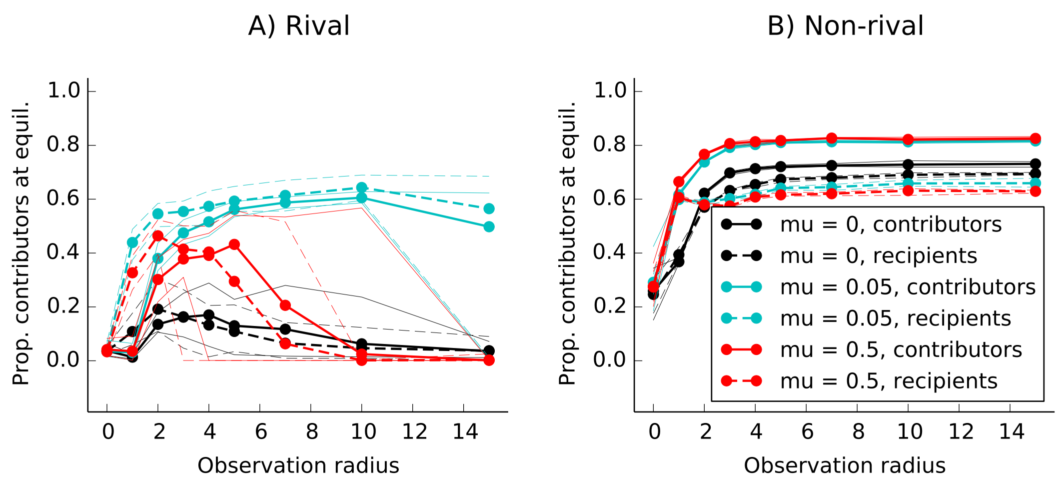


Figure 4.6: **The equilibrium proportion of contributors by observation radius when interaction is constrained to immediate neighbors only (interaction radius = 1).** Line colors show levels of mobility and line types differentiate the observation target. Results are shown for (A) rival and (B) non-rival contributions. The thick lines show the proportion averaged over 25 replications for that particular parameter combination. The thin lines show the minimum and the maximum proportions achieved in the replications.

When the exchanged contributions are non-rival, the global level of contribution is high over a large range of interaction radii. Widespread contribution fails to emerge only when the interaction radius and/or the observation radius are extremely large. This implies that non-rival exchange allows for relatively large contribution-based communities. For relatively large communities (interaction radius > 1), observed contribution has little effect, and 100% contribution is possible even when there is no observation (observation radius = 0). Overall, observing contributors has a larger positive effect than observing recipients (Figures 4.3B, 4.4B, 4.5B, and 4.6B). Community turnover does not affect outcomes except when the communities are small (interaction radius = 1) or when observation is widespread in large communities. In the first case, some mobility is better than no mobility (Figure 4.6B) and in the second case, too much mobility is bad (Figure 4.5B).

When the exchanged contributions are rival, only small communities can have high levels of contribution (optimal interaction radius ~ 2-3; Figures 4.3A, 4.4A, 4.5A). Further, observation is crucial for the emergence of contribution communities: the level of contribution is zero when there is no observation. As the observation radius increases, the level of contribution radically increases initially but eventually starts decreasing slowly (Figures 4.6A and 4.7A). The optimal observation radius is between 2 and 5, depending on the target of observation. Compared to observing contributors, observing recipients requires a smaller observation radius to achieve the maximum level of contribution. Finally, the effect of mobility is non-monotonic: low mobility ($\mu = 0.05$) is better than no mobility ($\mu = 0$) or too much mobility ($\mu = 0.5$).

Figure 4.8 identifies the reason for differences between rival and non-rival

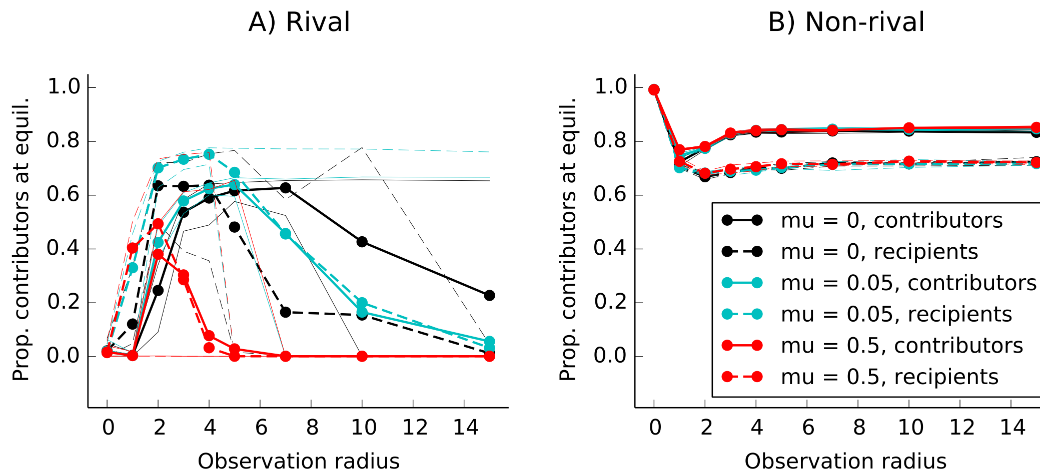


Figure 4.7: **The emergence of contribution by observation radius when interaction is constrained to immediate neighbors and neighbors of neighbors (interaction radius = 2).** Line colors show levels of mobility and line types differentiate the observation target. Results are shown for (A) rival and (B) non-rival contributions. The thick lines show the proportion averaged over 25 replications for that particular parameter combination. The thin lines show the minimum and the maximum proportions achieved in the replications.

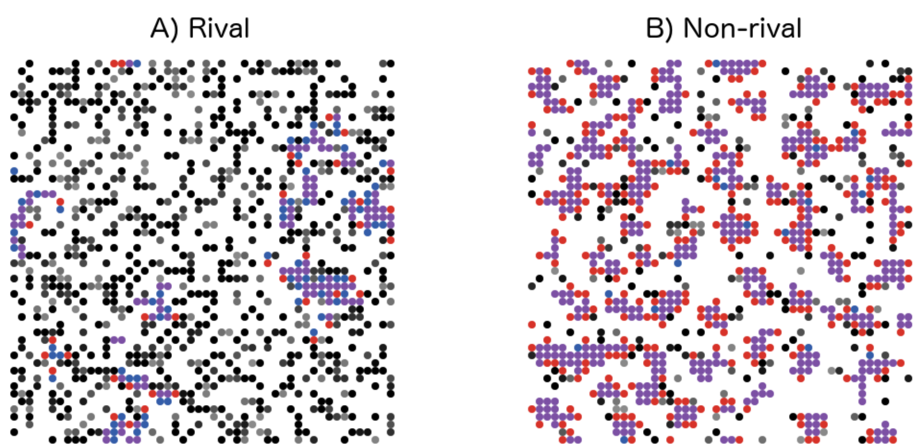


Figure 4.8: **The emergence of contribution communities for (A) rival and (B) non-rival contributions.** We show typical results for interaction radius = 1, observation radius = 5, observing recipients, and mobility $\mu = 0.05$. Agents in blue contribute but do not benefit, agents in red benefit but do not contribute, and agents in purple both contribute and benefit.

contributions. Non-rivalness implies that a larger number of individuals can benefit from a single contribution, as when a user is given advice that benefits many others in an online community. This leads to the easy formation of multiple small communities in which contributors benefit and hence continue contributing, despite free riders who benefit enough to hang around the periphery of the clusters. When contributions are rival and only one individual can benefit from each contribution, contribution-based communities are much less likely to emerge and persist. If they do, this usually happens around a core of unconditional altruists (agents with low θ_{UA} and high θ_{FR}) who form a critical mass. These agents (the blue agents in Figure 4.8A) continue contributing regardless of what others around them do. When outcome-based mobility is relatively low, the agents remain in the neighborhood long enough to have a chance to benefit from a contribution or to observe many others benefiting. (If they were observing contributors instead of recipients, they would have only observed the altruist or the few altruists that started contributing, not the many neighbors who benefit). As a result, a few clusters form around the handful of altruists in the population but the contagion does not spread to agents in other corners of the space.

The differences in the macro-outcomes between rival and non-rival contributions result from the structure of interactions and not from the differences in effect size. Assuming that the GR and TPI effects for non-rival contributions are weaker than the GR and TPI effects for rival contributions does not significantly affect the emergence of non-rival contributions.

4.4 DISCUSSION

Selfless acts of kindness and anonymous voluntary donations can be puzzling, even though they are not uncommon. In daily life, people donate blood, contribute money to charity, hold the door open for the person behind, or vacate a subway seat for an elderly passenger. In the online world, users review services, rank products, or answer strangers' questions on forums. Why do communities vary in the level of member contributions? This study suggests that the answer could lie in the contagion of prosocial behavior. We first presented empirical evidence from an online experiment for the existence and interaction of two distinct mechanisms of contagion — generalized reciprocity and third-party influence. We then implemented these mechanisms in an agent-based model to investigate the conditions under which they lead to high levels of contributions at the population level.

The empirical results showed that receiving and observing helping behavior can increase the likelihood to help a stranger. However, the willingness to contribute can be offset by lower perceived need when the level of helping is sufficiently high, particularly among those who have not themselves been helped.

We implemented these findings in a threshold model with dynamic interaction structure and adaptive behavior. The computational experiments suggested two alternative pathways for the emergence of contribution-based communities. It is useful to think of these two pathways in the context of rival face-to-face interactions on the one hand and non-rival online contributions on the other hand. In face-to-face interactions, acts of generosity are rival if the benefit is limited to the intended recipient, as happens when holding the door open or vacating one's seat for a stranger. The simulation results show that these contributions can emerge and spread in small and stable communities, that is, communities that are tightly knit and have little turnover. In such communities, hearing about or seeing other people who benefit from the kindness of strangers increases contributions. As a result, a relatively small number of persistent altruists can trigger the spread of helping behavior. In this situation, gossip and newspaper reports about anonymous acts of generosity play an important role. For example, in an office environment, a single active anonymous altruist could trigger a chain of generosity so long as there is sufficient gossip about the level of charitable behavior such that observers come to believe that generosity is normative and conform to this "office culture."

In comparison, non-rival contributions, such as writing a product review on the Web or answering a question in an online forum, are much more likely to emerge and spread across a wider range of conditions, including in much larger groups with high turnover. For example, small esoteric-interest groups and large general-topic online portals could be equally successful user-generated content communities. In such communities, hearing about or seeing other people who contribute sustains high levels of contribution, while awareness of the number of beneficiaries decreases contribution (perhaps due to the belief that there is little need for additional sacrifice).

Still, our conclusions have an important limitation. Our study provided a plausible explanation for the emergence and persistence of voluntary contribution-based communities. However, we only addressed the emergence of contribution communities among anonymous individuals. Undoubtedly, once a community forms and anonymity diminishes, cooperation-inducing mechanisms based on social sanctions (for example, reputation systems or long-term-membership privileges) become more prominent and more effective.

CHAPTER 5
CONCLUSION

Although we often prefer to think otherwise, we, humans, are much more interdependent than autonomous. The human body consists of ten times more bacterial cells than human ones (Wenner, 2007). Similarly, our ideas, beliefs, desires, and even memories (Patihis et al., 2013) are largely other people's instead of ours. We affect each other in numerous ways. And as our world is becoming increasingly more interconnected, understanding how we influence each other becomes both more challenging and more crucial.

This dissertation addressed one particular small piece of this challenge. We investigated how prosocial behavior and antisocial behavior are influenced by others and how they spread in social aggregates. We found that individuals who benefit from prosocial behavior become more likely to help others but those who observe widespread prosocial behavior become less likely to do so. Individuals who suffer from antisocial behavior are also more likely to harm others but those who observe limited harmful behavior become less likely to do harm. In other words, receiving acts of kindness could make the spread of prosocial behavior self-sustaining, while observing few mean acts could make the spread of antisocial behavior self-limiting. The self-sustaining contagion of prosocial behavior can lead to the emergence of voluntary contribution communities, particularly in the case of non-rival donations. This finding offers a plausible explanation for the rise in user-generated content communities online.

5.1 BROADER IMPLICATIONS

Our study carries a number of implications for theoretical research. First, it brings attention to the possibility that the same structural condition evokes different, even contradictory mechanisms. In the context of prosocial behavior, we found that information about others' behavior can be associated both with third-party influence and with the bystander effect, depending on the content of the information. These two mechanisms had opposing effects on human behavior. Thus, it is important not to conflate structural conditions with mechanisms. This distinction is one of the fundamental tenets of analytical sociology (Hedström, 2005) but it is often overlooked in sociological studies relying on regression analyses.

Second, the same finding also reminds us that the lack of an expected effect can be due to the simultaneous effects of two opposing mechanisms. The fact that third-party influence and the bystander effect could act under similar conditions could explain why Suri and Watts (2011) and Jordan et al. (2013) did not find evidence that prosocial behavior spreads beyond the direct interaction. It is likely that the repeated interactions in these studies induced a stronger bystander effect compared to the random matchings in the experiment analyzed by Fowler and Christakis (2010).

Third, we found that direct transmission upon contact and indirect transmission through observation are two distinct contagion mechanisms for directed social behavior. For example, new slang or non-verbal greetings can spread both when people come across them in conversation and when they observe third parties use them. Consequently, research on linguistic and cultural contagion would benefit from accounting for both contagion mechanisms.

Fourth, our research brings attention to social contagion as a mechanism for the spread of prosocial behavior. Previously, economists, psychologists, and evolutionary biologists have investigated prosocial behavior as an evolutionary predisposition (Trivers, 1971; Gintis, 2000), an instrumental calculation (Axelrod, 1984; Taylor, 1987), or an internalized social norm (Gouldner, 1960; Scott, 1971). These theoretical frameworks have generally focused on individuals responding independently, rather than as nodes of a social network. In contrast, we here concentrated on how prosocial behavior can be induced in social interaction. Below, we elaborate on how our work could be developed and extended.

Further, in a more general context, our research contributes to social science methodology by developing an experimental design and an online platform for studying the diffusion of behavior in large social groups under controlled conditions. Our experiment on prosocial behavior allowed crowd-sourced participants to interact with each other repeatedly over time in a novel way. Previous online experiments on Amazon Mechanical Turk had achieved repeated interaction by replicating the typical setup of traditional laboratory experiments: participants are gathered in a (virtual) waiting room until the desired group sizes are reached; then, they are asked to interact in periods over the next hour, waiting for all others to make their decision before their turn comes again (Suri and Watts, 2011; Mao et al., 2012). In contrast, in our experiment, participants were contacted by e-mail every time they were invited to participate and asked to respond within 24 hours. The experiment took place over two weeks and subjects participated up to six times. This design was better aligned with the capabilities of the recruitment platform (Amazon Mechanical Turk) and hence, easier to program and implement than previous designs.

Finally, our findings also have practical implications. Public broadcasting offers a good example of the potential power of appealing to listeners' feelings of normative obligation and appreciation for having benefited from the contributions of others. In general, our findings can be used to inform interventions that more effectively promote user contributions, prosocial behavior, and co-operative ventures, especially for large-scale online collaborative projects and content communities. With the advent of new digital media, we have become largely dependent on user-generated products and services — we work on a Linux operating system in the office, prepare reports with information from Wikipedia, unclog the drain at home following instructions from Yahoo! Answers, and meet friends at a restaurant found on Yelp. Harnessing generalized reciprocity and third-party influence to enhance creative participation and community formation in online communities will potentially have a positive impact on almost everyone's daily life.

5.2 RESEARCH LIMITATIONS

As is often the case, our research is not definitive and warrants replication with improved experimental designs, with entirely different research methods, and in other social contexts.

To begin with, the fact that we found either no or relatively weak empirical evidence for the hypothesized effects could be due to certain design shortcomings of the online experiments. First, despite all the measures we took to guarantee internal validity, our results suffered from the low level of attention characteristic of online subjects. This problem was exacerbated by the weakness of the stimuli. The only thing we manipulated was the text that participants read. So, participants understood the consequences of their actions to the extent to which they carefully read and understood our instructions. Undoubtedly, we could have invested more time and effort to strengthen the incentives by visualizing the instructions and introducing interactivity. For example, the experiments in Chapter 2 and Chapter 3 both involved a treatment in which the previous action was determined by chance. Rather than announcing the outcome of the random event in text, we could have implemented the draw in real time, for example, as a "lottery" button that the participant needs to press.

The high baselines in the two experiments were also problematic: subjects were already quite generous or mean in the control groups and hence, there was little variation to be manipulated. This was the case despite that fact that we had already adjusted the incentives and experiment designs once after we identified that problem in the pilot tests. In retrospect, we realize that we should have pilot tested the experiment incentives and design iteratively until we could guarantee a reasonable rate of "untreated" behavior.

Similarly to most experimental research, our empirical studies also suffer from limited generalizability. We used a sample that is not representative of the general population and future research needs to investigate whether our findings extend to populations with different demographic profiles. Our study was based on small monetary incentives and it is unclear to what extent we can expect similar effects when the gifts involved are more valuable, the sacrifices made are more impressive, and the losses incurred are more drastic. In general, it is unclear if the identified mechanisms are present in different social contexts or in daily life at all. And even if they are, our experiments cannot tell us about the actual impact of the mechanisms.

