ABSTRACT

Title of Dissertation: THE EMERGENCE OF SYMBOLIC NORMS

Xinyue Pan, Doctor of Philosophy, 2023

Dissertation directed by: Professor, Michele J. Gelfand, Stanford Graduate School of Business and Department of Psychology, Stanford University
Professor, Dana S. Nau, Department of Computer Science and Institute for Systems Research, University of Maryland, College Park

Human groups are experts in developing and maintaining social norms. Many social norms have clear practical functions, such as regulating competition or facilitating coordination. Some other norms, however, have arbitrary functions and limited direct material consequences for the self or the group, but are nevertheless enforced. I define such norms as *symbolic norms*.

Symbolic norms are prevalent across human societies. Given the discrepancy between the social importance and the functional opacity of these norms, it is important to understand how a seemingly neutral behavior can emerge as a symbolic norm and be adopted by the population. In this dissertation, I argue that a neutral behavior is more likely to evolve as a symbolic norm when it shows statistical correlation with a practical behavior on the population level. I call this the *norm*

spillover effect. The norm spillover effect predicts that if, on the population level, followers of a practically beneficial norm happen to conduct a certain neutral behavior more often than practical norm violators, the social norm will spill over from the practical domain to the neutral domain. Thus, people will adopt and enforce that neutral behavior, and a symbolic norm will emerge.

This dissertation uses agent-based models and an empirical experiment to test the norm spillover effect across two levels of analyses. First, agent-based models are used to test the evolutionary force behind the norm spillover effect on the population level. I argue that the statistical correlation between a practical and a neutral behavior creates an ecology that fosters symbolic norm following and enforcement. Second, an empirical experiment is conducted to examine the psychology of the norm spillover effect on the individual level. I argue that the perceived correlation between a practical and a neutral behavior increases the perceived direct function of and the pressure to conform to the symbolic norm.

THE EMERGENCE OF SYMBOLIC NORMS

by

Xinyue Pan

Dissertation submitted to the Faculty of the Graduate School of the University of Maryland, College Park, in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
2023

Advisory Committee:
Dr. Michele J. Gelfand, Chair
Dr. Dana S. Nau, Co-chair
Dr. Paul J. Hanges, Committee Member
Dr. James A. Grand, Committee Member
Dr. Rellie Derfler-Rozin, Dean's Representative

Dedication

To my parents, who showed me the beauty of science.

Acknowledgements

First and foremost, I would like to express my deepest gratitude to my advisors, Drs. Michele Gelfand and Dana Nau. Thank you for your unwavering support and guidance throughout my Ph.D. journey. Your belief in me has been a constant source of my courage. Thank you for leading me to the realm of interdisciplinary research, which has become my passion. Thank you for being my steadfast pillar and safety net, both in academics and in life.

I would also like to express my sincere appreciation to my dissertation committee, Drs. Paul Hanges, James Grand, and Rellie Derfler-Rozin. Thank you for your insightful comments on this project and consistent support for my dissertation and career. I could not have achieved this milestone without your help.

I would like to thank Vincent Hsiao, who has provided numerous valuable suggestions on the current model and on all my modeling work. I would also like to thank Dr. Soham De, who taught me Python and agent-based modeling, which has enabled me to do the current work that I love.

I am grateful to Dr. Ying Lin, Basti Weitz, colleagues in the Culture Lab at Stanford University, Alisha Tronetti, Alex Butler, Ruoxi Li, Dr. Hengduo Li, Alex Davis, Virginia Choi, and Huixian Yu, for their suggestions on experimental design, draft writing, simulation data collection, and experimental materials. I would also like to thank the participants in my pilot studies, who have provided valuable feedback on the experimental materials. I would also like to thank Archie Tablada, Chitra Lekhram, Joanne Leffson-Bryant, Koda Smith, and Tony Chan, for their administrative support. I acknowledge the University of Maryland supercomputing resources (http://hpcc.umd.edu/) made available for conducting the simulations reported in this dissertation.

I would also like to extend my heartfelt gratitude to Dr. Shihui Han and colleagues in the Culture and Social Cognitive Neuroscience Lab at Peking University. Thank you for believing in me and supporting me during the challenging time. You are the reason why I can start this Ph.D. journey. I would also like to thank Dr. Lihua Mao, the instructor of my first psychology course, Introduction to Psychology. Thank you for making me fall in love with psychology.

I would like to thank Dr. Chutian Wang, my steadfast comrade both in work and in life. Your unwavering understanding, encouragement, and support have been my anchor throughout this process.

Last but certainly not least, I extend my deepest gratitude to my parents. Thank you for showing me the beauty of science and fostering my love for the world. Your endless love has been the cornerstone of my academic pursuits. I am forever indebted to you.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

Human groups are experts in developing and maintaining social norms. Starting from 3 years old, children show clear understanding of the normative structure of conventional activities and spontaneously infer the presence of social norms (Rakoczy et al., 2008; Schmidt et al., 2016). From the etiquette of politeness to the regulations on traffic rules, from the food sharing systems in hunter-gatherer societies (Kameda et al., 2003) to the intricate agreements on international relations, social norms serve as the cornerstone of human interaction and shape human behaviors on all levels.

Many social norms have clear direct and practical functions. People respect each other's property rights to avoid ownership conflicts. They form queues to facilitate fairness and orderliness. They keep quiet in libraries to avoid distracting others. They drive on the same side of the road to ensure safety and efficiency. These norms, whether written or unwritten, serve clear direct functions for the group, such as regulating competition over resources, maintaining cooperation, and facilitating coordination (Gelfand et al., 2017).

However, there are also some other norms that have arbitrary functions and limited direct material consequences for the self or the group. For example, many societies have rules about dress codes, hair styles, or what ornaments to wear. Certain words or hand gestures can be considered as an offensive behavior in some cultures (Chandler & Schwarz, 2009). Many societies have food taboos that require them to eat only certain animals (Navarrete & Fessler, 2003). These norms seem to have arbitrary material consequences on oneself or others, but are usually treated as important and even sometimes imbued with great meaning (Köster et al., 2022). I define such norms as *symbolic norms*.

Symbolic norms are prevalent across human societies. While many of these norms are harmless, some others can be costly or even hazardous (Sterelny, 2014), with their practical functions remaining unclear. Given the discrepancy between the social importance and the functional opacity of these norms, it is important to understand how a seemingly neutral behavior can emerge as a symbolic norm and be adopted by the population.

In this dissertation, I use agent-based models and experiments to begin to shed light on this puzzle. Broadly, I hypothesize that a neutral behavior is more likely to evolve as a symbolic norm when it shows statistical correlation with a practical behavior on the population level, even if the neutral and the practical behaviors do not have a causal relationship. I call this the *norm spillover effect*. The norm spillover effect predicts that if, on the population level, followers of a practically beneficial norm (e.g., cooperation norm) happen to conduct a certain neutral behavior more often than practical norm violators, the social norm will spill over from the practical domain to the neutral domain. Thus, people will adopt and enforce that neutral behavior, and a symbolic norm will emerge.

I argue that the norm spillover effect can be interpreted from two levels of analyses. On the population level, the statistical correlation between a practical and a neutral behavior creates an ecology that fosters symbolic norm following and enforcement. On the individual level, the perceived correlation between a practical and a neutral behavior increases the perceived direct function of and the pressure to conform to the symbolic norm.

First, agent-based models are used to test the norm spillover effect on the population level and examine the emerging dynamics of it. In the model, agents perform two kinds of game-theoretic interactions with each other: one representing their practical behavior, and the other representing their symbolic choice. These two kinds of game-theoretic interactions will be referred to as interactions in *Domain I* and *Domain II*, respectively. *Domain I* is a *practical* domain. Following the norm in *Domain I* directly benefits others' well-being (i.e., increases others' payoffs). *Domain II* is a *symbolic* domain. Agents' behaviors in *Domain II* do not directly influence anyone's payoff. An agent can choose to conduct any combination of behaviors in the two domains and enforce any norm. The population then goes through an evolutionary process, in which agents interact with each other and update their strategies. I hypothesize that although an agent's behavior in the symbolic domain has no direct impact on anyone's payoff, when there is an initial correlation between the behaviors in the two domains on the population level, agents will adopt and enforce the symbolic choice that initially co-occurs with the practically beneficial behavior and the symbolic norm will emerge.

While the norm spillover effect accounts for the emergence of symbolic norms on the population level, such effect needs to be supported by concurrent psychological processes that can make the norm spillover effect happen. To understand these processes, I move to empirical experiments on human participants given that agent-based models are not able to address these questions. In the second part of the dissertation, an online experiment is conducted to examine the psychology of the norm spillover effect on the individual level. I hypothesize that the norm spillover effect is realized in psychological processes through two pathways. First is through the perceived direct function of the symbolic norm. I hypothesize that the perceived correlation increases the perceived direct function of the symbolic norm and thus increases its adoption. Second is through norm enforcement. I hypothesize that when a symbolic choice correlates with a practically beneficial behavior on the population level, performing the symbolic choice signals good traits in the performer so that people are more likely to select symbolic norm followers as partners. Such enforcement from others motivates people to adopt the symbolic norm themselves.

The rest of the dissertation is structured in the following way: The rest of Chapter 1 introduces theories that explain the existence and maintenance of symbolic norms on the population level. Chapter 2 introduces the agent-based models that examine the norm spillover effect on the population level. Chapter 3 introduces the psychological processes that enable the norm spillover effect to happen on the individual level and then describes the empirical experiment on the psychological processes. In Chapter 4, I conclude with a discussion and some future directions.