The limitations of the computational study on the contagion of prosocial behavior are even more apparent. Agent-based models of social phenomena are better suited for raising questions than for answering them; they are a tool for generating hypotheses and not for providing evidence (Macy and Willer, 2002). But even with such limited scope, the results of an agent-based model need to be interpreted cautiously since they critically depend on the behavioral assumptions the model is built upon. For example, our study used a somewhat complicated behavioral model that relied on adaptive thresholds to represent the mechanisms we found empirically. The thresholds allowed us to embed our work better in the existing literature. However, a probabilistic model might have proven simpler and more appropriate for our purposes. Further, the assumptions we made for the fixed values in the model (for instance, the length of memory) were dictated by what produces the highest variation in the results. However, such arbitrary calibration might result in values that are meaningless in reality. What is worse, the stylization of the model prevents us from judging the correspondence of our choices to real-world values. We acknowledge these problems in most social simulations, including the one proposed in this dissertation.

Finally, we are also fully aware of the limited impact of our theoretical contribution. Even though we found clear evidence for the contagion of prosocial behavior, this does not imply that there are no other motivations for helping behavior in natural settings. For example, most people who donate blood have never had surgery or an accident involving loss of blood, and even those who have may not be motivated by that experience. People who anonymously write book reviews, edit Wikipedia, or answer user inquiries might have done so because of generalized reciprocity or out of an altruistic desire to help others that is unrelated to having benefited from similar contributions by others. Nevertheless, the benefits provided by these online services may give users a feeling of appreciation that augments their other motives.

5.3 FURTHER RESEARCH

Despite its limitations, our study provides a preliminary answer to a question of significant theoretical and practical importance. Our work could be expanded and extended in numerous ways. In what follows, we discuss some future directions for research. We discuss the ideas in relation to prosocial behavior but they could easily be applied to antisocial behavior too.

5.3.1 Group size

Previous theory and research on direct reciprocity and generalized exchange suggest that generalized reciprocity may be more likely in small groups. Research on direct reciprocity has shown that the willingness to "pay it back" is more likely in small groups because of the "shadow of the future" (Axelrod, 1984), which refers to the possibility that the expected future benefits in an ongoing interaction can outweigh the immediate costs, leading to a positive net benefit in the long run. The future casts a longer shadow in smaller groups because there is a higher probability of repeated interaction. This game-theoretic principle has been confirmed in laboratory experiments showing that cooperation and collective action are more likely in small groups (Hamburger et al., 1975; Isaac and Walker, 1988). According to experimental research, generalized exchange is also more easily established in smaller groups (Greiner and Levati, 2005; Molm et al., 2007; Tsvetkova and Buskens, 2013).

This effect of group size may extend to generalized reciprocity as well. In smaller groups, the random walk of generalized reciprocity to strangers is more likely to revisit a given member. Simply put, in smaller groups, what goes around comes around sooner and more often. This argument has been theoretically demonstrated with the finding that generalized reciprocity is more likely to be evolutionarily stable in small groups and in long repeated interactions (Boyd and Richerson, 1989; Pfeiffer et al., 2005).

5.3.2 Branching generalized reciprocity

Goods and services that are non-rival can be enjoyed by multiple individuals with no increase in cost. For example, it costs the same to leave a product review on a shopping website read by thousands of people as it does to leave the same review on a website read by only a handful. Anyone who reads the review can benefit no matter how many others also benefit from the review. From a standard cost-benefit perspective, people are expected to be more likely to provide a public good the greater the number who benefit, relative to the cost (Oliver and Marwell, 1988). This principle can be extended to generalized reciprocity. The multiplier effect of providing non-rival benefits magnifies the downstream consequences of helping others and hence, one could expect that the larger the number who benefit, the greater the willingness to pay it forward.

And yet, it is also possible that receiving a non-rival benefit may not evoke the same degree of gratitude, positive affect, and moral obligation as receiving a rival benefit. Then, non-rival gifts may be less likely to be paid forward. In other words, the effect of multiple recipients for generalized reciprocity is not intuitively obvious and warrants empirical investigation. Such test will also be useful for calibrating the computational model we studied in Chapter 4.

5.3.3 Shared group identity

Another promising direction for extending our research is to test whether shared group identity fosters generalized reciprocity and the spread of generosity, as suggested by social identity theory, or on the contrary, whether group identity is instead concomitant with the expectation of generalized reciprocity, as Yamagishi and colleagues have argued (Yamagishi et al., 1999; Yamagishi and Kiyonari, 2000). A related possibility is to build upon the finding by Molm et al. (2007) that generalized exchange increases feelings of solidarity and evaluate the extent to which generalized reciprocity promotes social cohesion that offsets the segregating effects of homophily.

5.3.4 Mechanisms behind generalized reciprocity

There is some empirical evidence that for prosocial behavior, generalized reciprocity is driven by both gratitude and positive affect (Bartlett and DeSteno, 2006). However, the extent to which people perceive "paying it forward" as a normative obligation remains unknown. Further, in our experiments, participants were explicitly told that those after them will face the same decision situation but would participants have acted similarly if we had omitted that information? It will be worthwhile to find out the extent to which people recognize the contagious effect of their social actions.

5.3.5 Field experiments

Generalized reciprocity and third-party influence can and should be studied in the field too. To do this, one can make controlled interventions in new, short-term groupings such as movie sets, summer schools, or project groups in offices. In addition, one can conduct field experiments with mobile applications or on online social networks. For example, brand-related mobile applications can implement the possibility for emergent pay-it-forward chains, as those in drive-through restaurants. Further, one can trace the dissemination of pay-it-forward gift certificates along existing social relations, for instance on Facebook. Similar experiments could be done with non-rival gifts too. Thanks to the advance and spread of online technologies, such large-scale empirical research on social contagion is no longer unthinkable.

APPENDIX A

AMAZON MECHANICAL TURK

Amazon Mechanical Turk (AMT) is an online service that enables employers to distribute work in the form of small tasks (on the order of 1–30 minutes) to a large number of users. Users choose tasks that they want to work on and receive small financial rewards (typically in the range $0.01–1.00) if they complete them. Originally, the platform was intended for human intelligence tasks — tasks that a computer would find hard or impossible to do but that even a low-skill person could complete easily. However, the platform has become increasingly popular among academic researchers as a highly cost-effective method for conducting surveys and controlled experiments (Lawson et al., 2010; Bohannon, 2011; Dodds et al., 2011; Suri and Watts, 2011).

An online experiment on AMT has a number of advantages over traditional laboratory experiments involving undergraduate participants. AMT provides access to a much larger pool of hundreds of thousands of participants, simplifies participant recruitment, allows rapid implementation, reduces participant costs, and facilitates secure payments (Mason and Suri, 2012). In addition, while the AMT population is not a random sample, it is much more diverse in age, ethnicity, education, and income than the typical experimental-subjects pool of undergraduate students in American universities: age ranges from 18 to 81 with mean age of 31.6, and AMT users come from over 66 countries, with annual incomes ranging from $10,000 to more than $100,000 per year (with median income around $25,000) (Paolacci et al., 2010; Ross et al., 2010).

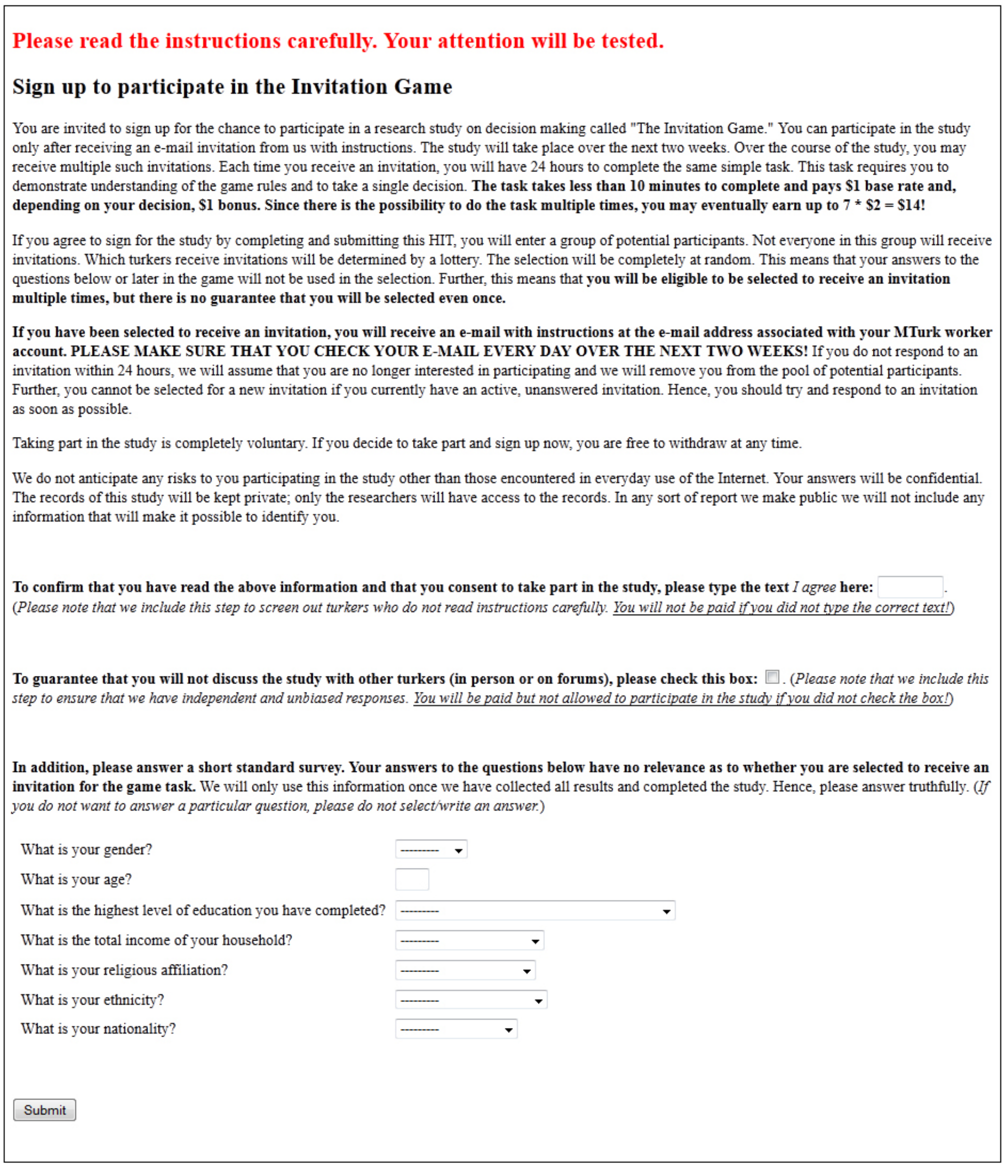
Despite these advantages, using AMT for social research poses several potential problems that also arise with traditional experiments but may be more challenging to overcome with an online platform. AMT users may not be highly motivated, are susceptible to distractions, and may drop out of a task or complete it without having fully understood the instructions. As with many traditional laboratory experiments, participants have self-selected into AMT and are not representative of the underlying population. AMT also poses problems that are unique to online experiments. In particular, some AMT users have been known to use multiple accounts and computer programs to complete tasks. As AMT has become increasingly utilized for social and behavioral studies, researchers have developed effective countermeasures that can address some of these problems. To remove computer-generated responses and reduce inattentiveness and attrition, we posted the recruitment HITs as single tasks, screened participants by past performance, and placed demographic questions and attention checks at the beginning of the studies. Previous AMT studies have demonstrated the effectiveness of these measures (Horton et al., 2011; Mason and Suri, 2012).

Finally, previous laboratory-based replications of AMT experiments have demonstrated that the decision-making and judgment behavior of AMT users does not qualitatively differ from that of subjects in traditional off-line laboratory settings (Paolacci et al., 2010; Horton et al., 2011). Importantly, AMT users exhibit prosocial values in Prisoner's Dilemma situations that are similar to those observed in equivalent laboratory experiments (Horton et al., 2011; Suri and Watts, 2011). These results suggest reasonable generalizability for the results from AMT-based behavioral studies.

APPENDIX B

APPENDIX TO THE CONTAGION OF PROSOCIAL BEHAVIOR

B.1 EXPERIMENT INSTRUCTIONS



Please read the instructions carefully. Your attention will be tested.

Sign up to participate in the Invitation Game

You are invited to sign up for the chance to participate in a research study on decision making called "The Invitation Game." You can participate in the study only after receiving an e-mail invitation from us with instructions. The study will take place over the next two weeks. Over the course of the study, you may receive multiple such invitations. Each time you receive an invitation, you will have 24 hours to complete the same simple task. This task requires you to demonstrate understanding of the game rules and to take a single decision. **The task takes less than 10 minutes to complete and pays $1 base rate and, depending on your decision, $1 bonus. Since there is the possibility to do the task multiple times, you may eventually earn up to 7 * $2 = $14!**

If you agree to sign for the study by completing and submitting this HIT, you will enter a group of potential participants. Not everyone in this group will receive invitations. Which turkers receive invitations will be determined by a lottery. The selection will be completely at random. This means that your answers to the questions below or later in the game will not be used in the selection. Further, this means that **you will be eligible to be selected to receive an invitation multiple times, but there is no guarantee that you will be selected even once.**

If you have been selected to receive an invitation, you will receive an e-mail with instructions at the e-mail address associated with your MTurk worker account. PLEASE MAKE SURE THAT YOU CHECK YOUR E-MAIL EVERY DAY OVER THE NEXT TWO WEEKS! If you do not respond to an invitation within 24 hours, we will assume that you are no longer interested in participating and we will remove you from the pool of potential participants. Further, you cannot be selected for a new invitation if you currently have an active, unanswered invitation. Hence, you should try and respond to an invitation as soon as possible.

Taking part in the study is completely voluntary. If you decide to take part and sign up now, you are free to withdraw at any time.

We do not anticipate any risks to you participating in the study other than those encountered in everyday use of the Internet. Your answers will be confidential. The records of this study will be kept private; only the researchers will have access to the records. In any sort of report we make public we will not include any information that will make it possible to identify you.

To confirm that you have read the above information and that you consent to take part in the study, please type the text ***I agree*** **here:** .
(*Please note that we include this step to screen out turkers who do not read instructions carefully. You will not be paid if you did not type the correct text!*)

To guarantee that you will not discuss the study with other turkers (in person or on forums), please check this box: . (*Please note that we include this step to ensure that we have independent and unbiased responses. You will be paid but not allowed to participate in the study if you did not check the box!*)

In addition, please answer a short standard survey. Your answers to the questions below have no relevance as to whether you are selected to receive an invitation for the game task. We will only use this information once we have collected all results and completed the study. Hence, please answer truthfully. (*If you do not want to answer a particular question, please do not select/write an answer.*)

What is your gender?	---------
What is your age?	
What is the highest level of education you have completed?	---------
What is the total income of your household?	---------
What is your religious affiliation?	---------
What is your ethnicity?	---------
What is your nationality?	---------

Submit

Figure B.1: **Recruitment HIT for the Invitation Game.**

From: Cornell SDL <mvt9@cornell.edu>
Subject: You have been invited to the Invitation Game

Message from Cornell SDL (mvt9@cornell.edu)

Dear turker,

You have been invited to complete the task associated with the MTurk HIT "Sign up to participate in the Invitation Game," which you submitted. Your invitation is valid for the next 24 hours.

To complete the task, please use the following information:

* MTurk Worker ID: A27L6Z6PBCE04Y
* Invitation ID: ILUS

and:

1. Go to https://sdlab.soc.cornell.edu/study11/igame/ and complete the task.
2. After you have completed the task, go to https://www.mturk.com/mturk/preview?groupId=2FH56XBAT2D5PP8RRYUQ7JG7YZP04I and submit the HIT.

Thank you for your participation!

Best regards,

Milena Tsvetkova, Cornell SDL

***

If you have questions, you may contact me at mvt9@cornell.edu. If you have any questions or concerns regarding your rights as a subject in this study, you may contact the Cornell Institutional Review Board (IRB) at 607-255-5138 or access their website at http://www.irb.cornell.edu. You may also report your concerns or complaints anonymously through Ethicspoint (www.hotline.cornell.edu) or by calling toll free at 1-866-293-3077. Ethicspoint is an independent organization that serves as a liaison between the University and the person bringing the complaint so that anonymity can be ensured.

Greetings from Amazon Mechanical Turk,

The message above was sent by an Amazon Mechanical Turk user.
Please review the message and respond to it as you see fit.

Sincerely,
Amazon Mechanical Turk
https://workersandbox.mturk.com
410 Terry Avenue North
SEATTLE, WA 98109-5210 USA

Figure B.2: **E-mail invitation for the Invitation Game.**

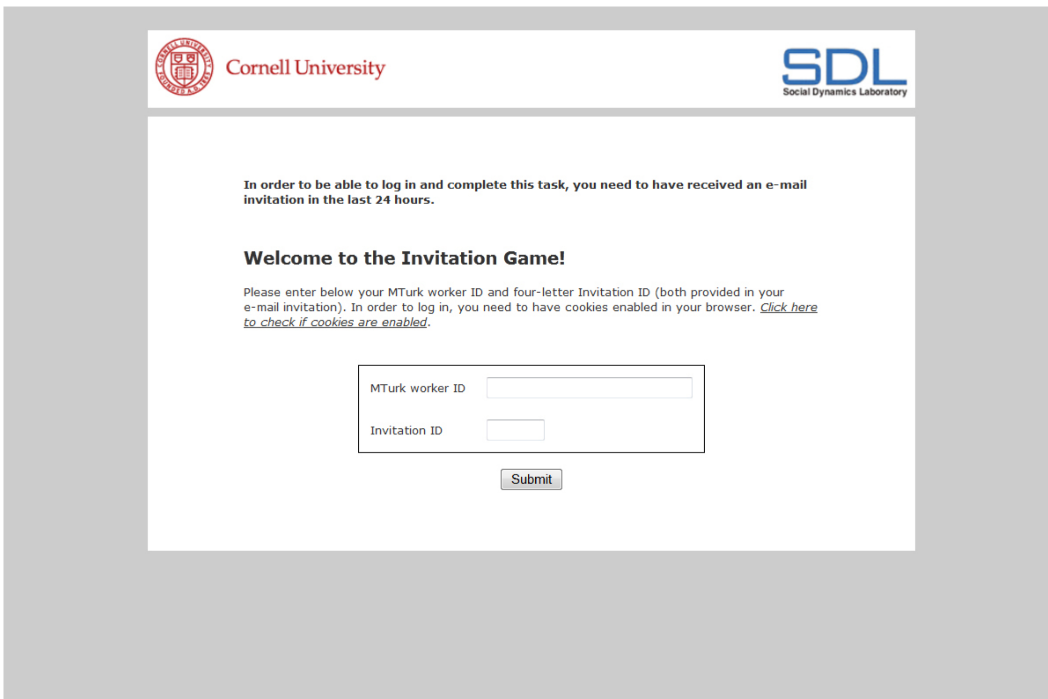


Figure B.3: **Login page for the Invitation Game.**

Cornell University

SDL
Social Dynamics Laboratory

Since the amount of money you receive in the game depends on the decisions that you and the other participants make, it is important that you understand and remember the game rules well. Please read the instructions below carefully and then answer five questions about them. You will need to demonstrate your understanding of the rules before you can proceed to make your decision.

How the Invitation Game Works

You are part of a group of 10 turkers who have signed up to participate in the game. To participate in the game, you need an invitation. Unfortunately, we do not have invitations available for everyone and hence, we send invitations to randomly selected participants. Each participant who receives and accepts an invitation receives a payment of $1 and a bonus of $1. However, they then have the choice to return their bonus, in which case we create a new invitation and allow one more person than otherwise to participate. The recipient of the new invitation is randomly selected from the other 9 turkers in the group. This means that a participant is not eligible for the invitation they create but otherwise, they may be selected to receive any invitation created by any of the other 9 participants in the group.

When a participant returns their bonus, we multiply the bonus amount by 2 and thus, the next invited participant receives the same $1 regular payment and $1 bonus. This next participant also faces the same choice: they need to decide whether to keep their bonus or return it and create a new invitation for one more turker. This means that if participant A decides to return their bonus and invite turker B, and then turker B also decides to return their bonus, the chance that participant A is selected to complete the task again increases. If participant A decides to keep their bonus, we will not create and send an invitation to turker B. However, regardless of participant A's decision, participant A remains eligible for any invitations currently outstanding or created in the future by others in the group.

The identities of all participants will be kept anonymous and the decisions made by individual participants will not be revealed to others.

Now, please answer the following questions. You will not be allowed to proceed until you have answered all questions correctly.

The game starts with a small number of invitations allocated by the researchers but recipients of invitations may create new invitations. ○ True ○ False

In order to receive an invitation,

- ○ one needs to have answered all questions correctly.
- ○ one needs to have already created many invitations.
- ○ one needs to be randomly selected from the members of one's group.

Every invited participant receives the same payment and bonus and faces the same choice options regarding what to do with their bonus. ○ True ○ False

As a participant, if you decide to return your bonus and create a new invitation,

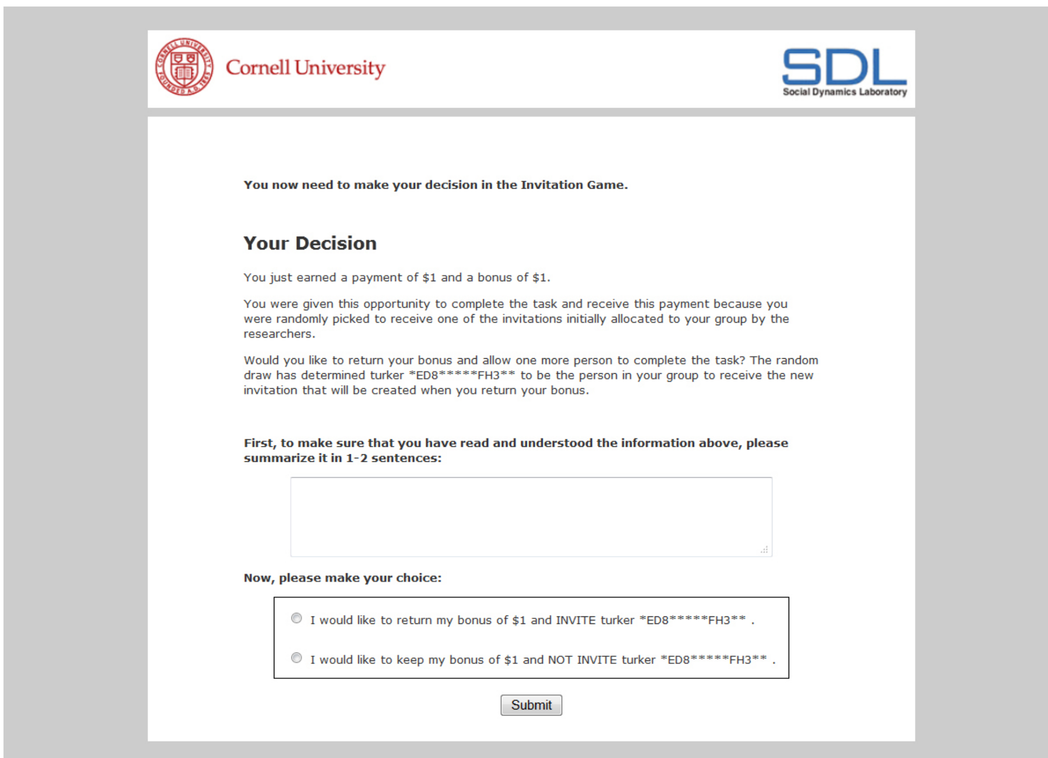
- ○ you may be selected to receive the invitation you create.
- ○ you cannot receive the invitation you create but you may be selected to receive other invitations, including the invitation created by the recipient of your invitation.

As a participant, if you decide to keep your bonus,

- ○ you may still be selected to receive invitations.
- ○ you exit the game and can no longer receive any invitations.

Submit

Figure B.4: **Instructions for the Invitation Game.**


Cornell University

SDL
Social Dynamics Laboratory

You now need to make your decision in the Invitation Game.

Your Decision

You just earned a payment of $1 and a bonus of $1.

You were given this opportunity to complete the task and receive this payment because you were randomly picked to receive one of the invitations initially allocated to your group by the researchers.

Would you like to return your bonus and allow one more person to complete the task? The random draw has determined turker *ED8*****FH3** to be the person in your group to receive the new invitation that will be created when you return your bonus.

First, to make sure that you have read and understood the information above, please summarize it in 1-2 sentences:

Now, please make your choice:

- I would like to return my bonus of $1 and INVITE turker *ED8*****FH3** .
- I would like to keep my bonus of $1 and NOT INVITE turker *ED8*****FH3** .

Submit

Figure B.5: **Decision page for a seed in the no-observation, low-payment treatment group in the Invitation Game.**

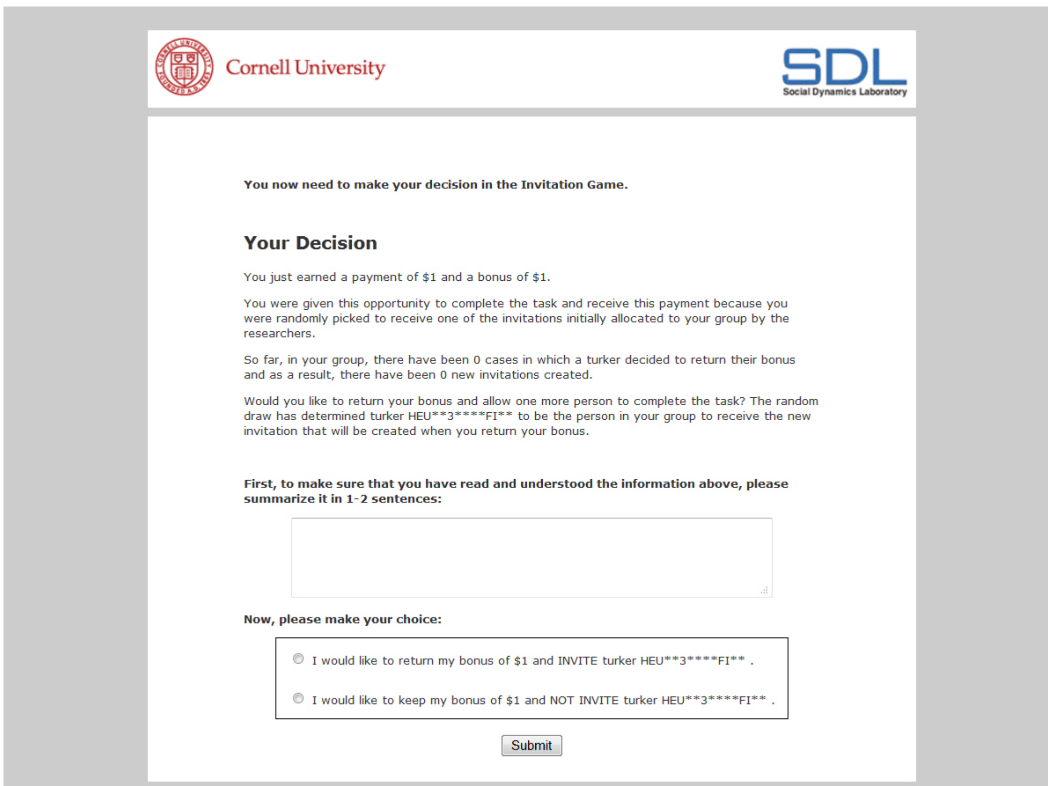
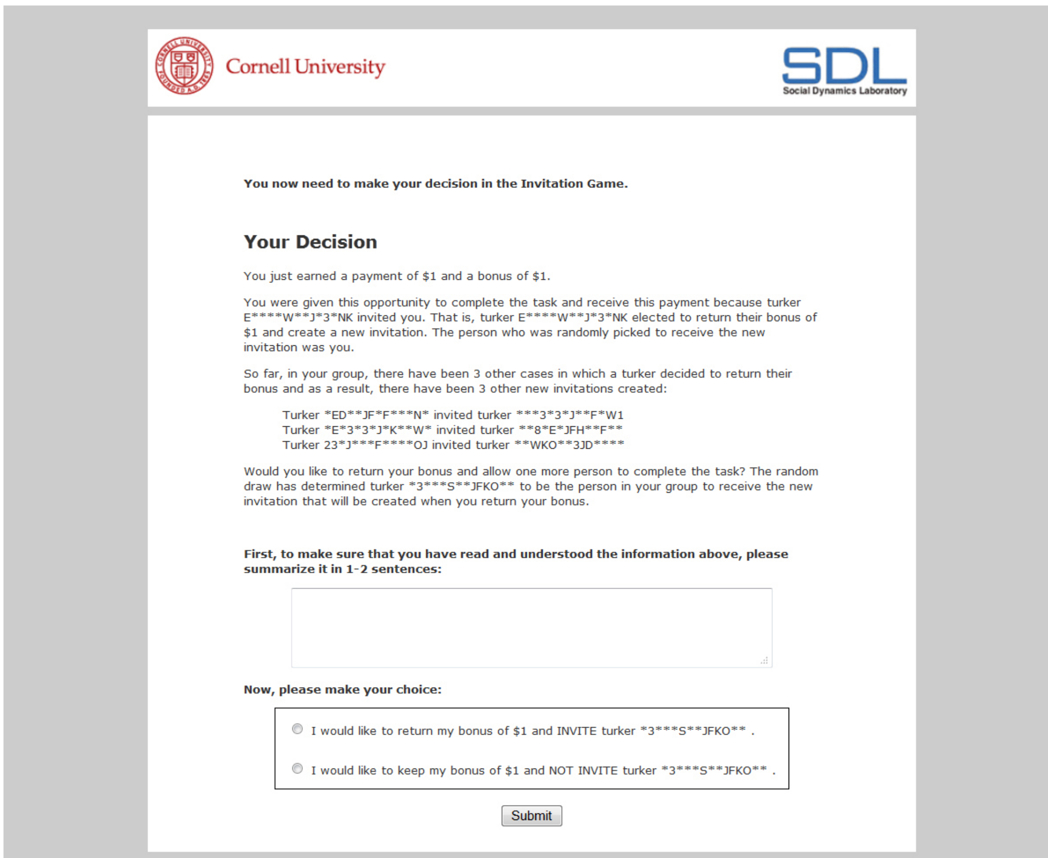


Figure B.6: **Decision page for a seed in the observation, low-payment treatment group in the Invitation Game.**



Cornell University

SDL
Social Dynamics Laboratory

You now need to make your decision in the Invitation Game.

Your Decision

You just earned a payment of $1 and a bonus of $1.

You were given this opportunity to complete the task and receive this payment because turker E****W**J*3*NK invited you. That is, turker E****W**J*3*NK elected to return their bonus of $1 and create a new invitation. The person who was randomly picked to receive the new invitation was you.

So far, in your group, there have been 3 other cases in which a turker decided to return their bonus and as a result, there have been 3 other new invitations created:

Turker *ED**JF*F***N* invited turker ***3*3*J**F*W1
Turker *E*3*3*J*K**W* invited turker **8*E*JFH**F**
Turker 23*J***F****OJ invited turker **WKO**3JD****

Would you like to return your bonus and allow one more person to complete the task? The random draw has determined turker *3***S**JFKO** to be the person in your group to receive the new invitation that will be created when you return your bonus.

First, to make sure that you have read and understood the information above, please summarize it in 1-2 sentences:

Now, please make your choice:

- I would like to return my bonus of $1 and INVITE turker *3***S**JFKO** .
- I would like to keep my bonus of $1 and NOT INVITE turker *3***S**JFKO** .

Submit

Figure B.7: **Decision page for an invitee in the observation, low-payment treatment group in the Invitation Game.**

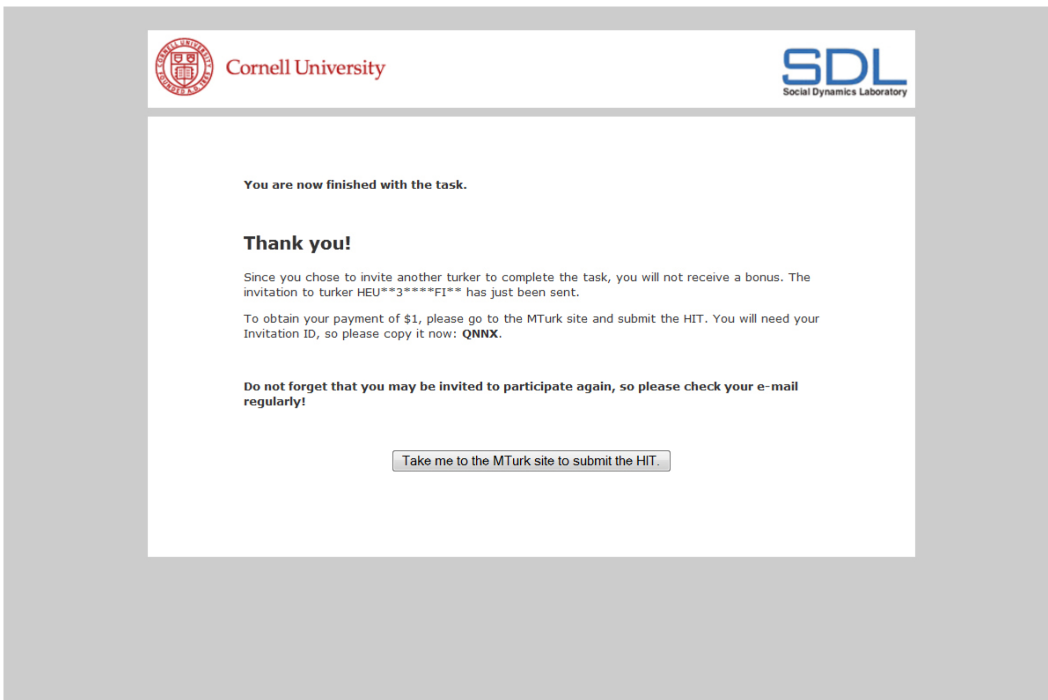


Figure B.8: **Final page in the Invitation Game.**

B.2 ADDITIONAL ANALYSES

B.2.1 Experimental procedure

The study was conducted over a period of six weeks in March-April, 2013, in two sessions, each lasting 10-14 days. The two sessions corresponded to the two different payment treatments. To avoid learning effects, we did not allow AMT users who participated in the first session to sign up for and participate in the second session. Further, we scheduled the two sessions two weeks apart in order to minimize carry-over effects due to participants obtaining a lower base rate than the rate they might remember from the recruitment advertisement for the previous session.

Since recipients of invitations were randomly selected, not all of the AMT users assigned to the four 150-person groups received an invitation. Further, not all of the AMT users who received invitations responded to them. 662 individuals received at least a first invitation, to which 89 did not respond, either because they did not check their e-mail on time, they did not have an opportunity to respond on time, or they were no longer interested in participating. If a participant did not respond to an invitation within 24 hours, we removed that participant from the group, added another randomly selected AMT user from the participant pool to the group to maintain 150 members, and forwarded the unanswered invitation to another randomly selected group member.