<u>Clarification of Concepts</u>

The Function of a Norm

A criterion to identify a symbolic norm is the lack of practical function of a behavior. However, although it is straightforward to create a theoretically unfunctional norm in a laboratory setting, it is hard to identify symbolic norms in real life. On the one hand, it is possible that a seemingly unfunctional behavior has an unknown practical function. For example, among some Latin American groups, there is a norm of adding a handful of ash when preparing maize. Although people in these groups cannot explain the function of this norm, adding ash (or limestone) turned out to increases the body's ability to absorb niacin, which helps prevent pellagra (Köster et al., 2020). In this case, the norm may seem symbolic to its practitioners, but it is practical. On the other hand, it is also possible that a norm is unfunctional, but people perceive functions in it. For example, in the Southwest of Nigeria, there is a taboo that people must not use bare hands to collect rainwater. People believe that it will avoid being struck by the thunder, although that's not the case scientifically (Omobola, 2013). In fact, logically, the function of a behavior is never falsifiable. In other words, it's hard to define any real-life social norm as purely symbolic because it is always possible that the function is unknown yet.

Given the complexity of defining a symbolic norm in the first place, this dissertation does not focus on what is or is not a symbolic norm. Instead, I ask the question from another angle: Given a behavior with a relatively ambiguous function, what makes it more likely to emerge as a symbolic norm? This is a testable question because the function of the symbolic norm is not an absolute criterion. Instead, we can hold the function of a behavior constant, and explore the factors that influence its emergence as a symbolic norm. The reason why we need the function of the symbolic norm to be relatively ambiguous is that, if a behavior has such a clear function that people have strong reasons to follow it, other factors will have little impact on its emergence. Nevertheless, theoretically, the norm spillover effect should be generalizable to both practical and unpractical norms, a point I will return in Chapter 4.

Defining a Norm

The term “norm” has been used to refer to a variety of different constructs. On the personal perspective, norms can be interpreted as “the feelings of moral obligation to perform or refrain from specific action” (Schwartz & Howard, 1984, p. 234). On the injunctive perspective, norms refer to “what most others approve or disapprove” (Cialdini et al., 1990, p. 1015). On the descriptive perspective, norms describe “what most others do” (Cialdini et al., 1990, p. 1015). On the social perspective, norms are “rules and standards that are understood by members of a group, and that guide and/or constrain social behavior without the force of laws,” which emphasizes its difference from formal institutions (Cialdini & Trost, 1998, p. 152). From an interactive perspective, norms are patterns of behaviors that “everyone conforms, everyone is expected to conform, and everyone wants to conform when they expect everyone else to conform” (Young, 2015, p. 359).

In agent-based models, since it is not plausible to assess individuals’ attitudes or expectations, norms are usually measured in a descriptive way on the population level. Nevertheless, there are still differences in definitions. Some work defines norm as a behavior that is costly to follow but benefit the group (Axelrod, 1986). In this scenario, norm itself has a valence (i.e., it is beneficial). On the contrary, in some other work, norms are descriptive, and simply refer to what the majority of the population does (De et al., 2018). In this case, the emergence of a norm describes that some behavior becomes “more characteristic (e.g., more uniform) of some sociocultural collective unit than of individuals observed at random” (Pepitone, 1976, p. 642).

In this dissertation, I measure a norm from two perspectives. First, I capture the descriptive perspective of a norm. Under this definition, when a symbolic norm emerges, it means that the behavior becomes more prevalent than its alternatives in the population. Second, I also capture the enforcement perspective of a norm. Under this definition, a symbolic norm emerges when the deviation from it is disapproved of by others. Such disapproval can either be a direct punishment, or an indirect enforcement by avoiding interacting with the violators, as social ostracism is universally used as an expression of disapproval and informal sanction (Eriksson et al., 2021).

Theories on Symbolic Norms on the Population Level

In this section, I summarize relevant theories that explain the existence and maintenance of symbolic norms. While "symbolic norm" may not be a common term in the existing literature, numerous theories explore similar constructs from diverse perspectives, including fields such as religious studies, evolutionary biology, anthropology, computer science, anthropology, and social and developmental psychology. In this section, I highlight four theories of them, including costly signaling, the evolution of ethnic markers, silly rule models, and cultural transmission, and show their connections with and differences from the new theory of norm spillover.

Costly Signaling

One branch of theories explain the existence of symbolic norms as costly signaling. An example of costly signaling is religion. The signaling theory of religion suggests that since performing religious belief, rituals, attire, and taboos are costly, those who are willing to incur time and energy to religious behaviors manifest their commitment and loyalty to the group (Sosis, 2009). If loyalty also indicates altruism, then costly religious behaviors can help individuals recognize reliable allies and thus benefit cooperation in a group.

Similar mechanisms have also been examined in evolutionary biology. The ornaments on a peacock's tail, for example, are metabolically costly, but signal the healthiness of their owner (Salahshour, 2019). A gazelle's leap is a wasteful move, but displays their enhanced fitness and escape abilities (Murray & Moore, 2009). For human beings, accordingly, if conducting a behavior signal one's good characters, people are more likely to conduct that behavior, and such signaling benefits the group (Murray & Moore, 2009). Evolutionary models have shown that costly signaling can be evolutionary stable and proliferate in a population (Gintis et al., 2001). And costly signaling and costly cooperation co-evolve (Salahshour, 2019).

The norm spillover effect is related to but different from the costly signaling theory. For a behavior to serve as a signal, there needs to be a *link* between the signal and the underlying qualities of the signaler. Take a gazelle's leap as an example, a link means that a fitter gazelle indeed has more energy to afford the leap. From a modeling perspective, such a link is usually built by implementing an internal trait that leads high-quality individuals to have different consequences after signaling (compared with low-quality individuals) (Gintis et al., 2001). However, the norm spillover effect indicates that such a link does not need to be an internal trait on the individual level. In fact, the symbolic norm can emerge when the correlation is on the population level.

Moreover, theories on costly signaling focus on how a costly behavior is maintained. They argue that for a signal to be reliable, there needs to be differential signaling costs or benefits for signalers. For instance, as peacocks have increasing tail size and ornamentation, it carries increasing costs that only the most fit peacocks can afford. Thus, the intensity of signaling will reach different equilibria for less fit and more fit individuals. In other words, a less fit peacock cannot fake having a big tail, and thus tail size can be a reliable signal for fitness. In another example, when nesting confraternal chicks want food from their parents, they beg as a signal for their need. Since more hungry chicks benefit more from food, those who are really hungry will beg more. Thus, begging becomes a reliable signal for hunger. Either way, a substantial cost of signaling is important for the signal to be reliable because otherwise individuals will fake the signal (Murray & Moore, 2009).

In contrast, the norm spillover effect focuses on which behavior can emerge as a signaling behavior in the first place. The emergence of a symbolic norm does not require the symbolic choice to be costly. In this sense, the norm spillover effect can be considered as an unsustainable case of costly signaling. As I show in the Results section, in the norm spillover effect, initially, an individual's symbolic choice may serve as a signal for their practical behavior. Then since performing the behavior is uncostly, the *dilemma of honest signaling* happens, in which the major population adopts the symbolic choice and the symbolic norm emerges. Eventually, one's symbolic choice can no longer serve as a signal and the symbolic norm fades away. The norm spillover effect covers the process from emergence to fading, and thus has a different focus from the cost signaling theories.

Symbolic Choice as an Ethnic Marker

Another branch of theories indicate that a symbolic norm may serve as an ethnic marker, which helps ingroup recognition and thus facilitates coordination and cooperation. Humans live in groups. Distinctive ethnolinguistic markers, symbols, clothing, and cultural rituals can serve as cues to detect group membership, so that members of a same group can better coordinate and limit social exchange to a small reliable ingroup members (Pan et al., 2023).

In a mathematical model on ethnic markers, the researchers assumed a population with individuals with different observable markers and playing a coordination game. They found that if individuals are predisposed to interact with similar others, and learn better-off others' markers and behaviors as a package, groups distinguished by both differences in practical norms and in arbitrary markers will emerge and remain stable (McElreath et al., 2003). In other words, different symbolic norms are chosen by different groups as a marker that differentiates them from each other. In another model, computer simulations also showed that when agents are categorized by visible tags and can make cooperation decisions based on the partners' tags, tag-based cooperation will emerge and facilitate cooperation in the population (Riolo et al., 2001).

Although an individual's symbolic choice can serve as an ethnic marker, the norm spillover effect generates different predictions from the ethnic marker theories. First, models of ethnic markers predict that a population will eventually divide into subgroups with different ethnic markers (McElreath et al., 2003). However, the norm spillover effect aims to explain why a neutral behavior eventually emerges as a global symbolic norm. Second, if an individual's symbolic choice serves as an ethnic marker, individuals should enforce their own symbolic choice because that indicates an ingroup membership (Hammond & Axelrod, 2006; McElreath et al., 2003; Riolo et al., 2001). However, as I will show in Section **The Norm Spillover Effect Is Different From Tag-based Cooperation**, under the norm spillover effect, even symbolic norm violators themselves may enforce the symbolic norm in some conditions. Such result is the opposite to the prediction of the ethnic marker theories. Therefore, the norm spillover effect examines a different mechanism from the theories on the evolution of ethnic markers.