The experiment did not involve deception of any kind. Invitations were actually created by participants. Hence, the number of donated invitations participants received or observed depended on the number of previous participants

Table B.1: **Number of observations and number of participants by experimental manipulation.**

		Seed	Invitee	Total
No observation	Low payment	112 (40)	136 (47)	248 (126)
	High payment	84 (26)	184 (61)	268 (126)
Observation	Low payment	93 (29)	185 (65)	278 (129)
	High payment	82 (27)	194 (68)	276 (137)
Total		371 (122)	699 (241)	1070 (518)

The brackets for seeds and invitees show the number of unique participants who interacted only as seeds or only as invitees. The brackets in the "Total" column count also the participants who interacted as both seeds and invitees.

who had chosen to donate their bonus. Thus, avoiding deception came at the cost of endogenizing these manipulations. However, we took concrete measures to reduce any confounding effects from the endogenous manipulations. First, we invited new seeds throughout the experiment in order to minimize the difference in waiting time for first invitation between seeds and invitees. On average, seeds received their first invitation 49 hours after signing up for the study (min = 0.8, max = 130); for invitees, the average waiting time was about 56.5 hours (min = 0.4, max = 198). Second, the analyses control for the time between interactions (for the first interaction, this is the time elapsed since signing up for the study) to account for any remaining difference in waiting time and for the fact that invitees interacted more often than seeds due to the high level of generosity. Third, we did not inform participants when the experiment in their group started, how many seed invitations had been already sent out, and when the experiment in their group was to end (participants only knew that they may be selected to participate anywhere between 0 and 7 times). This means that participants did not know what their chances were for receiving another invitation and hence, could not condition their behavior on such knowledge. Similarly, in the observation manipulation, participants did not know what the number of already created invitations implied for the number of future invitations. Since the effect of observation starts decreasing as early as 75 invitations, we do not believe that the non-monotonicity of TPI is driven by an "end-game effect."

B.2.2 Internal validity

To improve the internal validity of the study, we required participants to answer correctly five multiple-choice questions that tested their comprehension of the game rules before they could proceed. The questions emphasized that invitations were distributed randomly and that while inviting someone else could increase one's chance to be invited again, not inviting does not decrease it. In addition to the multiple-choice questions, participants were required to write a short summary to demonstrate that they understood the decision they were asked to make. (See Section B.1.)

On average, participants took 1.7 attempts to answer the five questions correctly but the distribution is extremely skewed to the left, with 35 participants who took more than 5 attempts and a maximum of 34 attempts. Participants who required a large number of attempts were likely randomly guessing the answers to the questions without having read or understood the instructions. The summaries written by participants were blindly hand-coded without knowledge of the participant's treatment or decision. Common errors included assuming that the participant exits the game if they do not return their bonus, that the turkers from the list of previous donations or that all other 149 group members will receive invitations if the participant returns their bonus, or that the participant was invited by another turker when in fact they were treated as a seed.

To improve the internal validity of the results, the analyses in Table 2.1 exclude data from the 55 participants (126 observations) who required more than five attempts to answer the five multiple-choice questions correctly or whose written summaries revealed an apparent lack of understanding of the instructions. In Table B.2, we have replicated the analyses for the complete data. The results are qualitatively the same. The major difference is that the effect from GR in A) is smaller and loses statistical significance. The GR treatment was less visible (a 4-line paragraph) compared to the observation treatment (a long list of donors and recipients) and hence, it was probably overlooked by participants who were not paying attention.

Table B.2: **Odds ratios for donating across treatments for the complete sample.**

Manipulation	**A) GR**	**B) TPI+**	**C) TPI−**	**D) GR × TPI**
Invitee	1.931			0.283
	(0.269)			(0.151)
Has been invitee	0.427			1.704
	(0.207)			(0.515)
Seeds				
Observes 0–75		7.933*	(baseline)	(baseline)
		(0.041)		
Observes 76–150		1.310	0.098	0.164
		(0.779)	(0.243)	(0.099)
Observes 151+		0.221	0.009	0.023*
		(0.237)	(0.199)	(0.020)
Invitees				
Observes 76–150				13.691*
				(0.043)
Observes 151+				45.937*
				(0.028)
High payment	25.920**	2.805	1.012	3.114
	(0.002)	(0.194)	(0.994)	(0.240)
Time waited (in hours)	0.981	0.992	0.998	0.975
	(0.052)	(0.501)	(0.933)	(0.042)
Previous participations	1.184	1.150	0.773	0.531
	(0.599)	(0.743)	(0.842)	(0.198)
Baseline odds	3.023	4.002	196.341	163.839***
	(0.165)	(0.104)	(0.092)	(0.000)
Number of observations	569	415	195	627
Number of participants	278	310	149	295
Wald χ^2	5 df, 14.03*	6 df, 6.75	5 df, 2.62	9 df, 13.41
	(0.016)	(0.354)	(0.758)	(0.145)

Two-sided tests: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
Replication of Table 2.1 for all 1,196 observations and 573 individuals, including individuals who demonstrated poor understanding of the game rules. The table reports odds ratios and p-values (in brackets) from random-intercept logistic regression models for A) seeds and invitees in the no-observation treatment by number of donated invitations received; B) seeds in the observation and no-observation treatments by number of donated invitations observed; C) seeds in the observation treatment by number of donated invitations observed; and D) seeds and invitees in the observation treatment by number of donated invitations observed by invitees compared to seeds.

B.2.3 Demographics

Table B.3: **Detailed demographics for the participant sample (*N* = 518).**

Characteristic	**Mean / Percent**
Female	38.77
Age	29.99 (*SD* = 9.56)
Income	
Less than $10,000	9.67
$10,000 – $19,999	9.86
$20,000 – $29,999	13.73
$30,000 – $39,999	13.73
$40,000 – $49,999	11.22
$50,000 – $59,999	10.83
$60,000 – $69,999	6.77
$70,000 – $79,999	6.58
$80,000 – $89,999	5.80
$90,000 – $99,999	2.90
$100,000 – $149,999	7.54
$150,000 or more	1.35
Education	
Less than high school	0.77
High school or GED	12.16
Some college	34.75
Associate's degree	7.53
Bachelor's degree	35.52
Graduate degree (Master's, Doctorate, etc.)	9.27
Nationality	
United States	91.31
India	5.98
Other	2.71
Ethnicity	
White, non-Hispanic	72.15
Asian-Pacific Islander	13.73
African-American	5.80
Hispanic	3.87
Native American	1.35
Other	3.09
Religion	
Non-religious	29.34
Atheist	25.48
Protestant	10.42
Roman Catholic	9.85
Other Christian	12.36
Hindu	5.79
Buddhist	1.74
Jewish	1.16
Muslim	0.77
Other non-Christian	3.09

Table B.4: **Odds ratios for donating as predicted by demographic variables.**

Characteristic	**Odds ratio (*p*-value)**
Age	1.125**
	(0.007)
Female	0.901
	(0.887)
Income	1.187
	(0.167)
Education: Associate's or some college	0.247
	(0.218)
Education: Bachelor's or graduate degree	0.671
	(0.730)
Religion: non-Christian	0.339
	(0.400)
Religion: non-religious	2.384
	(0.344)
Religion: atheist	0.740
	(0.749)
Nationality: non-USA	0.762
	(0.858)
Ethnicity: Asian or Pacific Islander	0.440
	(0.485)
Ethnicity: other non-White	0.835
	(0.863)
Baseline odds	19.525*
	(0.026)
Number of observations	1067
Number of participants	516
Wald χ^2	11 df, 15.90
	(0.145)

Two-sided tests: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
The table reports odds ratios and *p*-values (in brackets) from a random-intercept logistic regression model. The baseline odds are for a thirty-year-old white American Christian male with high-school education or less and household income of less than $10,000.

B.2.4 Between-individual and within-individual effects

We replicate the analyses in Table 2.1 with within-subject centering (van de Pol and Wright, 2009) in order to separate between-individual effects from within-individual effects. Between-individual effects represent time-invariant differences in "types" of participants, e.g. seeds and invitees. Within-individual effects refer to changes over time in a "representative" individual as this participant receives or observes additional invitations. The between-individual values were calculated by averaging the manipulation over all of the observations for a particular individual. The within-individual values were calculated by taking the deviation of the manipulation in the focal observation from the individual's mean manipulation (i.e. the between-individual value). Thus, participants who interacted only once did not contribute to the calculation of within-individual effects.

The results reported in Table B.5 reveal that the within-individual effects are generally stronger than the between-individual effects. Most strikingly, the effect of GR is entirely due to increased odds of donation among former seeds who become invitees and not due to time-invariant between-individual differences in behavior. In other words, participants needed to experience both the "seed" and "invitee" condition in order to activate conditionally generous behavior. Similarly, the effect of observation was more pronounced among participants who observed different levels of generosity in subsequent interactions. These results may be due to the fact that the experiment involved minimal GR and TPI stimuli, which might have become more prominent with repeated interaction.

Table B.5: **Odds ratios for donating across treatments with disaggregated between-individual and within-individual effects.**

Manipulation	**A) GR**	**B) TPI+**	**C) TPI−**	**D) GR × TPI**
Invitee				0.948
				(0.970)
Between individuals	0.692			
	(0.814)			
Within individuals	15.238**			
	(0.008)			
Has been invitee	1.017			0.760
	(0.985)			(0.782)
Seeds				
Observes 0–75			(baseline)	(baseline)
Between individuals		4.428		
		(0.295)		
Within individuals		140.502*		
		(0.035)		
Observes 76–150				
Between individuals		1.974	0.483	0.655
		(0.667)	(0.809)	(0.862)
Within individuals		7.170	0.007	0.041
		(0.235)	(0.197)	(0.082)
Observes 151+				
Between individuals		0.280	0.011	0.059
		(0.419)	(0.290)	(0.243)
Within individuals		- -	0.000	0.003*
			(0.207)	(0.026)
Invitees				
Observes 76–150				
Between individuals				23.012
				(0.211)
Within individuals				74.650
				(0.051)
Observes 151+				
Between individuals				5.125
				(0.515)
Within individuals				1200.184*
				(0.018)
High payment	77.251**	2.586	0.881	3.527
	(0.005)	(0.287)	(0.944)	(0.290)
Time waited (in hours)	0.972*	0.994	1.032	0.974
	(0.025)	(0.645)	(0.486)	(0.081)
Previous participations	0.517	0.876	3.618	0.415
	(0.147)	(0.794)	(0.550)	(0.172)
Baseline odds	16.470*	4.805	15.529	121.539*
	(0.049)	(0.122)	(0.377)	(0.015)
Number of observations	516	371	175	554
Number of participants	252	277	133	266
Wald χ^2	6 df, 13.84*	8 df, 6.87	7 df, 2.30	13 df, 11.42
	(0.032)	(0.551)	(0.942)	(0.575)

Two-sided tests: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The table reports odds ratios and p-values (in brackets) from random-intercept logistic regression models for A) seeds and invitees in the no-observation treatment by number of donated invitations received; B) seeds in the observation and no-observation treatments by number of donated invitations observed; C) seeds in the observation treatment by number of donated invitations observed; and D) seeds and invitees in the observation treatment by number of donated invitations observed by invitees compared to seeds.

B.2.5 Robustness by payment

To test the results for robustness by payment, we replicated the analyses in Table 1 separately for the low-payment condition (Table B.6) and the high-payment condition (Table B.7). Since we halve the sample size, the statistical power decreases and the tests are no longer significant. Nevertheless, the direction of the GR and TPI effects is consistent across the two payment conditions. The size of the effects varies but not significantly. Hence, we can conclude that there are no important differences in GR and TPI between the two payment conditions we investigated.

Table B.6: **Odds ratios for donating across treatments for the low payment condition.**

Manipulation	A) GR	B) TPI+	C) TPI−	D) GR × TPI
Invitee	4.493			0.331
	(0.249)			(0.400)
Has been invitee	0.441			0.899
	(0.511)			(0.922)
Seeds				
Observes 0–75		23.889	(baseline)	(baseline)
		(0.103)		
Observes 76–150		5.486	0.038	0.304
		(0.324)	(0.285)	(0.444)
Observes 151+		0.392	0.000	0.030
		(0.621)	(0.080)	(0.093)
Invitees				
Observes 76–150				5.427
				(0.353)
Observes 151+				34.583
				(0.129)
Time waited (in hours)	0.945*	0.979	1.062	1.010
	(0.022)	(0.296)	(0.419)	(0.620)
Previous participations	0.740	0.628	0.864	0.566
	(0.612)	(0.437)	(0.956)	(0.410)
Baseline odds	20.428*	7.707	801.230*	45.870*
	(0.047)	(0.129)	(0.046)	(0.015)
Number of observations	248	205	93	278
Number of participants	126	143	64	129
Wald χ^2	4 df, 5.99	5 df, 4.16	4 df, 4.51	8 df, 7.80
	(0.200)	(0.526)	(0.341)	(0.453)

Two-sided tests: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
The table reports odds ratios and p-values (in brackets) from random-intercept logistic regression models for A) seeds and invitees in the no-observation treatment by number of donated invitations received; B) seeds in the observation and no-observation treatments by number of donated invitations observed; C) seeds in the observation treatment by number of donated invitations observed; and D) seeds and invitees in the observation treatment by number of donated invitations observed by invitees compared to seeds.

Table B.7: **Odds ratios for donating across treatments for the high payment condition.**

Manipulation	**A) GR**	**B) TPI+**	**C) TPI−**	**D) GR × TPI**
Invitee	22.443*			0.323
	(0.049)			(0.680)
Has been invitee	1.351			2.843
	(0.814)			(0.766)
Seeds				
Observes 0–75		5.543	(baseline)	(baseline)
		(0.263)		
Observes 76–150		0.357	0.000	0.000
		(0.499)	(0.172)	(0.062)
Observes 151+		0.108	0.000	0.000
		(0.304)	(0.308)	(0.212)
Invitees				
Observes 76–150				92352.7
				(0.054)
Observes 151+				27786.6
				(0.335)
Time waited (in hours)	0.998	1.010	1.011	0.837**
	(0.922)	(0.649)	(0.852)	(0.002)
Previous participations	0.704	1.491	25.515	0.095
	(0.610)	(0.701)	(0.458)	(0.247)
Baseline odds	40.645	6.214	46799.9*	1.75×10^{12}***
	(0.056)	(0.247)	(0.016)	(0.000)
Number of observations	268	166	82	276
Number of participants	126	134	69	137
Wald χ^2	4 df, 4.40	5 df, 2.81	4 df, 2.66	8 df, 11.92
	(0.355)	(0.730)	(0.617)	(0.155)

Two-sided tests: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
The table reports odds ratios and p-values (in brackets) from random-intercept logistic regression models for A) seeds and invitees in the no-observation treatment by number of donated invitations received; B) seeds in the observation and no-observation treatments by number of donated invitations observed; C) seeds in the observation treatment by number of donated invitations observed; and D) seeds and invitees in the observation treatment by number of donated invitations observed by invitees compared to seeds.

APPENDIX C

APPENDIX TO THE CONTAGION OF ANTISOCIAL BEHAVIOR

C.1 EXPERIMENT INSTRUCTIONS

Participate in the Bonus Game

You are invited to participate in a research study on decision making called the "Bonus Game." You will receive a guaranteed $0.25 base rate if you answer a short demographic survey, read the game instructions, and complete a five-question quiz on the instructions. **If you correctly answer the quiz within three attempts, you will proceed to play the game.**

Playing the game will pay you an additional $0.50 for participation and give you the chance to earn an extra bonus of up to $0.75. The extra amount you earn depends on your own and other participants' decisions. Reading the instructions and playing the game takes about 10 minutes.

Taking part in the study is completely voluntary. If you decide to take part now, you are free to withdraw at any time.

We do not anticipate any risks to you participating in the study other than those encountered in everyday use of the Internet. Your answers will be confidential. The records of this study will be kept private; only the researchers will have access to the records. In any sort of report we make public we will not include any information that will make it possible to identify you.

If you agree to the above conditions, please accept the HIT and go to https://sdlab.soc.cornell.edu/study14/bgame/. You will need your MTurk worker ID so please copy it now (you can find it under "Your Account" tab). Once you have finished the game, you will be given a confirmation code. **In order to get paid, you need to enter the confirmation code here and then submit the HIT.**

CONFIRMATION CODE:

***

If you have questions, you may contact us at mvt9@cornell.edu. If you have any questions or concerns regarding your rights as a subject in this study, you may contact the Cornell Institutional Review Board (IRB) at 607-255-5138 or access their website at http://www.irb.cornell.edu. You may also report your concerns or complaints anonymously through Ethicspoint (www.hotline.cornell.edu) or by calling toll free at 1-866-293-3077. Ethicspoint is an independent organization that serves as a liaison between the University and the person bringing the complaint so that anonymity can be ensured.

Figure C.1: **Recruitment HIT for the Bonus Game.**

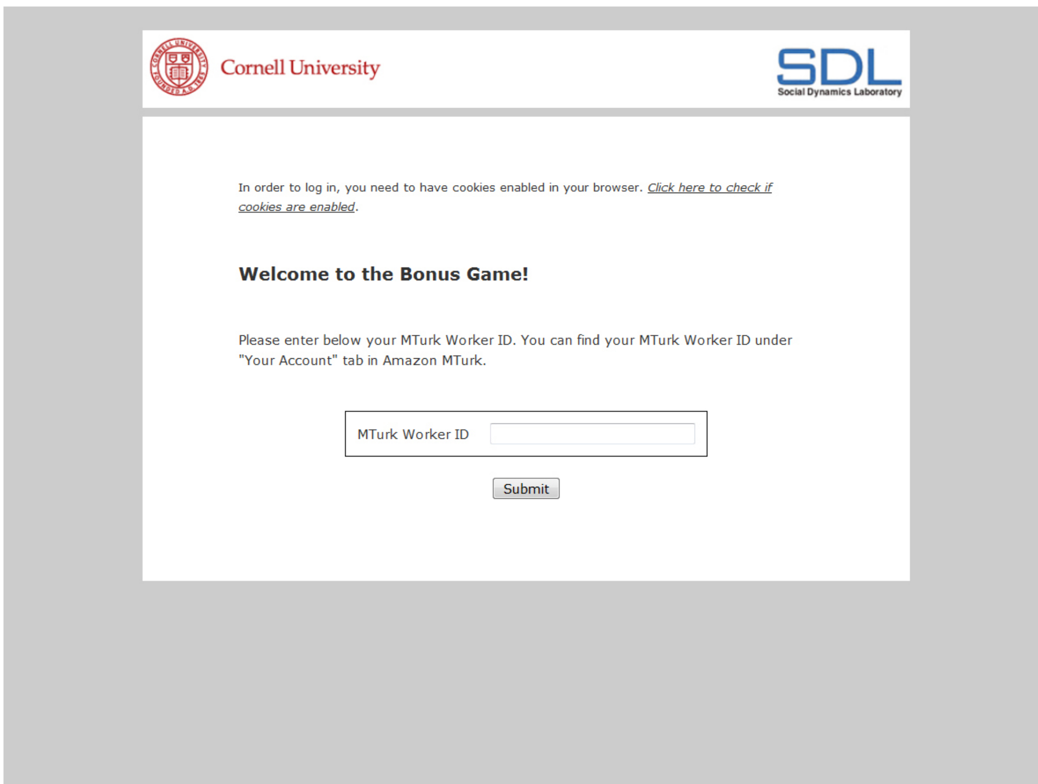


Figure C.2: **Login page for the Bonus Game.**

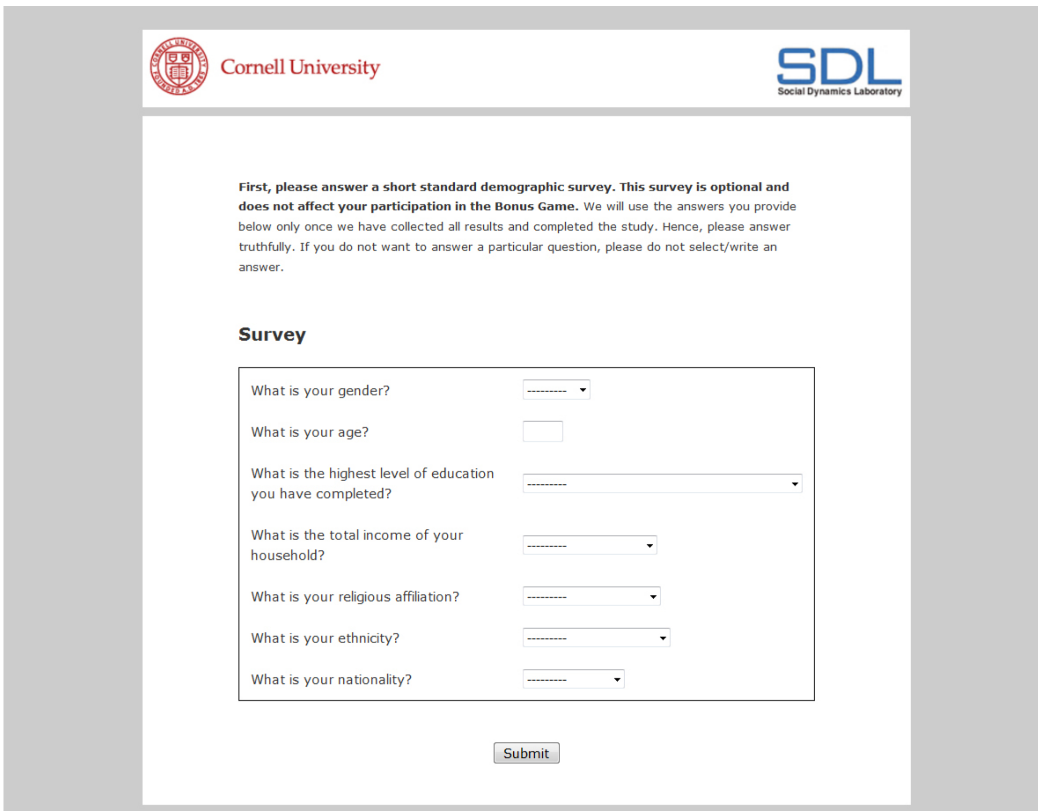


Figure C.3: **Survey for the Bonus Game.**

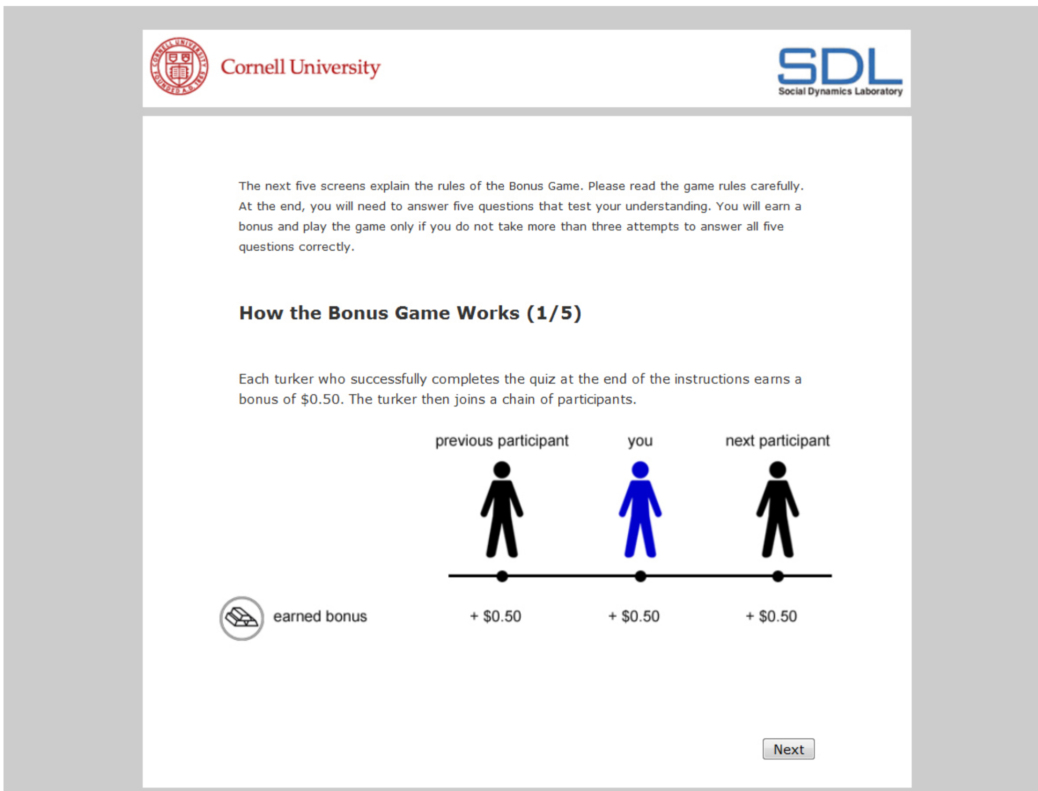


Figure C.4: **Page 1 of instructions for the Bonus Game.**

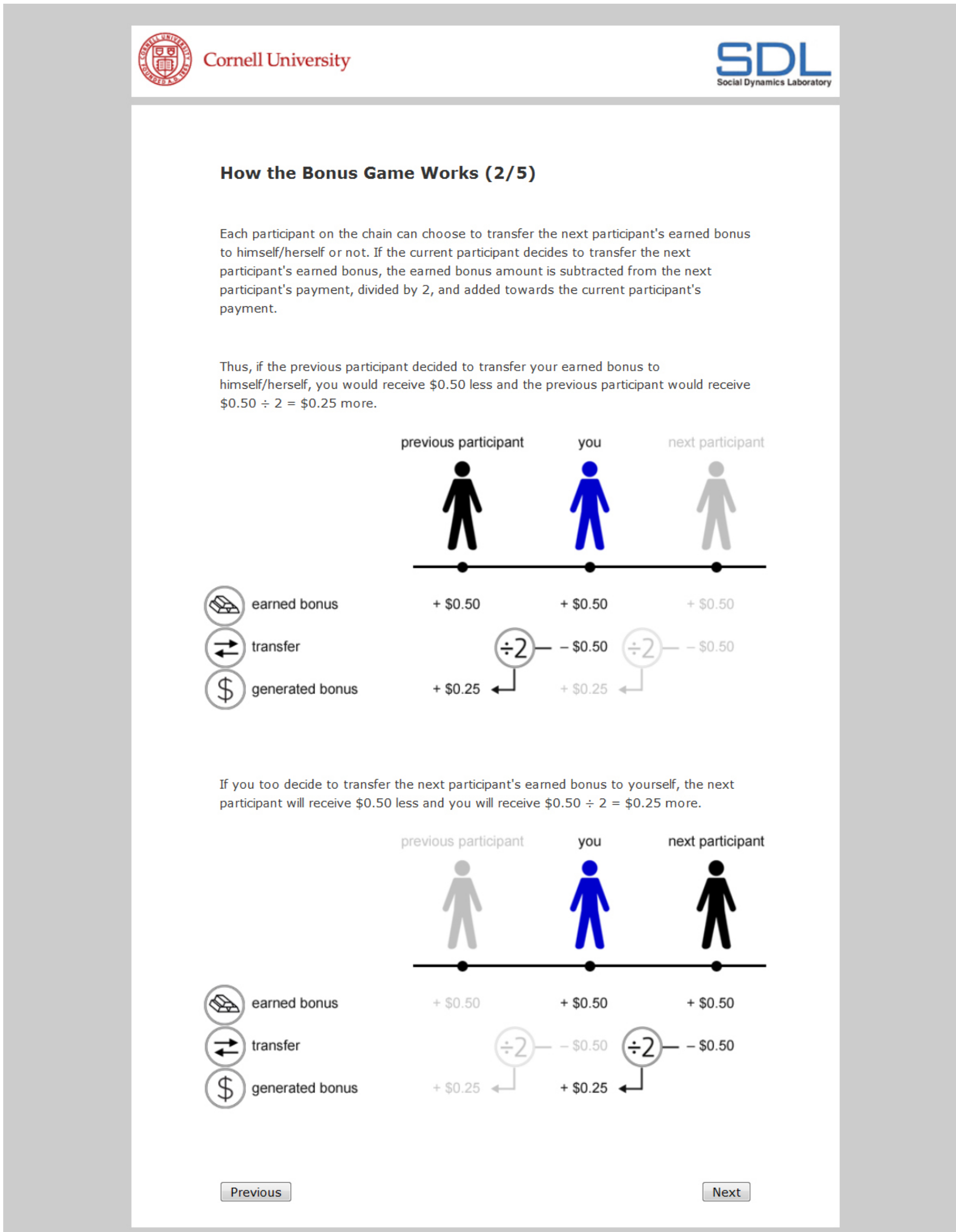


Figure C.5: **Page 2 of instructions for the Bonus Game.**

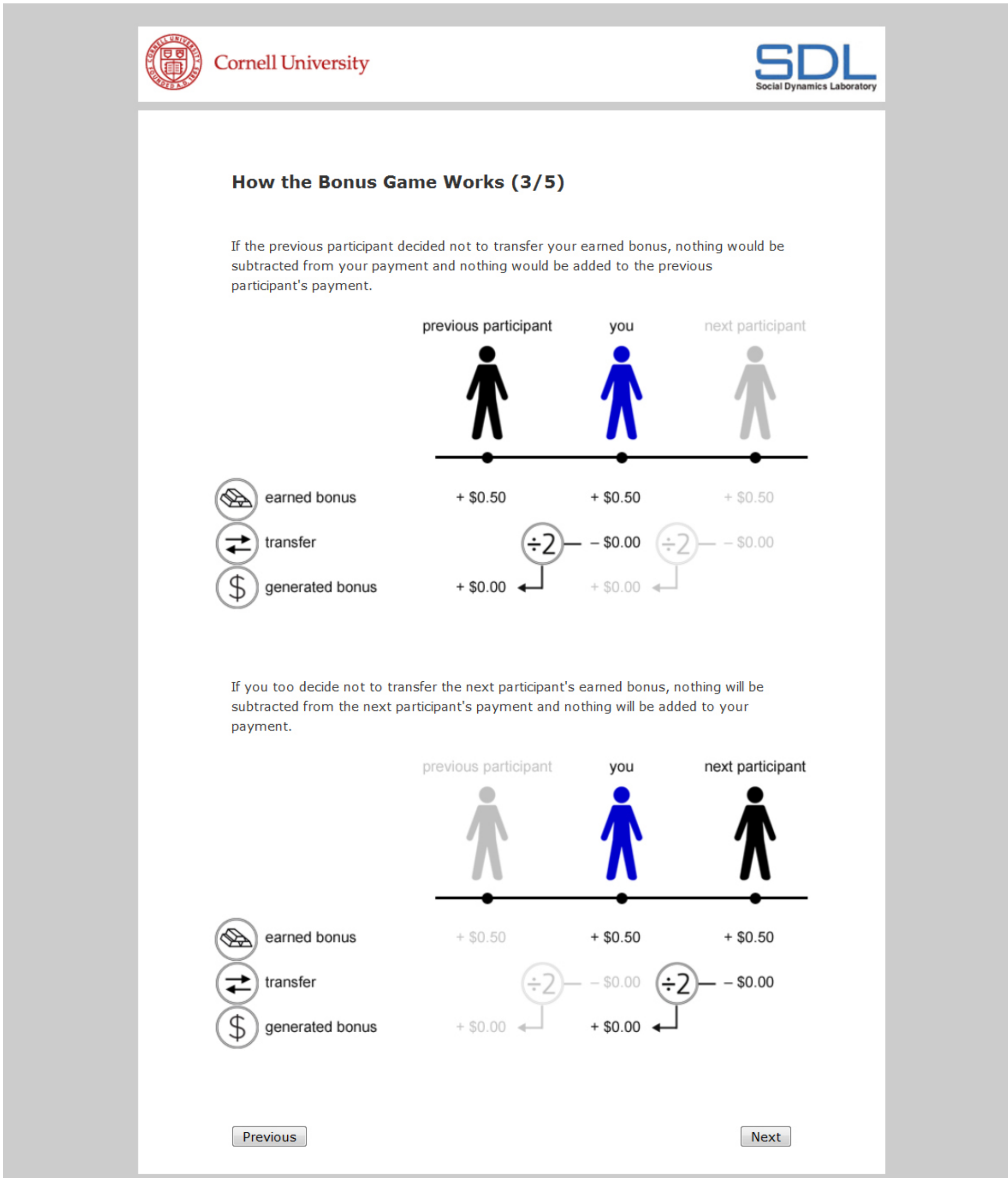


Figure C.6: **Page 3 of instructions for the Bonus Game.**

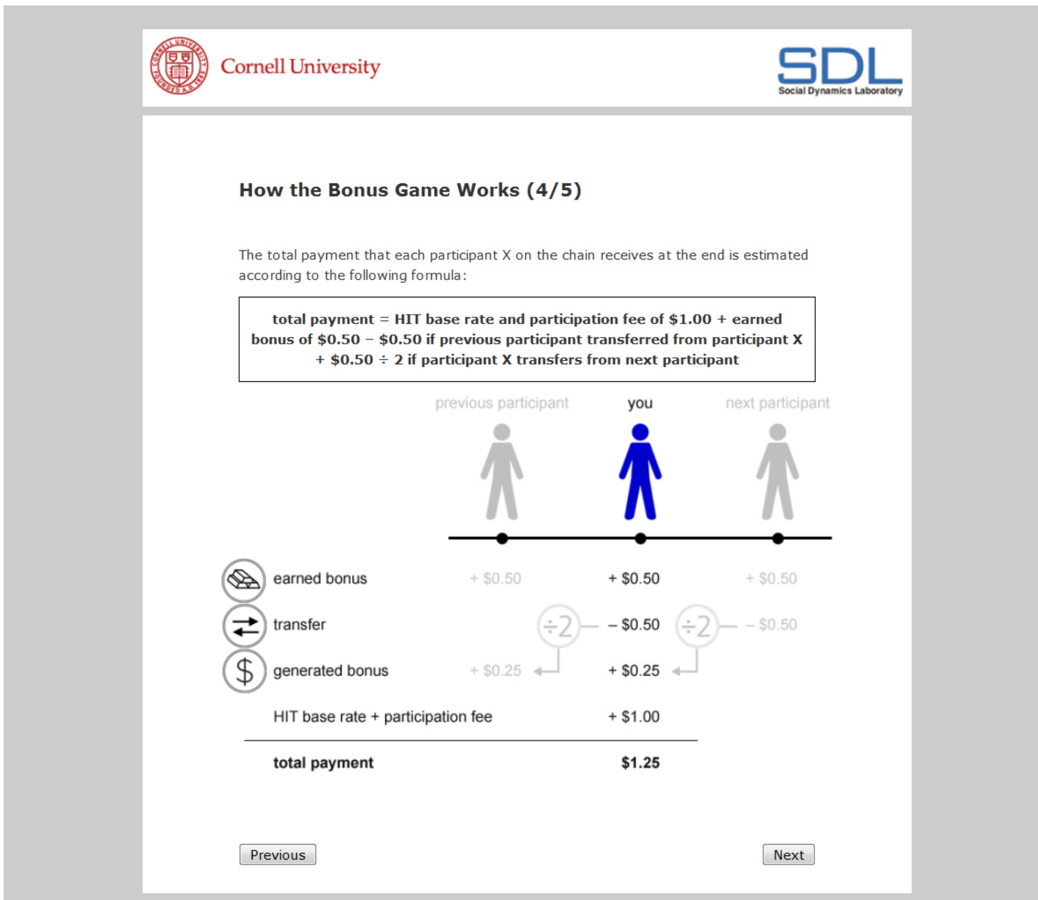


Figure C.7: **Page 4 of instructions for the Bonus Game.**

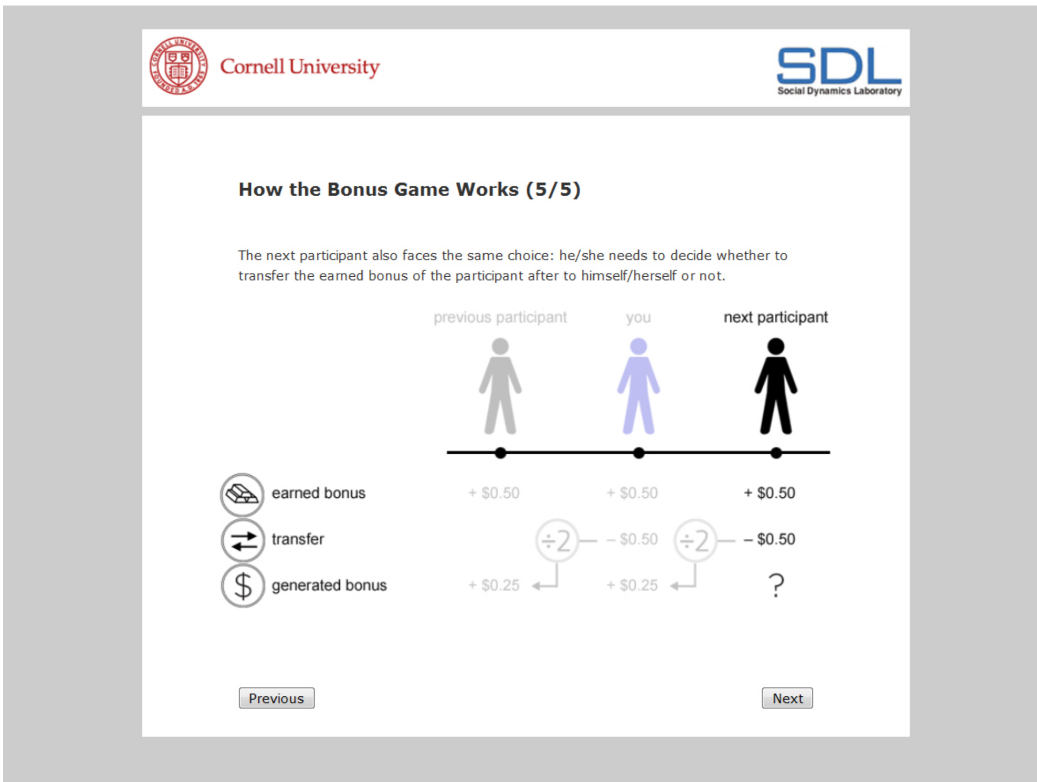
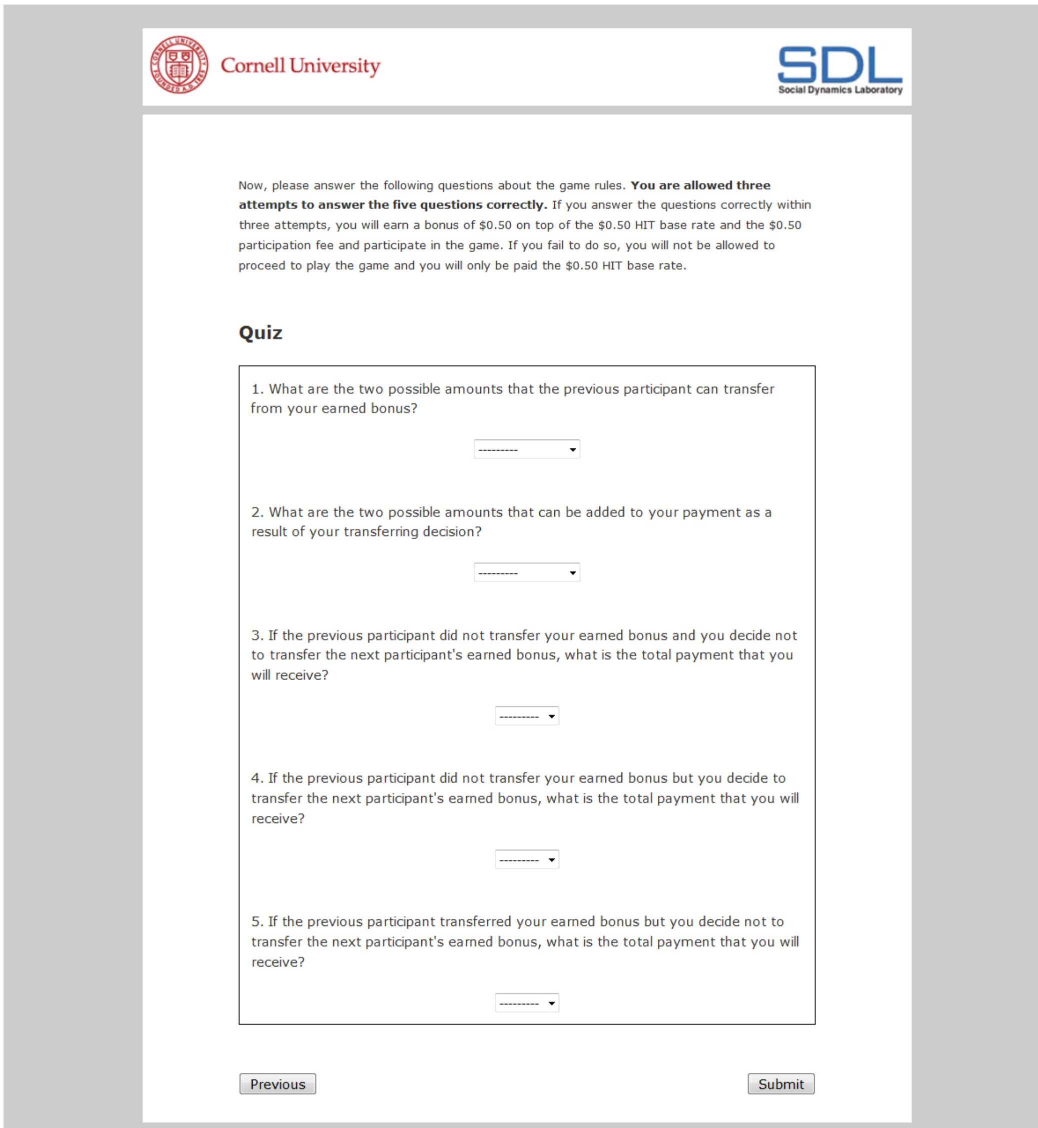


Figure C.8: **Page 5 of instructions for the Bonus Game.**



Cornell University

SDL
Social Dynamics Laboratory

Now, please answer the following questions about the game rules. **You are allowed three attempts to answer the five questions correctly.** If you answer the questions correctly within three attempts, you will earn a bonus of $0.50 on top of the $0.50 HIT base rate and the $0.50 participation fee and participate in the game. If you fail to do so, you will not be allowed to proceed to play the game and you will only be paid the $0.50 HIT base rate.

Quiz

1. What are the two possible amounts that the previous participant can transfer from your earned bonus?

----

2. What are the two possible amounts that can be added to your payment as a result of your transferring decision?

----

3. If the previous participant did not transfer your earned bonus and you decide not to transfer the next participant's earned bonus, what is the total payment that you will receive?

----

4. If the previous participant did not transfer your earned bonus but you decide to transfer the next participant's earned bonus, what is the total payment that you will receive?

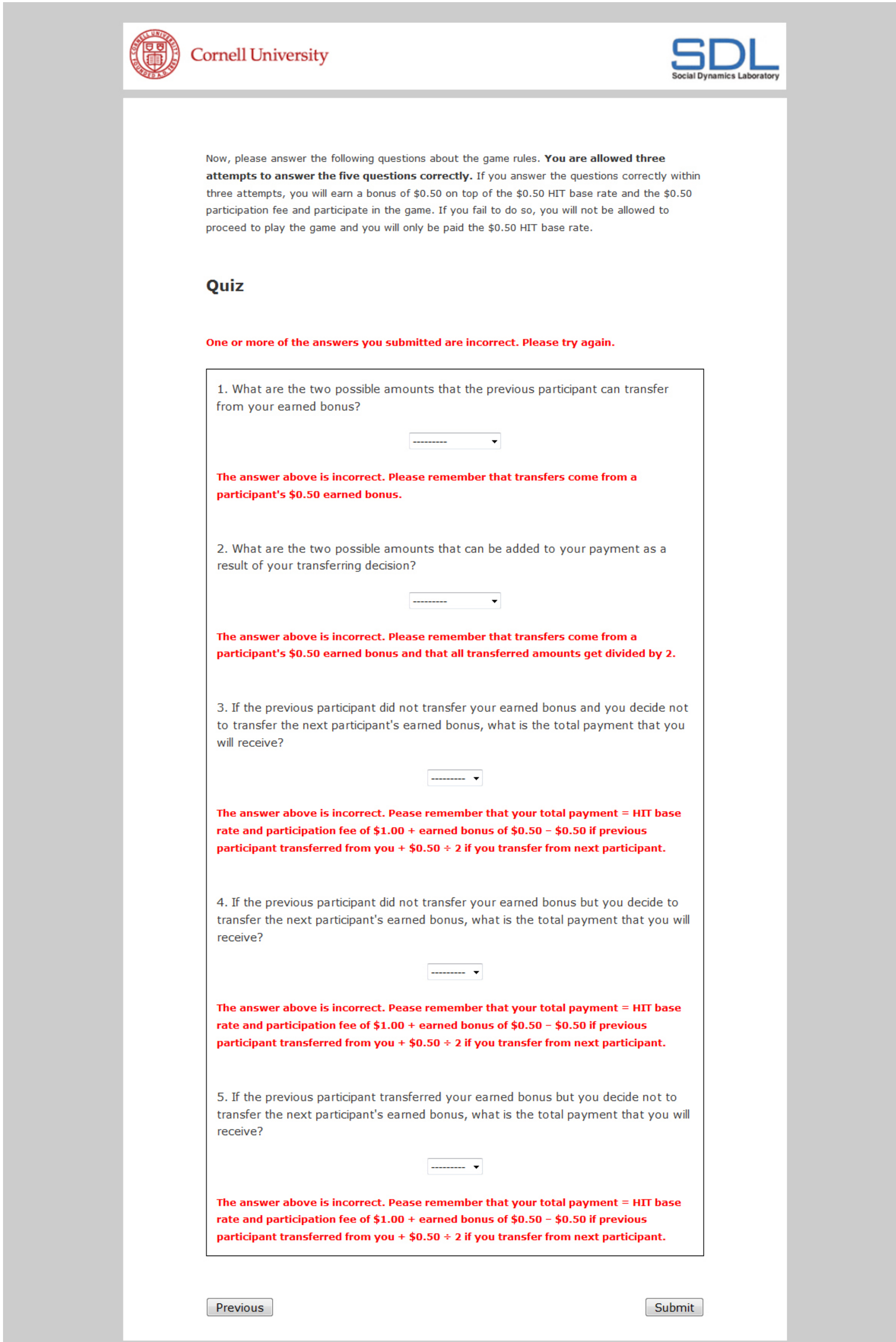
----

5. If the previous participant transferred your earned bonus but you decide not to transfer the next participant's earned bonus, what is the total payment that you will receive?

Previous

Submit

Figure C.9: **Quiz for the Bonus Game.**