Silly Rules Enhances Learning of Compliance and Norm Enforcement

Another theory argues that symbolic norms exist because it gives individuals more practice in learning rule enforcement. In an agent-based model, Koster et al. (2022) implemented a society where there is a centralized classification scheme that labels which behaviors are approved vs. transgressive. Individuals can punish each other and punishing the transgressors is rewarded by the system. The model found that when the classification scheme labels some unharmful behaviors as transgressive and encourages individuals to enforce them, it makes the group learn other beneficial behaviors more effectively. This is because adding an additional unfunctional rule (i.e., silly rule) to a normative system that already contains some practical rules helps individuals practice third-party norm enforcement, and thus benefits overall norm following in the population.

In this sense, the enforcement of symbolic norms can be considered as a practice of norm enforcement in general. In fact, literature on cultural tightness also indicates that when a society faces high ecological threat, strong social norms exist not only in domains that are critical for cultural coordination, but also in less functional domains as people bestow importance on following normative behavior in general (J. Jackson & Gelfand, 2016).

Nevertheless, the norm spillover effect is different from the model on silly rules. The model on silly rules assumes a *top-down* classification scheme that decides which symbolic choice is transgressive (Köster et al., 2022). However, the norm spillover effect focuses on the *bottom-up* process of norm emergence. In other words, there is no central institution that governs which symbolic choice should be the norm. Instead, the norm enforcement happens spontaneously without external rewards. Thus, although both theories are on symbolic norms, the norm spillover effect has a different assumption from the silly rule model.

Cultural Transmission and Social Learning

Another branch of theories explain the symbolic norms as a byproduct of cultural transmission. In cultural evolution, each generation adopts a large amount of behaviors from their prior generation and often does not understand why these behaviors are adaptive (Boyd et al., 2011; Derex et al., 2019).

Symbolic norms may be adopted because they are incorporated into a larger normative system that also includes other practical norms. Thus, when the generation learns from their prior generation, they learn both their practical and symbolic norms. Children tend to over-imitate adults—after seeing an action sequence, children copy not only the causally relevant, but also the perceivably unnecessary actions (Hoehl et al., 2019). They also show disapproval when seeing others fail to copy the unnecessary action and may protest with normative language (Kenward, 2012). In this sense, the symbolic norms may be maintained as a byproduct of adopting practical norms.

The norm spillover effect takes a different perspective from the theories on cultural transmission. Cultural transmission theories explain why, among all the norms in a society, some of them do not have clear functions (Boyd et al., 2011; Derex et al., 2019). However, the norm spillover effect aims to understand why, among all the behaviors with ambiguous functions, a specific behavior becomes the symbolic norm. Moreover, if a symbolic norm is inherited because it is embedded in a larger normative system, the symbolic norm should only be inherited when the other norms in the normative system are also inherited. However, as I will show in Section **The Impact of Practical Behavior on the Norm Spillover Effect**, the norm spillover model shows that when a symbolic choice is correlated with a practically beneficial behavior, the symbolic choice may become the norm even when the practically beneficial behavior itself is not adopted as a norm.

In summary, the norm spillover effect aims to examine how an uncostly neutral behavior proliferates in the population through a bottom-up process. This makes the norm spillover effect distinctive from previous literature on symbolic norms. In the rest of the dissertation, I use a series of agent-based models and an empirical experiment to illustrate the norm spillover effect.

Chapter 2: Evolutionary Game Theoretic Models

Methods

Introduction to the Evolutionary Game Theoretic Modeling Approach

In this first part of the dissertation, I use an agent-based model (ABM) based on an evolutionary game theoretic framework to test the norm spillover effect. An ABM is a handy tool to examine the emergence of social norms because it allows researchers to test their hypotheses through an "actor-thinking" approach (Macy & Willer, 2002). With an ABM, researchers implement their hypothesized intra- and interpersonal processes into the model and test whether the hypothesized mechanisms produce the predicted emerging states. If so, researchers can argue that the proposed process is one of the possible mechanisms behind the phenomenon of interest.

Moreover, an ABM can also be used to conduct "virtual experiments," in which researchers directly manipulate individual and environmental level factors and observe their impacts on model outcomes. Such virtual experiments can help make causal inferences among variables that are otherwise not easily manipulable in real life. Since norm emergence is a group-level process that evolves across time, it is difficult to conduct empirical experiments both time-wise and resource-wise. In this case, an ABM is a great exploratory first step to test hypotheses and identify the possible mechanisms behind group dynamics (Pan et al., 2021).

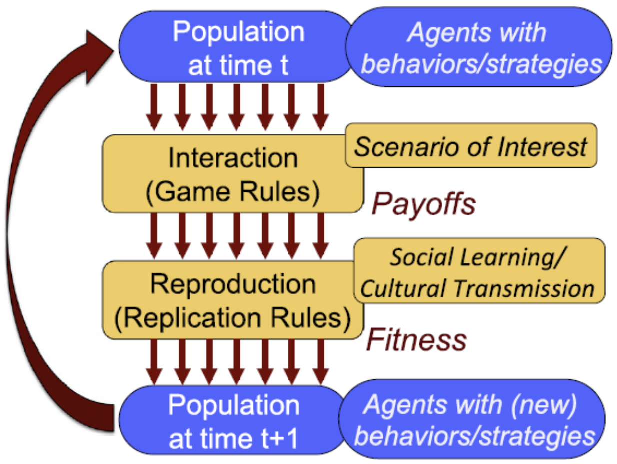
Specifically, this dissertation implements the ABM in an evolutionary game theoretic (EGT) framework. An EGT model is the application of population dynamical methods to game theory (Hofbauer & Sigmund, 2003). It was initially developed to model biological evolution but has been increasingly used to study the evolution of human behavior. In a typical agent-based EGT model, there is a population of agents, each of which has a strategy representing the behavioral options an individual can choose in interpersonal interactions. The population of agents interact in a formal game, the rule of which represents the interaction scenario that individuals are embedded in. From the interaction, agents gain payoffs, which represent the evolutionary fitness of different individuals (Sigmund & Nowak, 1999).

The evolutionary fitness related to different strategies decides whether a strategy will flourish or deteriorate in the population in the next generation. Strategies related to higher fitness are more likely to reproduce while strategies related to lower fitness are more likely to be replaced. In the context of human behavior, reproduction resembles cultural transmission or social learning (Cavalli-Sforza & Feldman, 1981), in which people are more likely to learn the behavior of those who are better-off.

After each interaction phase, there is a selection phase in which the agents' strategies are updated based on their payoffs. After the selection phase, agents interact again based on their new strategies. As this cycle repeats, the population dynamic is implemented to the game and the researchers can observe the trajectories of the proportions of different behaviors in the population. That way, an EGT can be used to study the emergence of group behaviors (see **Figure 1**).

An EGT framework is a great fit to studying norm emergence for the following reasons. First, an EGT model takes a "hypothesis-free" evolutionary perspective. Though some basic assumptions about human decision-making are still needed in an EGT model, it does not hypothesize which specific behavior an individual should choose or which norm should emerge. Instead, the researchers only need to put all the possible behavioral options in the strategy set and let evolution select which behavior proliferates. This makes it a handy tool to study bottom-up emergence of norms. Second, an EGT model focuses on behavioral changes on the population level. This makes it a perfect fit for studying social norms, which is also a group-level construct.

Figure 1 EGT Framework for Studying the Evolution of Behaviors



Note. From "Societal Threat and Cultural Variation in the Strength of Social Norms: An Evolutionary Basis," by P. Roos, M. Gelfand, D. Nau, and J. Lun, 2015, *Organizational Behavior and Human Decision*, *129*, p. 15. Copyright 2015 by Elsevier Inc.

Model Overview

In the agent-based EGT model of norm spillover, agents perform two kinds of game-theoretic interactions with each other: one representing their practical behavior, and the other representing their symbolic choice. These two kinds of game-theoretic interactions will be referred to as interactions in *Domain I* and *Domain II*, respectively.

Domain I is a practical domain. Following the norm in *Domain I* directly benefits others' well-being. The interaction in *Domain I* can be represented by a cooperation game. A cooperation game, also known as Prisoner's Dilemma, has been widely used to study norm compliance/violation. Similar to a norm following behavior, cooperation is a behavior that is costly to perform but benefits others (Axelrod, 1986). On the other hand, a non-cooperative behavior, or a defective behavior, is a violation of a practically functional norm.

Domain II is a symbolic domain. In *Domain II*, agents play a no-payoff game, in which they choose to conduct one of the two behavioral options, *X* or *Y*, but their choices will not directly influence anyone's payoff. In other words, the behavior in *Domain II* is symbolic. It does not have any first-order material consequence for one's own or other's well-being.

An agent can choose to conduct any combination of behaviors in the two domains and choose to enforce any behavior or combination of behaviors in *Domain I* and *Domain II*. The population then goes through an evolutionary process. If the norm spillover effect exists, I predict that when there is an initial correlation between the behaviors in *Domain I* and *Domain II* at the population level, a symbolic norm in *Domain II* will gradually evolve: Most agents will start to perform the symbolic choice that originally co-occurs more often with the cooperative behavior. In the following sections, I will specify the model setups.

Games and Agents' Strategies

In *Domain I*, agents play a pairwise cooperation game. There are two strategies an agent can choose—1) to cooperate (*C*) or 2) to defect (*D*). If an agent chooses to cooperate, they will spend a cost of c for their interaction partner to receive a benefit of b, where $b > c$. If an agent chooses to defect, they will spend no cost, and their partner will not receive any benefit, either (Nowak, 2006). Thus, a payoff matrix for a pairwise cooperation game can be represented by **Table 1**. By default, $c = 1$ and $b = 5$. In Section **The Impact of Practical Behavior on the Norm Spillover Effect**, I also explored how this payoff matrix influences the results. Although the payoff matrix does influence the strength of the symbolic norm, the norm spillover effect happens across a broad range of c and b.