Cornell University

SDL
Social Dynamics Laboratory

Now, please answer the following questions about the game rules. **You are allowed three attempts to answer the five questions correctly.** If you answer the questions correctly within three attempts, you will earn a bonus of $0.50 on top of the $0.50 HIT base rate and the $0.50 participation fee and participate in the game. If you fail to do so, you will not be allowed to proceed to play the game and you will only be paid the $0.50 HIT base rate.

Quiz

One or more of the answers you submitted are incorrect. Please try again.

1. What are the two possible amounts that the previous participant can transfer from your earned bonus?

The answer above is incorrect. Please remember that transfers come from a participant's $0.50 earned bonus.

2. What are the two possible amounts that can be added to your payment as a result of your transferring decision?

The answer above is incorrect. Please remember that transfers come from a participant's $0.50 earned bonus and that all transferred amounts get divided by 2.

3. If the previous participant did not transfer your earned bonus and you decide not to transfer the next participant's earned bonus, what is the total payment that you will receive?

The answer above is incorrect. Pease remember that your total payment = HIT base rate and participation fee of $1.00 + earned bonus of $0.50 – $0.50 if previous participant transferred from you + $0.50 ÷ 2 if you transfer from next participant.

4. If the previous participant did not transfer your earned bonus but you decide to transfer the next participant's earned bonus, what is the total payment that you will receive?

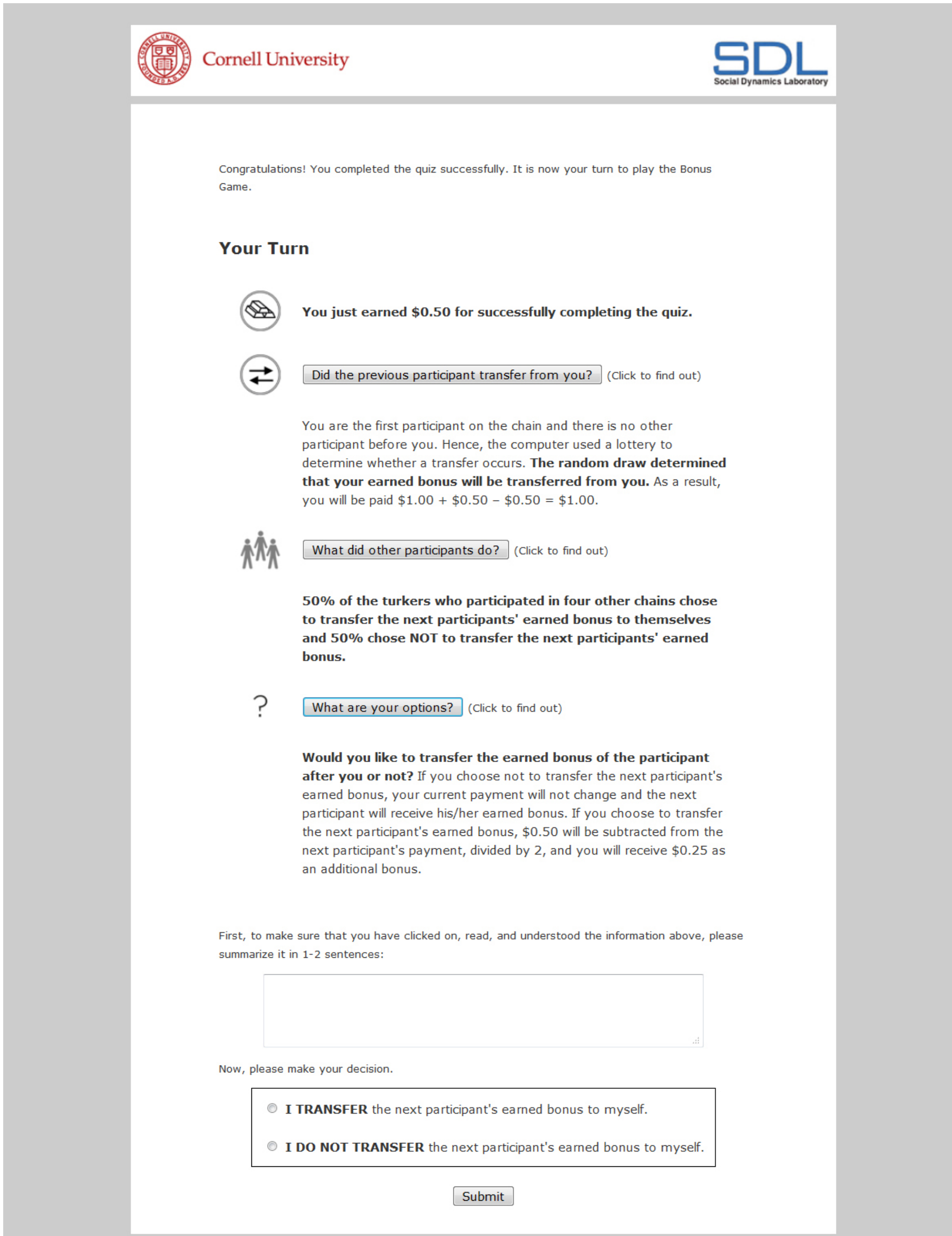
The answer above is incorrect. Pease remember that your total payment = HIT base rate and participation fee of $1.00 + earned bonus of $0.50 – $0.50 if previous participant transferred from you + $0.50 ÷ 2 if you transfer from next participant.

5. If the previous participant transferred your earned bonus but you decide not to transfer the next participant's earned bonus, what is the total payment that you will receive?