Table 1 Payoff Matrix for the Cooperation Game

	Action of Agent B	
	Cooperate	Defect
Action of Agent A		
Cooperate	(A: $b - c$, B: $b - c$)	(A: $-c$, B: b)
Defect	(A: b, B: $-c$)	(A: 0, B: 0)

Note. b is the benefit one receives from the partner's cooperation, and c is the cost of cooperating.

In *Domain II*, agents play a no-payoff symbolic game. There are two strategies in this game—*X* or *Y*. Each agent can choose to perform either *X* or *Y* as their symbolic choice. No matter which one they choose, it does not influence anyone's payoff directly.

In addition to the strategies in these two games, agents also have a third strategy which decides how they enforce the norms. In the default model, agents enforce a norm by *ostracizing* others who conduct certain behavior(s).

There are ten options of *ostracism strategies* that include all the possible combinations of behaviors that one can ostracize. Specifically, they are 1) *NQ*, who never ostracizes any agents, 2) *CQ*, who ostracizes only cooperators, 3) *DQ*, who ostracizes only defectors, 4) *XQ*, who ostracizes only agents who perform *X* in the symbolic domain, 5) *YQ*, who ostracizes only the agents who perform *Y* in the symbolic domain, 6) *CXQ*, who ostracizes both cooperators and X-performers, 7) *CYQ*, who ostracizes both cooperators and Y-performers, 8) *DXQ*, who ostracizes both defectors and X-performers, 9) *DYQ*, who ostracizes both defectors and Y-performers, and 10) *AQ*, who ostracizes everyone (see **Table 2**).

Though some of these strategies seem irrational in the real world (e.g., *CQ* who ostracizes only cooperators, *AQ* who ostracizes everyone), all combinations are included to make the strategy set exhaustive, so that it is not biased toward any behavior (García & Traulsen, 2019). This is to make the hypothesis falsifiable—the model gives agents equal opportunities to ostracize any behavior, but let evolutionary process decide which behaviors proliferate. In other words, implementing strategies that are the opposite of my hypotheses gives the model a chance to fail, so that the evolution of symbolic norm is not built in the model set-up.

When *Agent A* ostracizes *Agent B*, *Agent A* will quit the cooperation game whenever the two agents are assigned to interact in a cooperation game. That way, neither party will have the chance to contribute to or receive anything from the interaction. Because there are two parties in a cooperation game, if either party quits, the interaction will cease.

How does an agent decide whether their interaction partner is a cooperator, defector, X-performer, or Y-performer? In the default model, I assume that an agent's symbolic strategy is *transparent*. That means, every agent knows everyone else's symbolic strategy even before they interact. Such set-up is based on the assumption that if a symbolic strategy serves as an ethnic marker or signal, it has to be easily visible. On the other hand, I assume that an agent's practical behavior (i.e., cooperative behavior) is not transparent. That means, an agent can only randomly guess another agent's cooperation strategy if they never interacted before, but once they've interacted in a cooperation game, they will know each other's cooperation and symbolic strategies, as people get to know each other through interactions.

In fact, the transparency of symbolic strategy is a key to the norm spillover effect. In Section **Transparency of Cooperation Strategy and Symbolic Strategy**, I will discuss the impacts of the transparency of the strategies. In Supplementary Section ***2.1 Partial Strategy Transparency***, I also explore other model options in which the strategies are between completely non-transparent and completely transparent.

In addition, I also explore another model option in which agents enforce a norm by *punishing* others who conduct certain behavior(s). As I will describe in Supplementary Section ***2.9 Norm Enforcement by Punishment***, the norm spillover effect generally holds when the norms are enforced by direct punishment, although the direction of the norm spillover effect depends on the effectiveness of punishment.

Table 2 Ostracism Strategies

Ostracism strategy	Description
NQ	Never ostracize any agents
CQ	Ostracize cooperators
DQ	Ostracize defectors
XQ	Ostracize agents who perform *X* in the symbolic domain
YQ	Ostracize agents who perform *Y* in the symbolic domain
CXQ	Ostracize both cooperators and X-performers
CYQ	Ostracize both cooperators and Y-performers
DXQ	Ostracize both defectors and X-performers
DYQ	Ostracize both defectors and Y-performers
AQ	Ostracize everyone

Evolutionary Phases

To initialize each simulation, $N = 1024$ agents are embedded in a social network. This is a large enough population to sustain the diversity of strategies even if the proportion of a certain strategy runs low, but a small enough population for computational convenience. In the default model, a grid network was used, but different structures and mobility levels of the network will be explored in robustness tests. As I will describe later in Supplementary Sections ***2.7 Network Structure*** and ***2.8 Network Mobility***, the norm spillover effect remains robust across a broad range of mobility levels and network structures.

The simulation repeats iterations consisting of the following two steps: 1) interaction phase and 2) evolutionary dynamic.

In the interaction phase, in each iteration, every pair of agents connected by the network is assigned to interact. The interaction phase consists of two subphases in sequence: 1) ostracism and 2) cooperation.

In the ostracism subphase, each agent in the pair chooses whether to ostracize its partner based on its own strategy and its impression of its interaction partner. If either of the pair chooses to ostracize the other, the interaction phase of this pair is ceased for this iteration. Otherwise, they will continue to the cooperation subphase.

In the cooperation subphase, each agent in the pair chooses whether to cooperate with its partner based on its cooperation strategy. Then both agents in the pair gain payoffs as specified in **Table 1**. Agents know each other's cooperation and symbolic strategies after one single interaction. This is a simplified assumption that people learn about each other's strategy through direct interactions.

At the end of each iteration, agents update their strategies based on the payoffs that they gained from the current iteration. According to previous literature, there are three commonly used evolutionary dynamic rules: 1) pairwise comparison (Traulsen et al., 2007), 2) death-birth (Allen et al., 2020), and 3) birth-death (Tkadlec et al., 2020). The different evolutionary dynamic rules correspond to different scenarios of behavioral evolution in real life and may sometimes lead to qualitatively different evolutionary trajectories (Hindersin & Traulsen, 2015; Liu et al., 2017; Ohtsuki et al., 2006). In the default model, the death-birth rule will be used. In robustness tests, I also use pairwise comparison and test whether different evolutionary dynamic rules lead to different results. As I will describe in Supplementary Section ***2.6 Pairwise Comparison as Evolutionary Dynamic***, the norm spillover effect remains robust when pairwise comparison is used as the evolutionary dynamic. Although the birth-death rule has been widely applied to mathematical models (Hindersin & Traulsen, 2015), it is rarely used in agent-based models. Thus, it is not used in this dissertation.

A death-birth evolutionary dynamic mimics biological evolution in natural selection and is widely used to study the evolution of social behaviors (De et al., 2015; Roos et al., 2015). A death-birth evolutionary dynamic contains two subphases in sequence: 1) a death phase and 2) a birth phase.

In the death phase, a proportion of δ existing agents are randomly selected without replacement to "die." The parameter δ represents the death rate. By default, δ is set at 0.1, but in robustness tests, other values will also be explored. As I will describe in Supplementary Section ***2.4 Death Rate***, the norm spillover effect holds robust as long as the death rate is not too high. If an agent "dies," it will be removed from the population, leaving behind an empty spot in the network. All the other agents will also remove the memories of that dead agent.

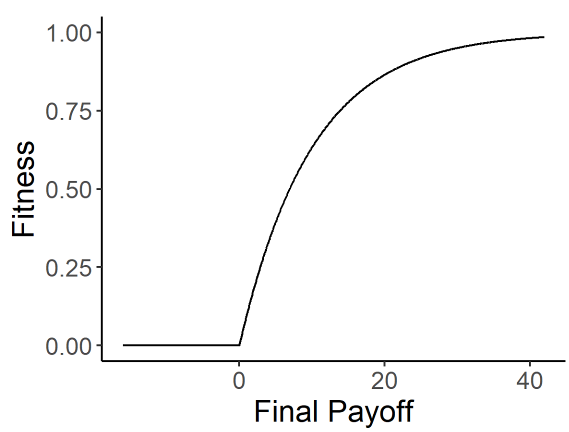
In the birth phase, each existing agent's fitness will be calculated. An agent's fitness represents their evolutionary adaptiveness in the current environment (Sigmund & Nowak, 1999). An agent's fitness is calculated as follows: After getting their *interaction payoff* in the interaction phase, each agent gets a *base payoff* = 30 from the environment.

In addition, each agent is also subject to a deduction of τ from the payoff (Roos et al., 2015). Such a parameter controls the base payoff of the agents. By default, $\tau = 15$. Thus, an agent's *final payoff*, π, in the current iteration can be represented in Equation (1). This *final payoff* is transformed into an agent's *fitness*, $f(\pi)$, based on the well-established principle of diminishing marginal utility (Diener & Biswas-Diener, 2002; Foster, 2004), as shown in Equation (2) and **Figure *2***. The parameter τ can also be interpreted as a manipulation of ecological threat in the environment. As I describe in Supplementary Section ***2.3 Ecological Threat***, the norm spillover effect remains robust across a broad range of τ.

$$\pi = \textit{interaction payoff} + \textit{base payoff} - \tau \tag{1}$$

$$f(\pi) = \begin{cases} 1 - e^{-0.1 \cdot \pi}, \text{if } \pi \geq 0; \\ 0 \qquad\qquad\;\, , \text{if } \pi < 0 \end{cases} \tag{2}$$

Figure 2 Relationship Between Final Payoff and Fitness



After calculating the fitness, each agent is chosen in a random order and given a chance to reproduce with a probability equal to its fitness. Reproduction means creating an offspring in a randomly selected adjacent empty site, if there is any. The offspring is a new agent that usually has the same cooperation, symbolic, and norm enforcement strategies as its parent. However, for each strategy, there is a small probability μ (i.e., the *exploration rate*) that this strategy will be randomly selected from all the possible options. By default, $\mu = 0.005$. As I will describe in Supplementary Section ***2.5 Mutation***, the norm spillover effect remains robust as long as μ is not too large.