The answer above is incorrect. Pease remember that your total payment = HIT base rate and participation fee of $1.00 + earned bonus of $0.50 – $0.50 if previous participant transferred from you + $0.50 ÷ 2 if you transfer from next participant.

Previous Submit

Figure C.10: **Additional clues for wrong answers to the quiz for the Bonus Game.**


Cornell University

SDL
Social Dynamics Laboratory

Congratulations! You completed the quiz successfully. It is now your turn to play the Bonus Game.

Your Turn

You just earned $0.50 for successfully completing the quiz.

Did the previous participant transfer from you? (Click to find out)

You are the first participant on the chain and there is no other participant before you. Hence, the computer used a lottery to determine whether a transfer occurs. **The random draw determined that your earned bonus will be transferred from you.** As a result, you will be paid $1.00 + $0.50 – $0.50 = $1.00.

What did other participants do? (Click to find out)

50% of the turkers who participated in four other chains chose to transfer the next participants' earned bonus to themselves and 50% chose NOT to transfer the next participants' earned bonus.

?

What are your options? (Click to find out)

Would you like to transfer the earned bonus of the participant after you or not? If you choose not to transfer the next participant's earned bonus, your current payment will not change and the next participant will receive his/her earned bonus. If you choose to transfer the next participant's earned bonus, $0.50 will be subtracted from the next participant's payment, divided by 2, and you will receive $0.25 as an additional bonus.

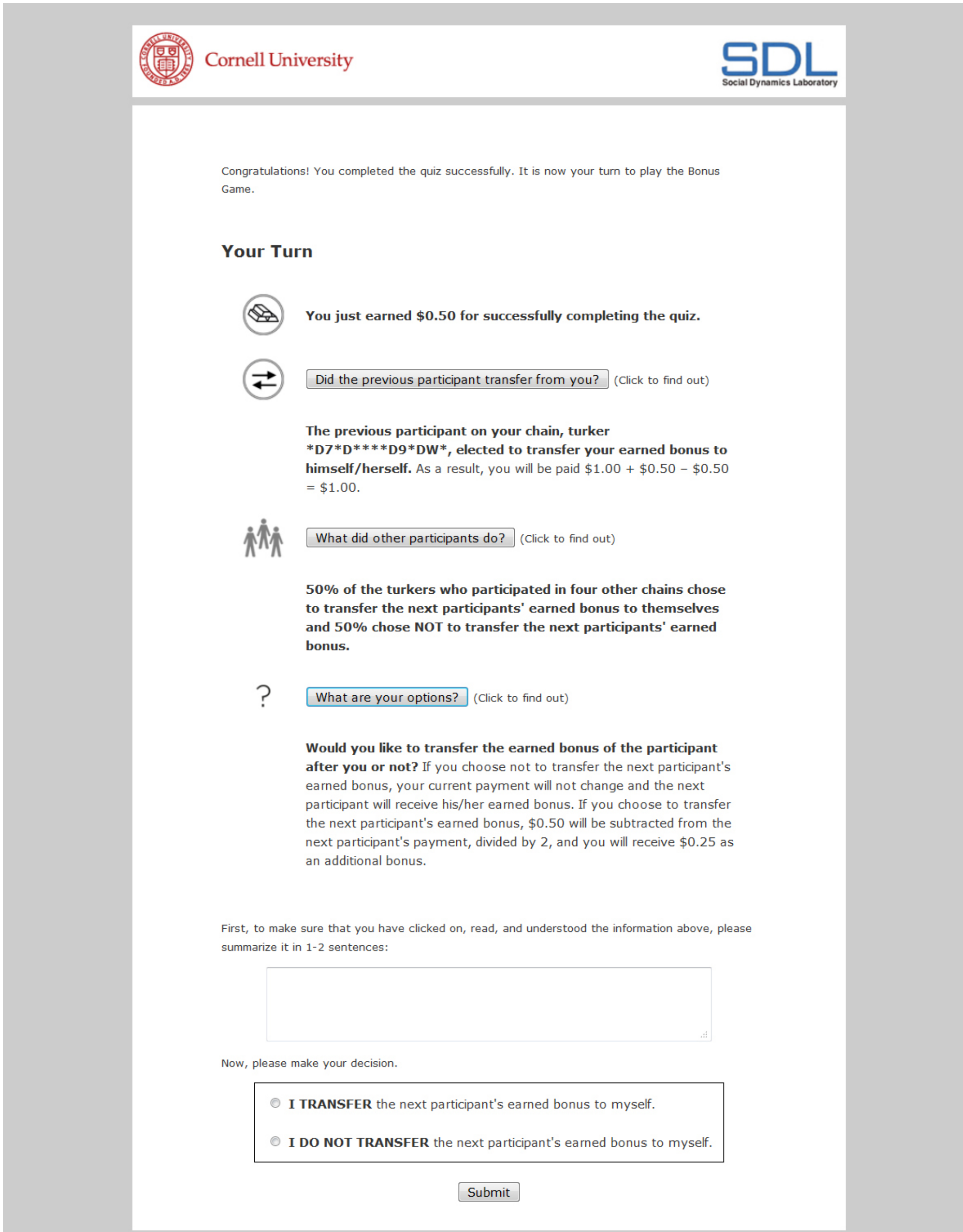
First, to make sure that you have clicked on, read, and understood the information above, please summarize it in 1-2 sentences:

Now, please make your decision.

- **I TRANSFER** the next participant's earned bonus to myself.
- **I DO NOT TRANSFER** the next participant's earned bonus to myself.

Submit

Figure C.11: **Decision page for seed in the observation condition in the Bonus Game.**



Cornell University

SDL
Social Dynamics Laboratory

Congratulations! You completed the quiz successfully. It is now your turn to play the Bonus Game.

Your Turn

You just earned $0.50 for successfully completing the quiz.

Did the previous participant transfer from you? (Click to find out)

The previous participant on your chain, turker *D7*D**D9*DW*, elected to transfer your earned bonus to himself/herself.** As a result, you will be paid $1.00 + $0.50 − $0.50 = $1.00.

What did other participants do? (Click to find out)

50% of the turkers who participated in four other chains chose to transfer the next participants' earned bonus to themselves and 50% chose NOT to transfer the next participants' earned bonus.

?

What are your options? (Click to find out)

Would you like to transfer the earned bonus of the participant after you or not? If you choose not to transfer the next participant's earned bonus, your current payment will not change and the next participant will receive his/her earned bonus. If you choose to transfer the next participant's earned bonus, $0.50 will be subtracted from the next participant's payment, divided by 2, and you will receive $0.25 as an additional bonus.

First, to make sure that you have clicked on, read, and understood the information above, please summarize it in 1-2 sentences:

Now, please make your decision.

- **I TRANSFER** the next participant's earned bonus to myself.
- **I DO NOT TRANSFER** the next participant's earned bonus to myself.

Submit

Figure C.12: **Decision page for link in the observation condition in the Bonus Game.**

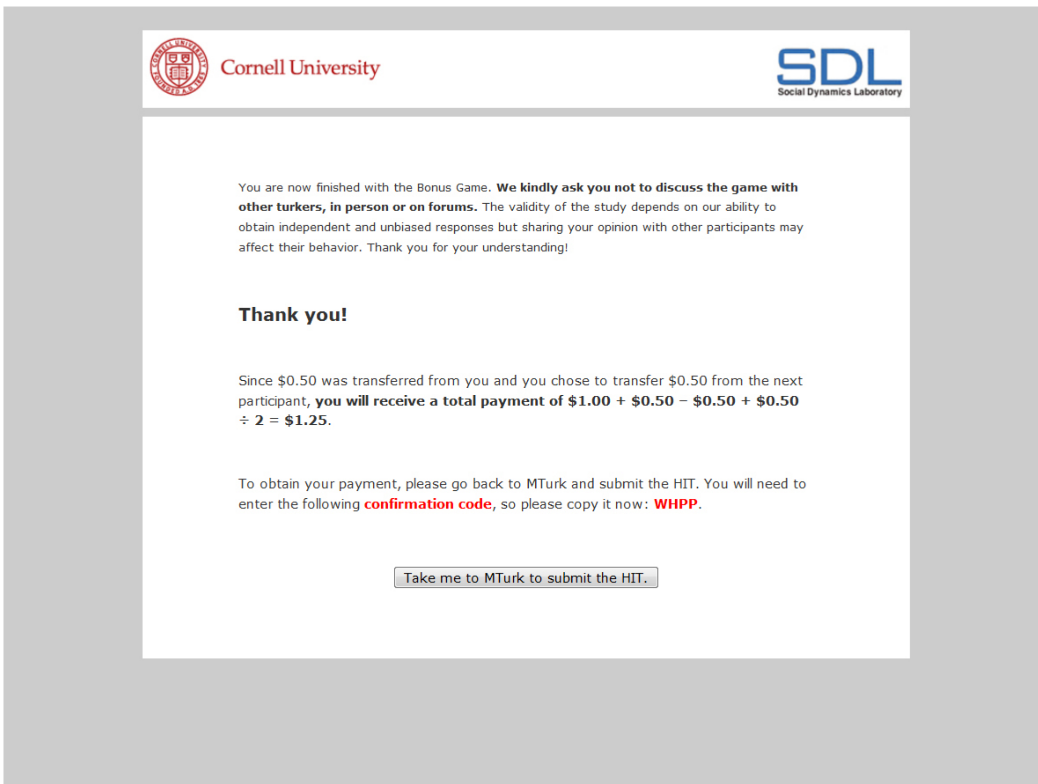


Figure C.13: **Final page in the Bonus Game.**

C.2 ADDITIONAL ANALYSES

To improve the internal validity of the study, we required participants to answer correctly five multiple-choice questions that tested their comprehension of the game rules. Participants were allowed three attempts (two mistakes) to answer the quiz in order to be able to participate. The quiz required simple mathematical operations and was thus also intended to convince participants that they have earned their payment, rather than received it as a gift, with the ultimate goal to strengthen the incentives. 1,198 AMT users attempted to answer the quiz and 438 did not manage to do so within the allowed number of attempts (failure rate of 36.6%).

Figure C.14 shows the power analysis used to determine the number of seeds (and hence, chains) in the game. The total sample size from the test represents the desired number of seeds in the no-observation condition and half of the desired number of seeds in the observation condition (since we wanted to test the effects of both low and high observation). The test suggested about 150 chains for a power level of around 0.9, assuming a transfer level of 50% in the no-observation condition and a relatively large effect size from observation.

Table C.1 shows the distribution of participants in the observation/no-observation and seed/link treatments. Table C.2 tests for a difference in the effect of experiencing a loss between seeds and links. The difference is not statistically significant ($\chi^2 = 0.88$, $p = 0.349$).

Table C.3 shows detailed demographics for the participant sample. Participants had a mean age of 30.4 (ranging from 18 to 67), were 37.3% female, with a median household income of $40,000–$49,999. The sample consisted of 72.3%

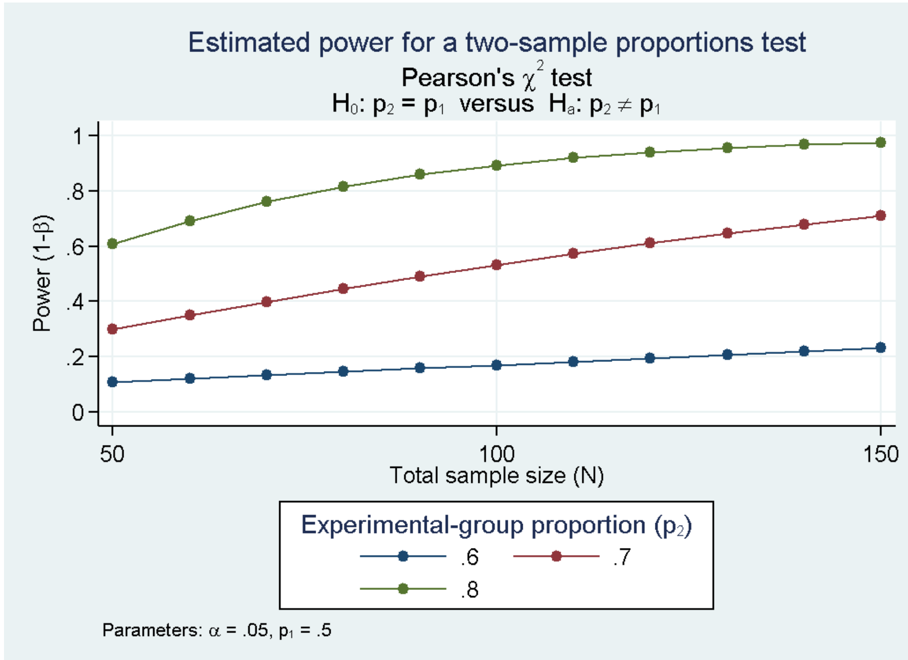


Figure C.14: **Power analysis for the number of seeds in the Bonus Game.** The analysis is for a two-sample proportions test assuming a transfer proportion of 0.5 in the no-observation condition and a significance level of 0.05.

US citizens and 24.2% Indian citizens, the remaining being from other countries. The most common ethnicities were 60.9% white and 25.1% Asian. 22.8% reported being non-religious and 19.3% atheists, while Christianity was the most common religion. 7.1% reported educational attainment of high school or less, 32.8% some college or Associate's degree, and 60.2% Bachelor's or graduate degree.

Table C.4 uses the demographic data to predict the log-odds that the participant transfers.

Table C.1: **Number of participants by experimental treatment.**

		Seed	Links	Total
No observation		50	200	250
Observation	Low observation	50	200	250
	High observation	50	200	250
Total		150	600	750

Table C.2: **Differences between links and seeds in the log odds of transfer from the next participant.**

	Coefficient
Loss	1.542
	(0.837)
Link	−0.518
	(0.470)
Link × Loss	−0.060
	(0.909)
Constant	0.598
	(0.375)
Number of observations	250
LR χ^2	3 df, 21.58***

Two-sided tests: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
The table reports coefficients and standard errors (in brackets) from logistic regression for participants in the no-observation treatment. Results do not show a statistically significant difference in the effect of experiencing a loss between seeds and links.

Table C.3: **Detailed demographics for the participant sample (*N* = 750).**

Characteristic	Mean / Percent
Female	37.33
Age	30.42 (*SD* = 8.80)
Income	
Less than $10,000	16.69
$10,000 – $19,999	12.89
$20,000 – $29,999	13.84
$30,000 – $39,999	13.70
$40,000 – $49,999	10.31
$50,000 – $59,999	6.65
$60,000 – $69,999	6.65
$70,000 – $79,999	4.48
$80,000 – $89,999	3.53
$90,000 – $99,999	3.93
$100,000 – $149,999	5.29
$150,000 or more	2.04
Education	
Less than high school	0.27
High school or GED	6.82
Some college	25.80
Associate’s degree	6.95
Bachelor’s degree	43.18
Graduate degree (Master’s, Doctorate, etc.)	16.98
Nationality	
United States	72.29
India	24.23
Other	3.48
Ethnicity	
White, non-Hispanic	60.86
Asian-Pacific Islander	25.07
African-American	3.89
Hispanic	3.08
Native American	0.67
Other	6.43
Religion	
Non-religious	22.82
Atheist	19.33
Hindu	19.06
Protestant	10.87
Roman Catholic	10.47
Other Christian	10.74
Muslim	2.28
Jewish	1.48
Buddhist	1.21
Other non-Christian	1.74

Table C.4: **Demographic differences in log odds of transfer.**

Characteristic	**Coefficient** **s.e.**
Age	−0.009
	(0.009)
Female	−0.617***
	(0.167)
Income	0.000
	(0.000)
Education: Associate's or some college	0.157
	(0.330)
Education: Bachelor's or graduate degree	0.292
	(0.321)
Religion: Hindu	0.934*
	(0.407)
Religion: other non-Christian	0.153
	(0.356)
Religion: non-religious	−0.280
	(0.222)
Religion: atheist	−0.267
	(0.234)
Nationality: India	−0.662
	(0.442)
Nationality: other	0.222
	(0.464)
Ethnicity: Asian or Pacific Islander	0.219
	(0.310)
Ethnicity: other non-White	−0.076
	(0.260)
Constant	0.893**
	(0.327)
Number of observations	745
LR χ^2	13 df, 29.65**

Two-sided tests: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
The table reports coefficients and standard errors (in brackets) from logistic regression with a baseline of a thirty-year-old white American Christian male with high-school education or less and household income of less than $10,000. Results show that females are less prone to antisocial behavior.

BIBLIOGRAPHY

Axelrod, Robert. 1984. *The Evolution of Cooperation*. New York, NY: Basic Books.

Bartlett, Monica Y. and David DeSteno. 2006. "Gratitude and prosocial behavior." *Psychological Science* 17:319–325.

Bearman, Peter. 1997. "Generalized exchange." *American Journal of Sociology* 102:1383–1415.

Ben-Ner, Avner, Louis Putterman, Fanmin Kong, and Dan Magan. 2004. "Reciprocity in a two-part Dictator game." *Journal of Economic Behavior & Organization* 53:333–352.

Biely, Christoly, Klaus Dragosits, and Stefan Thurner. 2007. "The Prisoner's Dilemma on co-evolving networks under perfect rationality." *Physica D: Nonlinear Phenomena* 228:40–48.

Bikhchandani, Sushil, David Hirshleifer, and Ivo Welch. 1992. "A theory of fads, fashion, custom, and cultural change as informational cascades." *Journal of Political Economy* 100:992–1026.

Bohannon, John. 2011. "Social science for pennies." *Science* 334:307.