Each simulation run repeats the interaction and evolutionary dynamic phases alternatively for 30,000 iterations. For each condition, 50 simulation runs are run and averaged across.

Manipulation of Initial Correlation Between Domain I and Domain II

To test whether the initial correlation between the behaviors in *Domain I* and *Domain II* at the population level leads to the evolution of the symbolic norm, a key independent variable is the *initial correlation* between behaviors in *Domain I* and *Domain II* on the population level. The magnitude of initial correlation is controlled by parameter $0 \leq \sigma \leq 0.25$, as in **Table *3***. When $\sigma = 0$, the behaviors in *Domains I* and *II* are initially independent at the population level. As σ gets larger, cooperative behavior co-occurs more often with the symbolic strategy *X*. When $\sigma = 0.25$, this means at the beginning of the simulation, all the cooperators are X-performers and all the defectors are Y-performers. Note that because $\sigma \geq 0$, cooperative behavior always co-occurs more often with *X* than *Y*. Since *X* and *Y* are equivalent to each other, results should be exactly symmetric if cooperative behavior co-occurs more often with *Y*.

To examine the random effect of σ, I randomly select ten levels of σ using R (RStudio Team, 2020). The ten levels are selected to be $\sigma \in [0.032, 0.240, 0.023, 0.079, 0.136, 0.195, 0.097, 0.240, 0.187, 0.085]$. Then I run 50 simulations under each level of σ. Since $\sigma = 0.240$ was selected twice, there ended up being 100 simulations under $\sigma = 0.240$ and 50 simulations under each of the other eight σ levels. In addition, since the nine randomly selected σ levels do not cover the full range of σ, I run 50 additional simulations under each level of $\sigma \in [0, 0.05, 0.1, 0.15, 0.2, 0.25]$.

Table 3 Initial Proportions of Strategy Combinations

	Cooperator (C)	Defector (D)
X-performer (X)	0.25 + σ	0.25 - σ
Y-performer (Y)	0.25 - σ	0.25 + σ

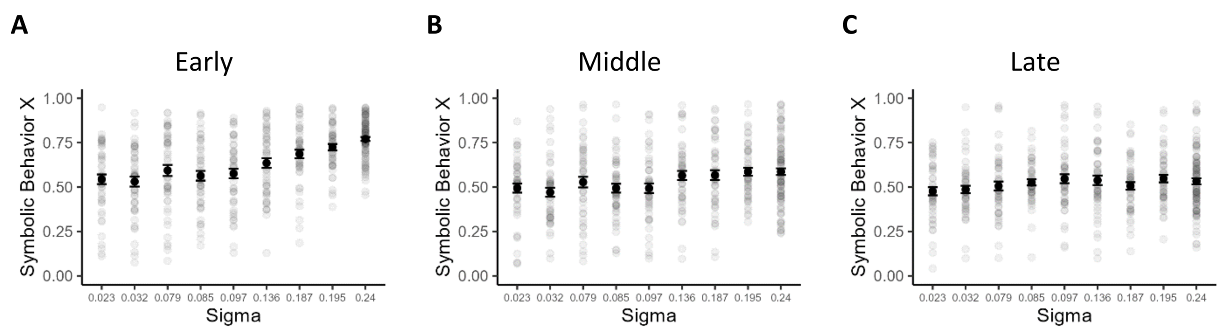
Results

The Effect of σ on Norm Spillover

As predicted, as the symbolic strategy *X* co-occurs more often with the cooperative behavior, more agents adopt the symbolic strategy *X* and enforce the symbolic norm *X* by ostracizing its violators (i.e., ostracizing Y-performers).

Figure *3* compares the strength of symbolic norm across different levels of *σ*. As *σ* gets stronger, the strength of symbolic norm increases, especially in the early phase of the simulation. However, the symbolic norm fades away in the middle and late phases of the simulation.

Figure 3 Strength of Symbolic Norm in Early, Middle, and Late Phases of the Simulation

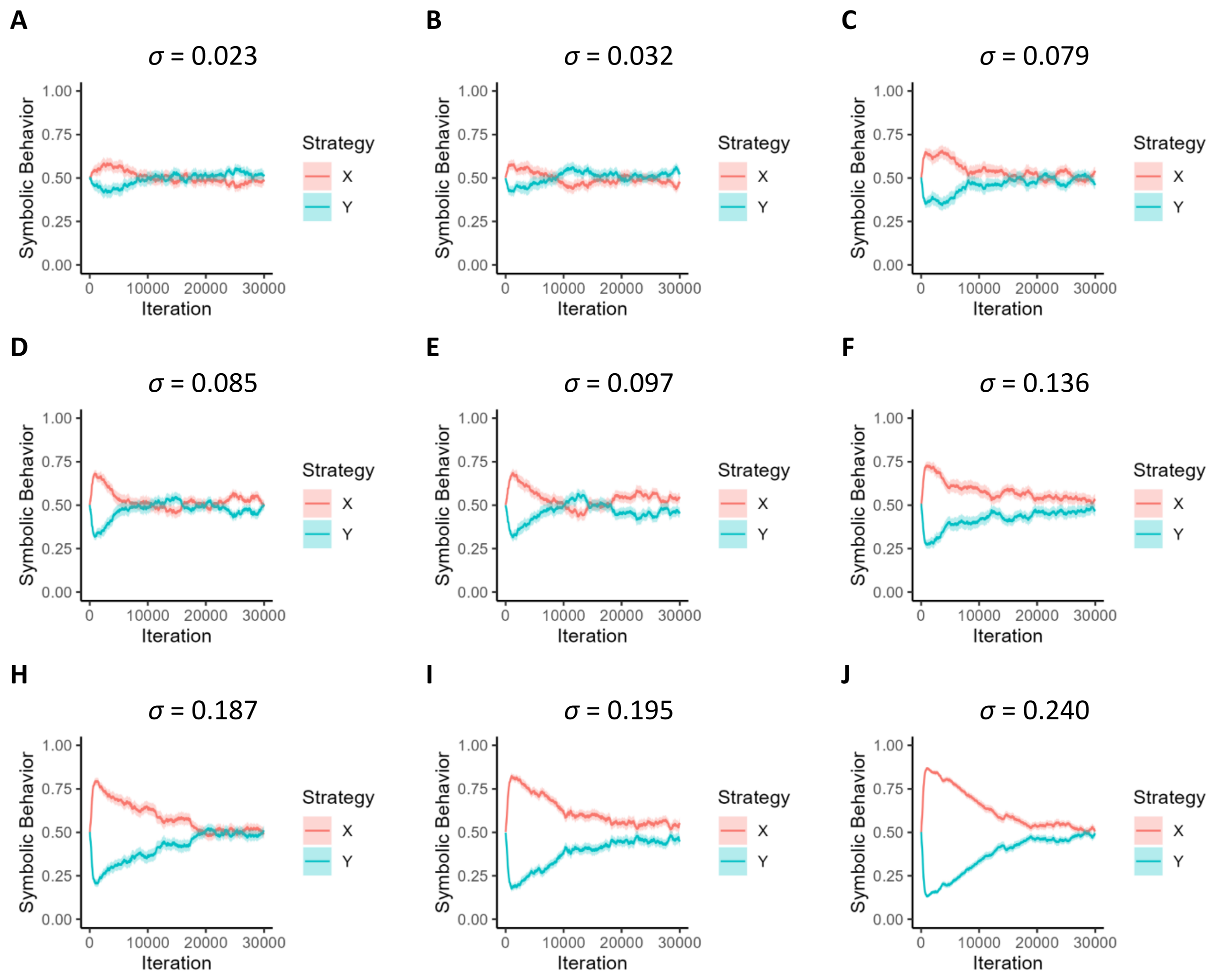


Note. The strength of symbolic norm in the early, middle, and late phases of the simulation. In the early phase of the simulation, the symbolic norm *X* evolves. As *σ* gets stronger, the strength of symbolic norm increases. However, the symbolic norm fades away in the middle and late phases of the simulation. Each data point in Plot A is an average of the 1-10000 iterations of that simulation run. Each data point in Plot B is an average of the 10001-20000 iterations of that simulation run. Each data point in Plot C is an average of the 20001-30000 iterations of that simulation run. For this figure and all the following figures, the error bars show the standard errors across the simulations under the same condition.

To zoom into the evolutionary trajectories across the time course of the simulations, **Figure *4*** shows the change of proportions of symbolic strategies across time under different levels of *σ*. When the symbolic strategy *X* co-occurs more often with the cooperative behavior, the symbolic norm *X* evolves.

Comparing across Plots A to J, results show that as the correlation (i.e., σ value) gets stronger, the peak level of symbolic norm *X* is higher. However, as time goes on, the symbolic norm gradually decreases. This is not surprising because as both practical norm followers (i.e., cooperators) and violators (i.e., defectors) start to adopt the same symbolic choice, the symbolic strategy no longer correlates with any practical behavior. Because of random mutation, the symbolic strategy will eventually become random with half X-performers and half Y-performers. In real life, social norms change. For instance, once being a social taboo, tattoos have become more socially acceptable and even fashionable over time (*Tattoos*, 2022). The simulation predicts that, although it takes time, a symbolic norm with no direct benefit may eventually fade away.