Bond, Robert M., Christopher J. Fariss, Jason J. Jones, Adam D. I. Kramer, Cameron Marlow, Jaime E. Settle, and James H. Fowler. 2012. "A 61-million-person experiment in social influence and political mobilization." *Nature* 489:295–298.

Boyd, Robert and Peter J. Richerson. 1989. "The evolution of indirect reciprocity." *Social Networks* 11:213–236.

Centola, Damon. 2010. "The spread of behavior in an online social network experiment." *Science* 329:1194–1197.

Christakis, Nicholas A. and James H. Fowler. 2009. *Connected: The Surprising Power of Our Social Networks and How They Shape Our Lives*. New York, NY: Little, Brown and Company.

Cialdini, Robert B. 2008. *Influence: Science and Practice*. Boston, MA: Allyn & Bacon.

Cialdini, Robert B., Raymond R. Reno, and Carl A. Kallgren. 1990. "A focus theory of normative conduct: Recycling the concept of norms to reduce littering in public places." *Journal of Personality and Social Psychology* 58:1015–1026.

Coleman, James. 1990. *Foundations of Social Theory*. Cambridge, MA: Belknap Press.

Darley, John M. and Bibb Latané. 1968. "Bystander intervention in emergencies: Diffusion of responsibility." *Journal of Personality and Social Psychology* 8:377–383.

Deutsch, Morton and Harold B. Gerard. 1955. "A study of normative and informational social influences upon individual judgment." *Journal of Abnormal and Social Psychology* 51:629–636.

Diekmann, Andreas. 1985. "Volunteer's Dilemma." *Journal of Conflict Resolution* 29:605–610.

Dodds, Peter Sheridan, Kameron Decker Harris, Isabel M. Kloumann, Catherine A. Bliss, and Christopher M. Danforth. 2011. "Temporal patterns of happiness and information in a global social network: Hedonometrics and Twitter." *PLoS ONE* 6:e26752.

Eguíluz, Víctor M., Martín G. Zimmermann, Maxi San Miguel, and Camilo J. Cela-Conde. 2005. "Cooperation and the emergence of role differentiation in the dynamics of social networks." *American Journal of Sociology* 110:977–1008.

Ekeh, Peter P. 1974. *Social Exchange Theory: The Two Traditions*. Cambridge, MA: Harvard University Press.

Elster, Jon. 1989. *Nuts and Bolts for the Social Sciences*. Cambridge; New York: Cambridge University Press.

Fagan, Jeffrey, Deanna L. Wilkinson, and Garth Davies. 2007. "Social contagion of violence." In *The Cambridge Handbook of Violent Behavior*, edited by Daniel Flannery, Alexander Vazsonyi, and Irwin Waldman, pp. 688–723. Rochester, NY: Cambridge University Press.

Falk, Armin and Urs Fischbacher. 2002. "'Crime' in the lab – detecting social interaction." *European Economic Review* 46:859–869.

Fehl, Katrin, Daniel J. van der Post, and Dirk Semmann. 2011. "Co-evolution of behaviour and social network structure promotes human cooperation." *Ecology Letters* 14:546–551.

Fehr, Ernst and Herbert Gintis. 2007. "Human motivation and social cooperation: Experimental and analytical foundations." *Annual Review of Sociology* 33:43–64.

Fischbacher, Urs, Simon G achter, and Ernst Fehr. 2001. "Are people conditionally cooperative? Evidence from a public goods experiment." *Economics Letters* 71:397–404.

Fiske, Alan Page and Tage Shakti Rai. 2014. *Virtuous Violence*. Cambridge University Press.

Fowler, James H. and Nicholas A. Christakis. 2010. "Cooperative behavior cascades in human social networks." *Proceedings of the National Academy of Sciences* 107:5334–5338.

Gintis, Herbert. 2000. "Strong reciprocity and human sociality." *Journal of Theoretical Biology* 206:169–179.

Gouldner, Alvin W. 1960. "The norm of reciprocity: A preliminary statement." *American Sociological Review* 25:161–178.

Granovetter, Mark. 1978. "Threshold models of collective behavior." *American Journal of Sociology* 83:1420–1443.

Granovetter, Mark and Roland Soong. 1983. "Threshold models of diffusion and collective behavior." *Journal of Mathematical Sociology* 9:165–179.

Granovetter, Mark and Roland Soong. 1986. "Threshold models of interpersonal effects in consumer demand." *Journal of Economic Behavior & Organization* 7:83–99.

Greiner, Ben and M. Vittoria Levati. 2005. "Indirect reciprocity in cyclical networks: An experimental study." *Journal of Economic Psychology* 26:711–731.

Hamburger, Henry, Melvin Guyer, and John Fox. 1975. "Group size and cooperation." *Journal of Conflict Resolution* 19:503–531.

Hanaki, Nobuyuki, Alexander Peterhansl, Peter S. Dodds, and Duncan J. Watts. 2007. "Cooperation in evolving social networks." *Management Science* 53:1036–1050.

Hedström, Peter. 2005. *Dissecting the Social: On the Principles of Analytical Sociology*. Cambridge University Press.

Hedström, Peter and Petri Ylikoski. 2010. "Causal mechanisms in the social sciences." *Annual Review of Sociology* 36:49–67.

Helbing, Dirk and W. Yu. 2009. "The outbreak of cooperation among success-driven individuals under noisy conditions." *Proceedings of the National Academy of Sciences* 106:3680–3685.

Hoobler, Jenny M. and Daniel J. Brass. 2006. "Abusive supervision and family undermining as displaced aggression." *Journal of Applied Psychology* 91:1125–1133.

Horton, John J., David G. Rand, and Richard J. Zeckhauser. 2011. "The online laboratory: conducting experiments in a real labor market." *Experimental Economics* 14:399–425.

Isaac, R. Mark and James M. Walker. 1988. "Group size effects in public goods provision: The voluntary contributions mechanism." *The Quarterly Journal of Economics* 103:179–199.

Jones, Edward E and Victor A Harris. 1967. "The attribution of attitudes." *Journal of Experimental Social Psychology* 3:1–24.

Jones, Marshall B. and Donald R. Jones. 2000. "The contagious nature of antisocial behavior." *Criminology* 38:25–46.

Jordan, Jillian J., David G. Rand, Samuel Arbesman, James H. Fowler, and Nicholas A. Christakis. 2013. "Contagion of cooperation in static and fluid social networks." *PLoS ONE* 8:e66199.

Karau, Steven J. and Kipling D. Williams. 1993. "Social loafing: A meta-analytic review and theoretical integration." *Journal of Personality and Social Psychology* 65:681–706.

Keizer, Kees, Siegwart Lindenberg, and Linda Steg. 2008. "The spreading of disorder." *Science* 322:1681–1685.

Keizer, Kees, Siegwart Lindenberg, and Linda Steg. 2013. "The importance of demonstratively restoring order." *PLoS ONE* 8:e65137.

Kollock, Peter. 1999. "The economies of online cooperation: Gifts and public goods in cyberspace." In *Communities in Cyberspace*, edited by Marc A. Smith and Peter Kollock, pp. 220–242. London: Routledge.

Kurzban, Robert and Daniel Houser. 2005. "Experiments investigating cooperative types in humans: A complement to evolutionary theory and simulations." *Proceedings of the National Academy of Sciences* 102:1803–1807.

Lawler, Edward J. 2001. "An affect theory of social exchange." *American Journal of Sociology* 107:321–352.

Lawson, Chappell, Gabriel S. Lenz, Andy Baker, and Michael Myers. 2010. "Looking like a winner: Candidate appearance and electoral success in new democracies." *World Politics* 62:561–593.

Loftin, Colin. 1986. "Assaultive violence as a contagious social process." *Bulletin of the New York Academy of Medicine* 62:550–555.

López-Pintado, Dunia and Duncan J. Watts. 2008. "Social influence, binary decisions and collective dynamics." *Rationality and Society* 20:399–443.

Macy, Michael W. 1991. "Chains of cooperation: Threshold effects in collective action." *American Sociological Review* 56:730–747.

Macy, Michael W. and Andreas Flache. 2002. "Learning dynamics in social dilemmas." *Proceedings of the National Academy of Sciences* 99:7229–7236.

Macy, Michael W. and Andreas Flache. 2009. "Social dynamics from the bottom up: Agent-based models of social interaction." In *The Oxford Handbook of Analytical Sociology*, pp. 245–268. Oxford: Oxford University Press.

Macy, Michael W. and Milena Tsvetkova. 2013. "The signal importance of noise." *Sociological Methods & Research* p. 0049124113508093.

Macy, Michael W. and Robert Willer. 2002. "From factors to actors: Computational sociology and agent-based modeling." *Annual Review of Sociology* 28:143–166.

Malinowski, Bronislaw. 1920. "51. Kula; The circulating exchange of valuables in the archipelagoes of Eastern New Guinea." *Man* 20:97–105.

Mallough, Ryan. 2013. "An extra-large sized order of generosity." Avaialable: `http://www2.macleans.ca/2013/01/16/an-extra-large-sized-order-of-generosity/`. Accessed 10 October 2012.

Mao, Andrew, Yiling Chen, Krzysztof Z. Gajos, David C. Parkes, Ariel D. Procaccia, and Haoqi Zhang. 2012. "TurkServer: Enabling synchronous and longitudinal online experiments." In *Workshops at the Twenty-Sixth AAAI Conference on Artificial Intelligence*.

Mason, Winter and Siddharth Suri. 2012. "Conducting behavioral research on Amazon's Mechanical Turk." *Behavior Research Methods* 44:1–23.

Meloni, S., A. Buscarino, L. Fortuna, M. Frasca, J. Gómez-Gardeñes, V. Latora, and Y. Moreno. 2009. "Effects of mobility in a population of Prisoner's Dilemma players." *Physical Review E* 79:067101.

Memmott, Mark. 2013. "55 Customers pay for next car's order at Mass. doughnut shop." Available: `http://www.npr.org/blogs/thetwo-way/2013/07/15/202365926/55-customers-pay-for-next-cars-order-at-mass-doughnut-shop`. Accessed 8 October 2013.

Molm, Linda D. 2010. "The structure of reciprocity." *Social Psychology Quarterly* 73:119–131.

Molm, Linda D., Jessica L. Collett, and David R. Schaefer. 2007. "Building solidarity through generalized exchange: A theory of reciprocity." *American Journal of Sociology* 113:205–242.

Muchnik, Lev, Sinan Aral, and Sean J. Taylor. 2013. "Social influence bias: A randomized experiment." *Science* 341:647–651.

Murphy, Kate. 2013. "Ma'am, your burger has been paid for." Available: `http://www.nytimes.com/2013/10/20/opinion/sunday/maam-your-burger-has-been-paid-for.html`. Accessed 20 October 2013.

Nowak, Martin A. and Sébastien Roch. 2007. "Upstream reciprocity and the evolution of gratitude." *Proceedings of the Royal Society B* 274:605–610.

Nowak, Martin A. and Karl Sigmund. 2005. "Evolution of indirect reciprocity." *Nature* 437:1291–1298.

Oliver, Pamela, Gerald Marwell, and Ruy Teixeira. 1985. "A theory of the critical mass. I. Interdependence, group heterogeneity, and the production of collective action." *American Journal of Sociology* 91:522–556.

Oliver, Pamela E. 1993. "Formal models of collective action." *Annual Review of Sociology* 19:271–300.

Oliver, Pamela E. and Gerald Marwell. 1988. "The paradox of group size in collective action: A theory of the critical mass. II." *American Sociological Review* 53:1–8.

Paolacci, Gabriele, Jesse Chandler, and Panagiotis G. Ipeirotis. 2010. "Running experiments on Amazon Mechanical Turk." *Judgment and Decision Making* 5:411–419.

Patihis, Lawrence, Steven J. Frenda, Aurora K. R. LePort, Nicole Petersen, Rebecca M. Nichols, Craig E. L. Stark, James L. McGaugh, and Elizabeth F. Loftus. 2013. "False memories in highly superior autobiographical memory individuals." *Proceedings of the National Academy of Sciences* 110:20947–20952.

Pfeiffer, Thomas, Claudia Rutte, Timothy Killingback, Michael Taborsky, and Sebastian Bonhoeffer. 2005. "Evolution of cooperation by generalized reciprocity." *Proceedings of the Royal Society B* 272:1115–1120.

Rand, David G., Samuel Arbesman, and Nicholas A. Christakis. 2011. "Dynamic social networks promote cooperation in experiments with humans." *Proceedings of the National Academy of Sciences* 108:19193–19198.

Rand, David G. and Martin A. Nowak. 2013. "Human cooperation." *Trends in Cognitive Sciences* 17:413–425.

Raub, Werner and Jeroen Weesie. 1990. "Reputation and efficiency in social interactions: An example of network effects." *The American Journal of Sociology* 96:626–654.

Rogers, Everett M. 2003. *Diffusion of Innovations*. New York: Free Press.

Ross, Joel, Lilly Irani, M. Six Silberman, Andrew Zaldivar, and Bill Tomlinson. 2010. "Who are the crowdworkers? Shifting demographics in Mechanical Turk." In *Proceedings of the 28th of the international conference extended abstracts on Human factors in computing systems*, CHI EA '10, pp. 2863–2872, Atlanta, GA. ACM. Avaialable: `http://doi.acm.org/10.1145/1753846.1753873`. Accessed 10 October 2012.

Ross, Lee. 1977. "The intuitive psychologist and his shortcomings: Distortions in the attribution process." *Advances in Experimental Social Psychology* 10:173–220.

Salganik, Matthew J., Peter Sheridan Dodds, and Duncan J. Watts. 2006. "Experimental study of inequality and unpredictability in an artificial cultural market." *Science* 311:854–856.

Sapolsky, Robert M. 2006. "Culture in animals: The case of a non-human primate culture of low aggression and high affiliation." *Social Forces* 85:217–233.

Schelling, Thomas C. 1971. "Dynamic models of segregation." *The Journal of Mathematical Sociology* 1:143–186.

Scott, John Finley. 1971. *Internalization of Norms: A Sociological Theory of Moral Commitment*. Englewood Cliffs, NJ: Prentice Hall Press.

Seinen, Ingrid and Arthur Schram. 2006. "Social status and group norms: Indirect reciprocity in a repeated helping experiment." *European Economic Review* 50:581–602.

Stanca, Luca. 2009. "Measuring indirect reciprocity: Whose back do we scratch?" *Journal of Economic Psychology* 30:190–202.

Suri, Siddharth and Duncan J. Watts. 2011. "Cooperation and contagion in web-based, networked public goods experiments." *PLoS ONE* 6:e16836.

Sutherland, Edwin Hardin and David Luckenbill. 1992. *Principles of Criminology*. Dix Hills, N.Y.: General Hall.

Taylor, Michael. 1987. *The Possibility of Cooperation*. Cambridge, UK: Cambridge University Press.

Trivers, Robert L. 1971. "The evolution of reciprocal altruism." *The Quarterly Review of Biology* 46:35–57.

Tsvetkova, Milena and Vincent Buskens. 2013. "Coordination on egalitarian networks from asymmetric relations in a social game of Chicken." *Advances in Complex Systems* 16:1350005.

Tsvetkova, Milena and Michael W. Macy. 2014a. "The science of 'paying it forward'." *The New York Times* (March 14).

Tsvetkova, Milena and Michael W. Macy. 2014b. "The social contagion of generosity." *PLoS ONE* 9:e87275.

Tsvetkova, Milena and Michael W. Macy. 2015. "The social contagion of antisocial behavior." *Sociological Science* 2:36–49s.

Tversky, Amos and Daniel Kahneman. 1991. "Loss aversion in riskless choice: A reference-dependent model." *The Quarterly Journal of Economics* 106:1039–1061.

Uehara, Edwina. 1990. "Dual exchange theory, social networks, and informal social support." *American Journal of Sociology* 96:521–557.

Valente, Thomas W. 1996. "Social network thresholds in the diffusion of innovations." *Social Networks* 18:69–89.

van de Pol, Martijn and Jonathan Wright. 2009. "A simple method for distinguishing within- versus between-subject effects using mixed models." *Animal Behaviour* 77:753–758.

Weber, J. Mark and J. Keith Murnighan. 2008. "Suckers or saviors? Consistent contributors in social dilemmas." *Journal of Personality and Social Psychology* 95:1340–1353.

Wedekind, Claus and Manfred Milinski. 2000. "Cooperation through image scoring in humans." *Science* 288:850–852.

Wenner, Melinda. 2007. "Humans carry more bacterial cells than human ones." Available: `http://www.scientificamerican.com/article/strange-but-true-humans-carry-more-bacterial-cells-than-human-ones/` Accessed 1 January 2015.

Widom, Cathy S. 1989. "Does violence beget violence? A critical examination of the literature." *Psychological Bulletin* 106:3–28.

Wilson, James Q. and George L. Kelling. 1982. "Broken windows." *Atlantic Monthly* 249:29–38.

Yamagishi, Toshio and Karen S. Cook. 1993. "Generalized exchange and social dilemmas." *Social Psychology Quarterly* 56:235–248.

Yamagishi, Toshio and Toko Kiyonari. 2000. "The group as the container of generalized reciprocity." *Social Psychology Quarterly* 63:116–132.

Yamagishi, Toshio, Nobuhito Jin, and Toko Kiyonari. 1999. "Bounded generalized reciprocity: Ingroup boasting and ingroup favoritism." *Advances in Group Processes* 16:161–97.

Young, H. Peyton. 2009. "Innovation diffusion in heterogeneous populations: Contagion, social influence, and social learning." *The American Economic Review* 99:1899–1924.