Figure 4 The Evolution of Symbolic Strategy



Note. The evolutionary trajectories of symbolic strategies *X* and *Y* overtime. In each plot, the symbolic norm *X* evolves in the early phase of the simulation and gradually decreases afterwards. As the initial correlation between *X* and cooperative behavior (i.e., σ value) gets stronger, the peak proportion of *X* is higher. For this figure and all the figures below, unless specified otherwise, each trajectory is the average across 50 simulation runs under the same condition. The shadows show standard errors. Plot J is averaged across 100 simulation runs under $\sigma = 0.240$, because this σ value was randomly chosen twice by an R program.

Figure *5* shows the proportions of different norm enforcement strategies in the early, middle, and late phases of the simulation. Plot A shows that a substantial proportion of agents evolve to enforce symbolic norm *X* by ostracizing Y-performers in the early phase of the simulation (i.e., YQ). As σ gets stronger, the enforcement of the symbolic norm becomes stronger. However, Plots B and C show that the proportions of YQs decrease in the middle and late phases of the simulation. Compared with Plots A-B, Plots D-F show that there are less agents who enforce *Y* (i.e., XQ).

Interestingly, Plots G-O in **Figure *5*** show that there are few agents who ostracize defectors (i.e., DYQ, DXQ, DQ). This is counterintuitive because logically, the most effective strategy should be ostracizing practical norm violators (i.e., ostracizing defectors) directly, rather than ostracizing symbolic norm violators. The reason for the lack of DQs is that in the default model, agents' cooperation strategies are not transparent. Agents do not know each other's cooperation strategy until they actually interact at least once in a cooperation game. As a result, an agent cannot decide whether to ostracize another agent based on their cooperation strategy. Section **Transparency of Cooperation Strategy and Symbolic Strategy** will show that if agents' cooperation strategies are transparent, most agents will ostracize defectors directly. Nevertheless, no matter whether cooperation strategy is transparent, the symbolic norm still evolves as long as agents' symbolic strategies are easily observable.

Figure 5 Norm Enforcement Strategies in Early, Middle, and Late Phases of the Simulation

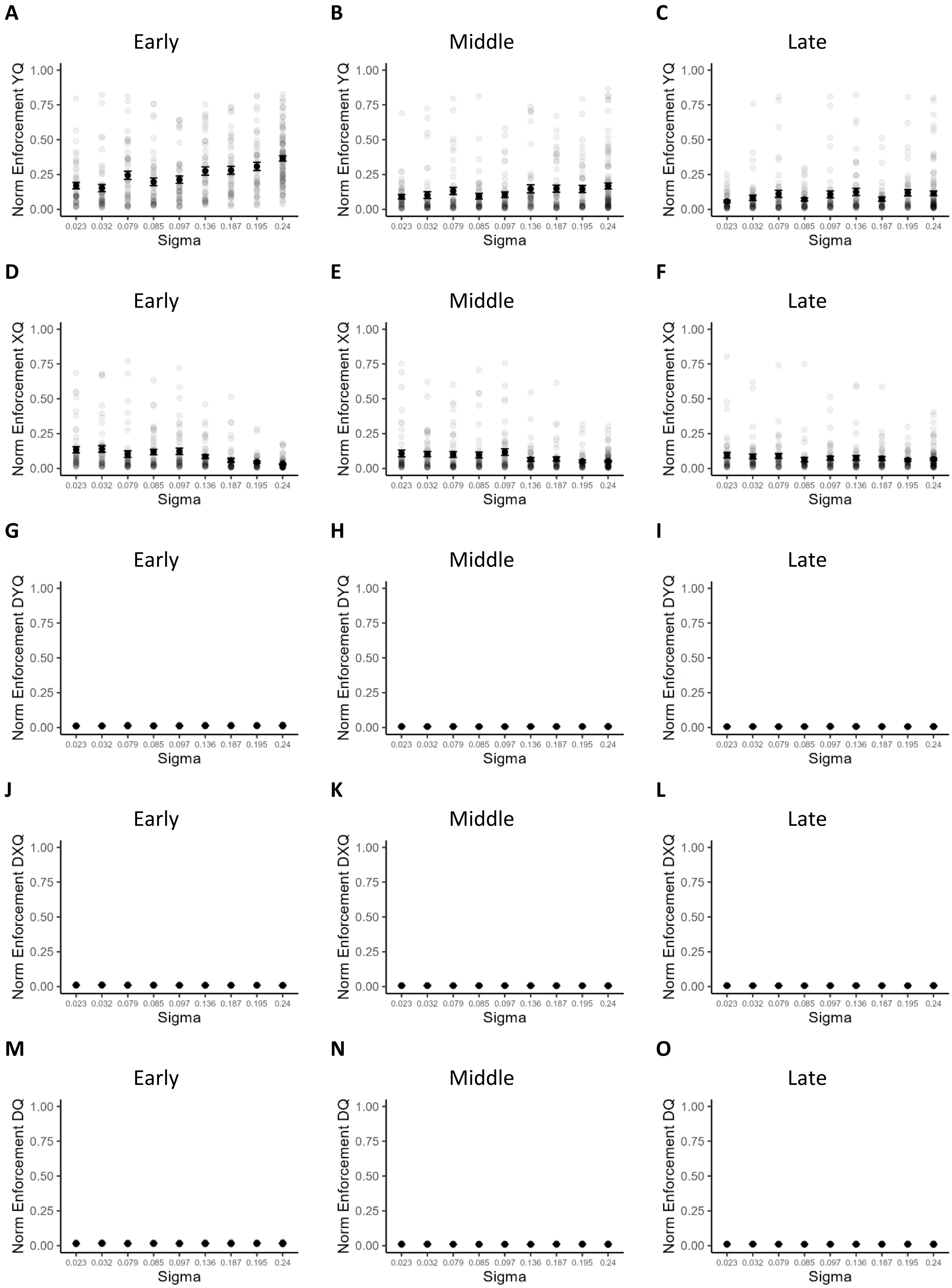
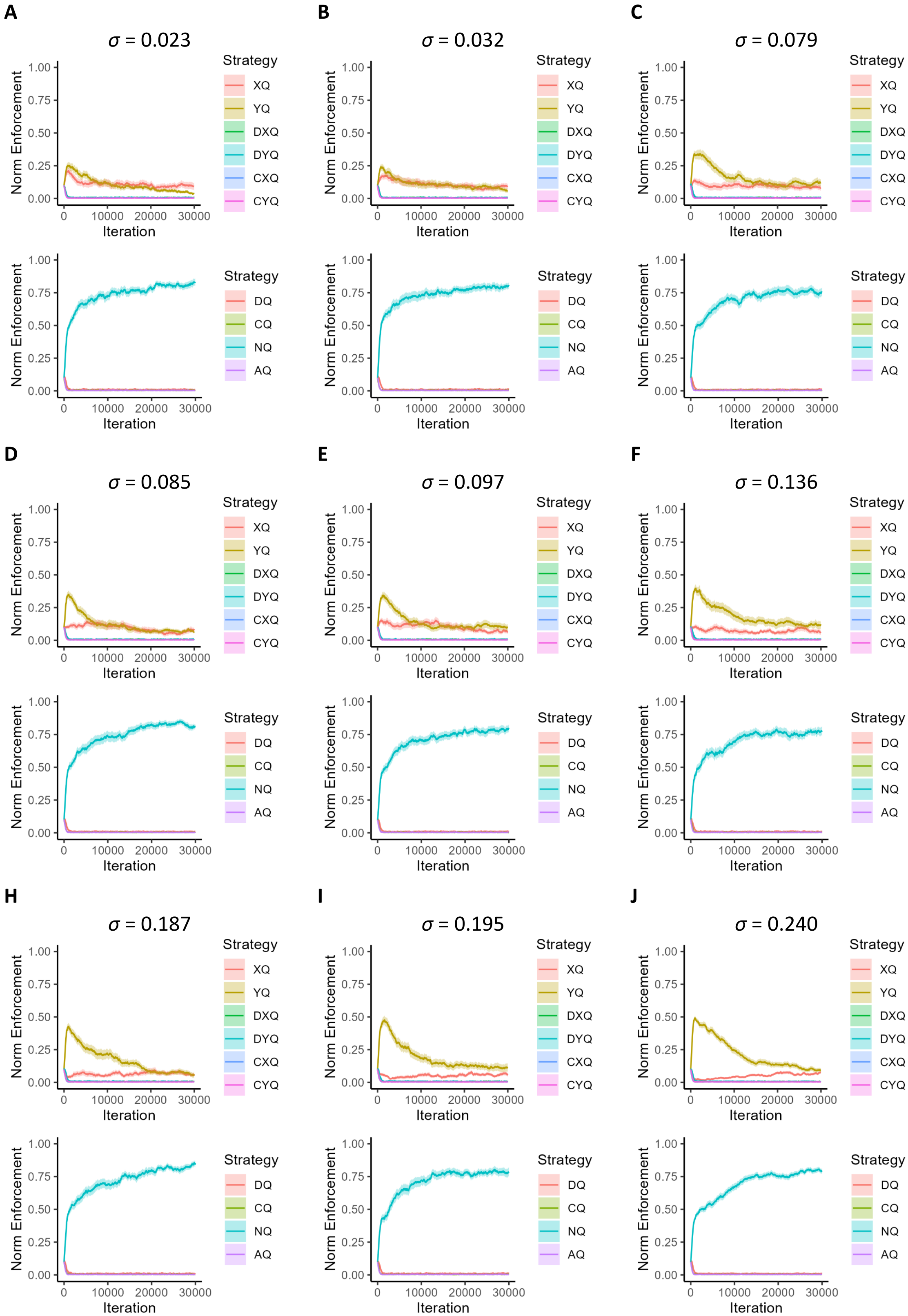


Figure *6* shows the evolutionary trajectories of different norm enforcement behaviors across time. Each plot shows the proportions of the ten ostracism strategies (see **Table 2**). The yellow line, YQ, shows the proportion of agents who enforce the symbolic norm *X* by ostracizing Y-performers. When the symbolic strategy *X* co-occurs more often with the cooperative behavior, a substantial proportion of agents evolve to enforce the symbolic strategy *X*. Moreover, as the initial correlation between *X* and cooperative behavior (i.e., σ value) becomes stronger, the peak level of the enforcement of *X* is higher.

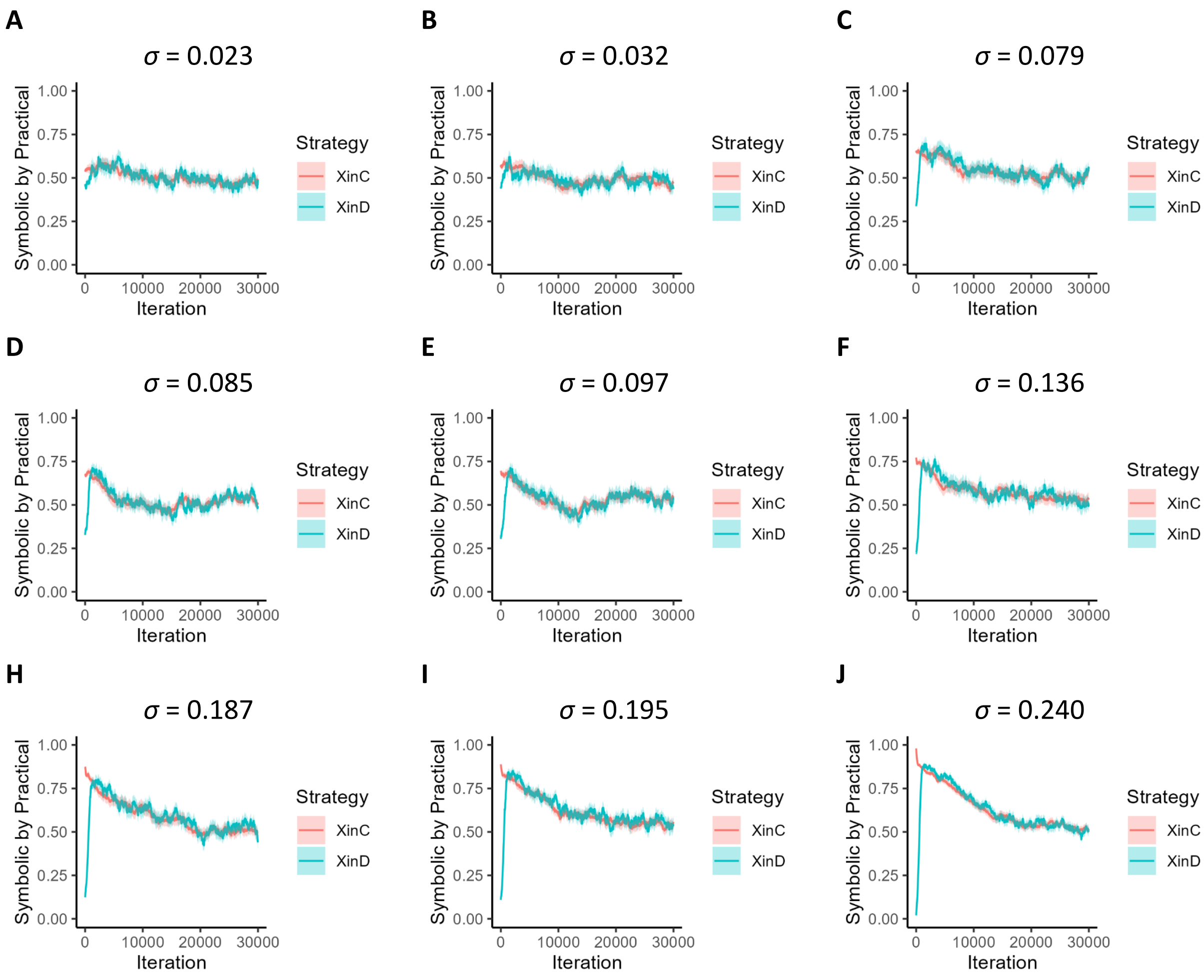
Figure 6 The Evolution of Norm Enforcement Behavior



Note. The evolutionary trajectories of different norm enforcement behaviors overtime. The yellow lines, YQ, show the proportions of agents who enforce the symbolic norm *X* by ostracizing Y-performers. In each plot, a substantial proportion of YQ evolves in the early phase of the simulation and gradually decreases afterwards. As the initial correlation between *X* and cooperative behavior (i.e., σ value) gets stronger, the peak proportion of YQ is higher. Plot J is averaged across 100 simulation runs under σ = 0.240, because this σ value was randomly chosen twice by an R program.

Figure *7* shows that because of the enforcement of the symbolic norm, defectors quickly adopt the symbolic norm. Once defectors have adopted the symbolic norm to the same extent as cooperators do, there is no longer a correlation between the symbolic and practical behaviors. Since then, the symbolic norm starts to decrease. This result also shows that the evolution of symbolic norm is *not* due to that cooperators outperform defectors and cooperators also happen to be X-performers (see **Figure 49** and **Figure 50** in Supplementary Section ***1.1 The Evolution of Cooperation*** in **Supplementary Materials** for more results on cooperation rates).

Figure 7 The Evolution of Symbolic Strategy Among Cooperators and Defectors



Note. The proportion of X-performers among cooperators and defectors, respectively. XinC represents the proportion of X-performers among cooperators. XinD represents the proportion of X-performers among defectors. The nine plots show that the symbolic strategy *X* evolves among both cooperators and defectors. Plot J is averaged across 100 simulation runs under $\sigma = 0.240$, because this σ value was randomly chosen twice by an R program.

Norm Spillover Effect Under Additional σ Levels

Since the nine randomly selected σ levels do not cover the full range of σ, six additional levels of σ are chosen to cover a broader range of σ. In this section, I run 50 simulation runs under each level of $\sigma \in$ [0, 0.05, 0.1, 0.15, 0.2, 0.25].

Results replicate the findings above. **Figure 8** shows that when $\sigma > 0$, the symbolic norm X first evolves and then disappears. Moreover, as the initial correlation (i.e., σ value) gets stronger, the peak level of the symbolic norm X is higher.

Figure 8 The Evolution of Symbolic Strategy Under Additional σ Values

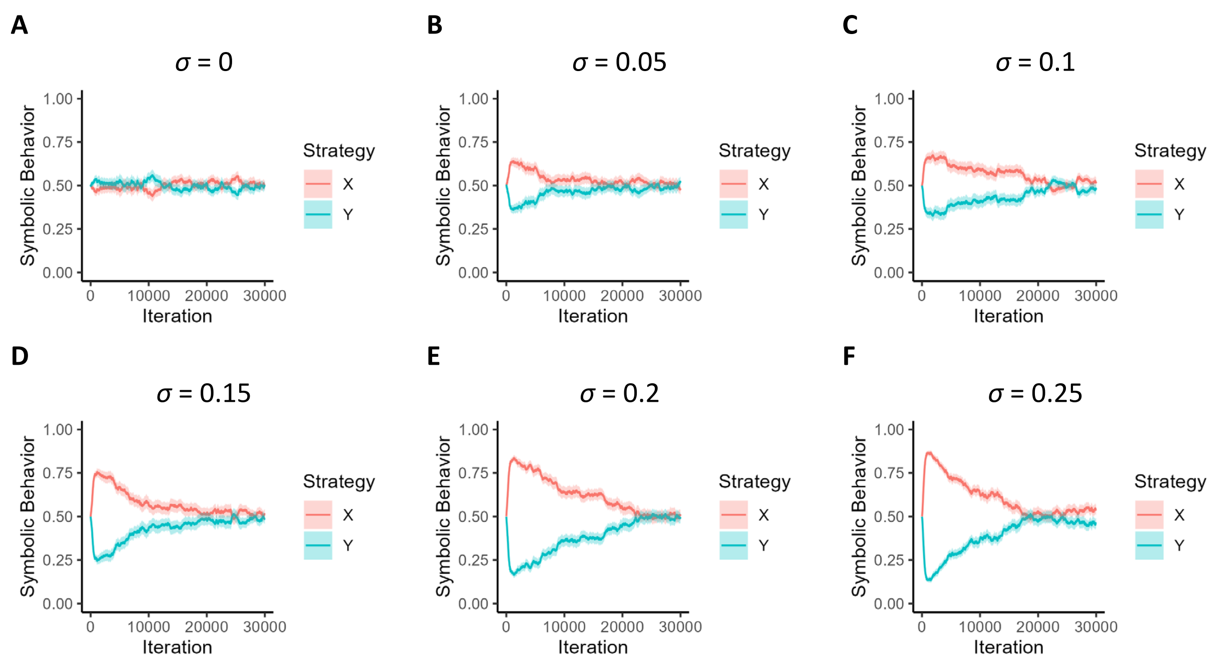


Figure 9 shows the trajectories of different norm enforcement behaviors under different levels of σ. The yellow line, YQ, shows the proportion of agents who enforce the symbolic norm X by ostracizing Y-performers. When the symbolic strategy X is correlated with the practical behavior C, a substantial proportion of agents evolve to enforce the symbolic norm X. Moreover, as the initial correlation (i.e., σ value) gets stronger, the peak level of the enforcement of X is higher.

Figure 9 The Evolution of Norm Enforcement Under Additional σ Levels

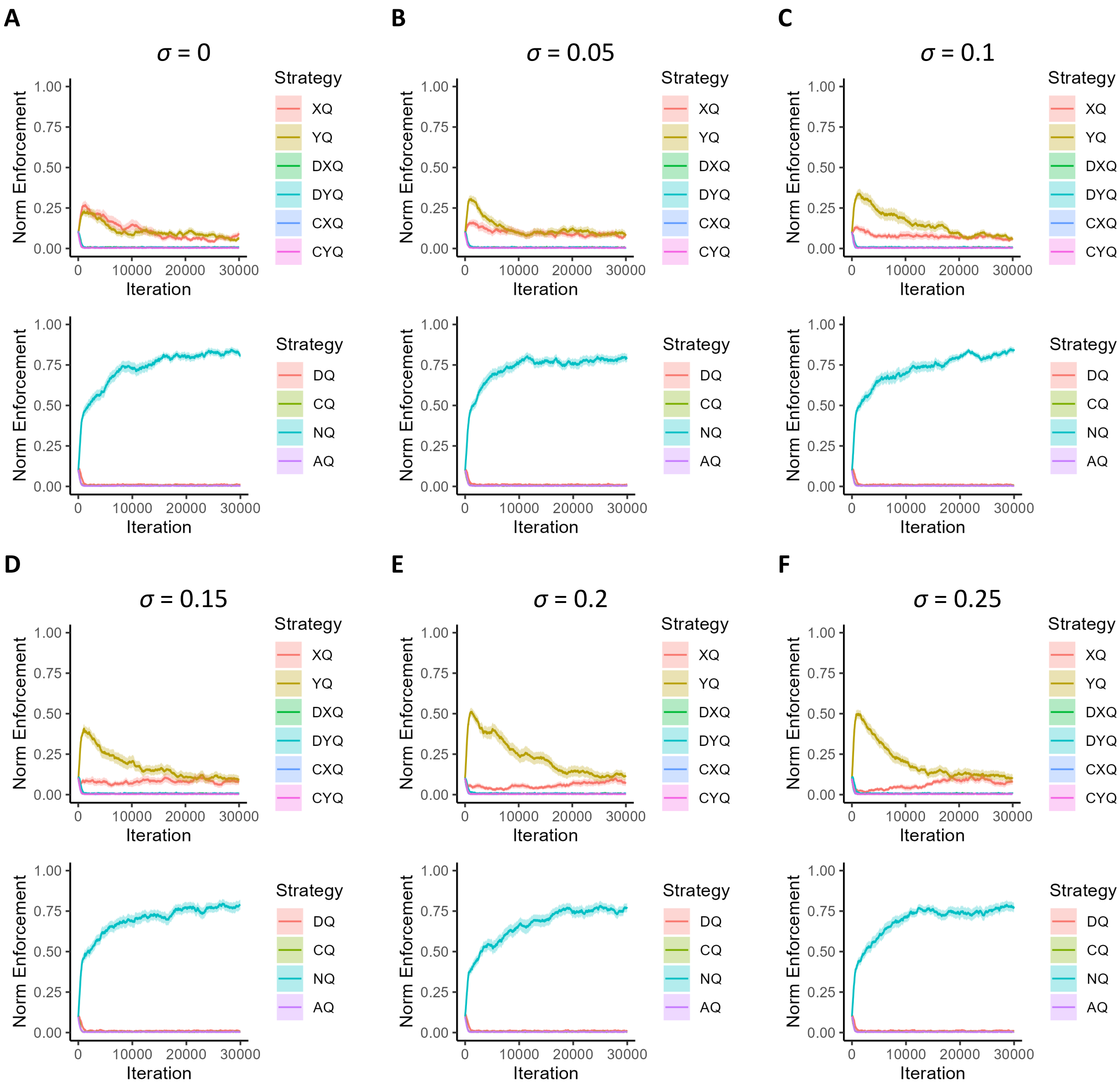
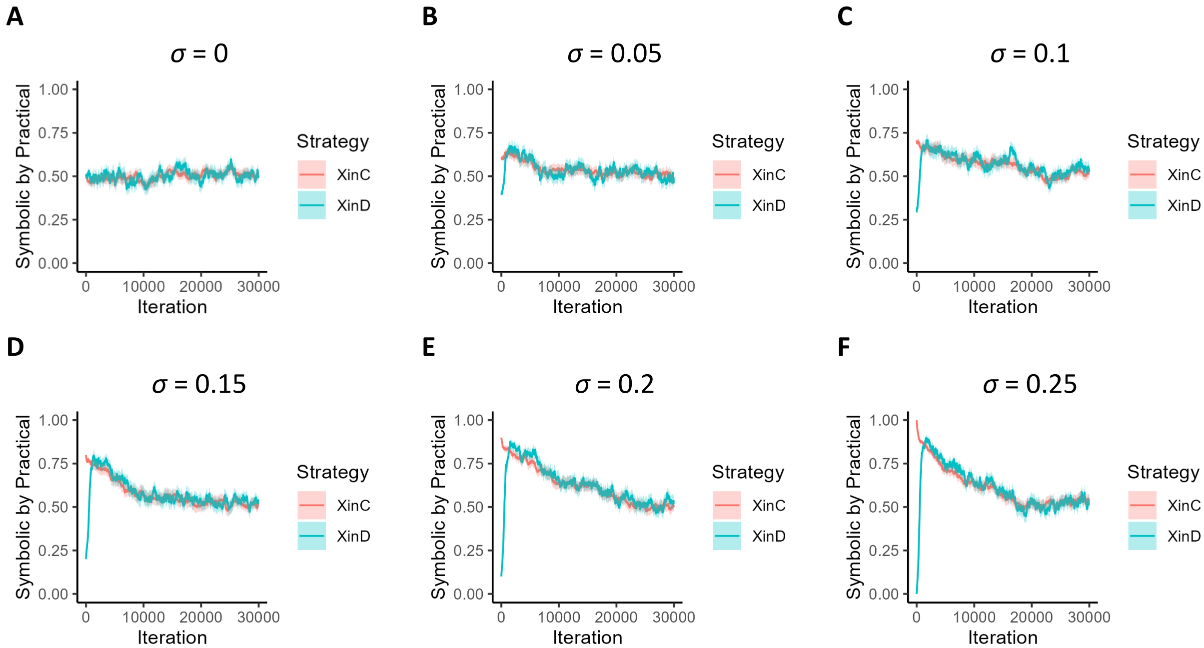


Figure 10 shows that because of the enforcement of the symbolic norm, defectors quickly adopt the symbolic norm. Once defectors have adopted the symbolic norm to the same extent as cooperators do, there is no longer a correlation between the symbolic and practical behaviors and the symbolic norm starts to decrease.

Figure 10 Symbolic Strategy Among Cooperators and Defectors Under Additional σ Values



Mechanisms Behind the Norm Spillover Effect

Transparency of Cooperation Strategy and Symbolic Strategy

Why do agents ostracize symbolic norm violators but do not ostracize practical norm violators directly? In this section, I show that the transparency of agents' cooperation and symbolic strategies play an important role.

To test the effects of the transparency of strategies, two variables are manipulated: 1) the transparency of cooperation strategy, and 2) the transparency of symbolic strategy. If a strategy is transparent, agents know each other's strategy in that domain even before their first interaction. If a strategy is non-transparent, agents randomly guess each other's strategy in that domain and use that guess as their initial impression. Nevertheless, after an agent has interacted with another in a cooperation game once, they know each other's both cooperation and symbolic strategies. Thus, the transparency of a strategy only influences agents' behaviors if they never interacted before.

This setup resembles how people learn each other's behavioral pattern through direct interactions. In this model, because only cooperators cooperate and only defectors defect, it makes sense that an agent can learn their partner's strategy after only a single interaction. In real life, this learning process may take longer. Thus, this model is a simplified assumption.

In the default model, the cooperation strategy is non-transparent and the symbolic strategy is transparent. In this section, three additional conditions are tested to examine the effects of strategy transparency: 1) when cooperation strategy and symbolic strategy are both transparent, 2) when both strategies are non-transparent, and 3) when cooperation strategy is transparent but symbolic strategy is non-transparent. In each condition, I run 50 simulation runs under $\sigma \in [0, 0.05, 0.1, 0.15, 0.2, 0.25]$. Six fixed levels of σ are used to make it easier to compare across conditions.

Figure 11 shows that the symbolic norm X evolves even when both cooperation and symbolic strategies are transparent. The key difference between this condition and the default model is that, when the cooperation strategy is transparent, the majority of the population become DQs who ostracize practical norm violators (i.e., ostracize defectors). In addition, there are more DYQs who ostracize both practical norm violators and symbolic norm violators (see **Figure 12**). This difference is because when agents' cooperation strategies are as transparent as their symbolic strategies, agents can easily differentiate between cooperators and defectors. In this case, a more effective strategy is to ostracize defectors directly.

Nevertheless, even when agents' cooperation strategies are completely transparent, there are still a substantial proportion of agents who ostracize symbolic norm violators (i.e., DQs and DYQs). Together with **Figure 11**, this result shows that agents adopt and enforce the symbolic norm even when they can easily know each other's practical behavior.

Figure 11 Symbolic Strategy When Both Cooperation and Symbolic Strategies are Transparent

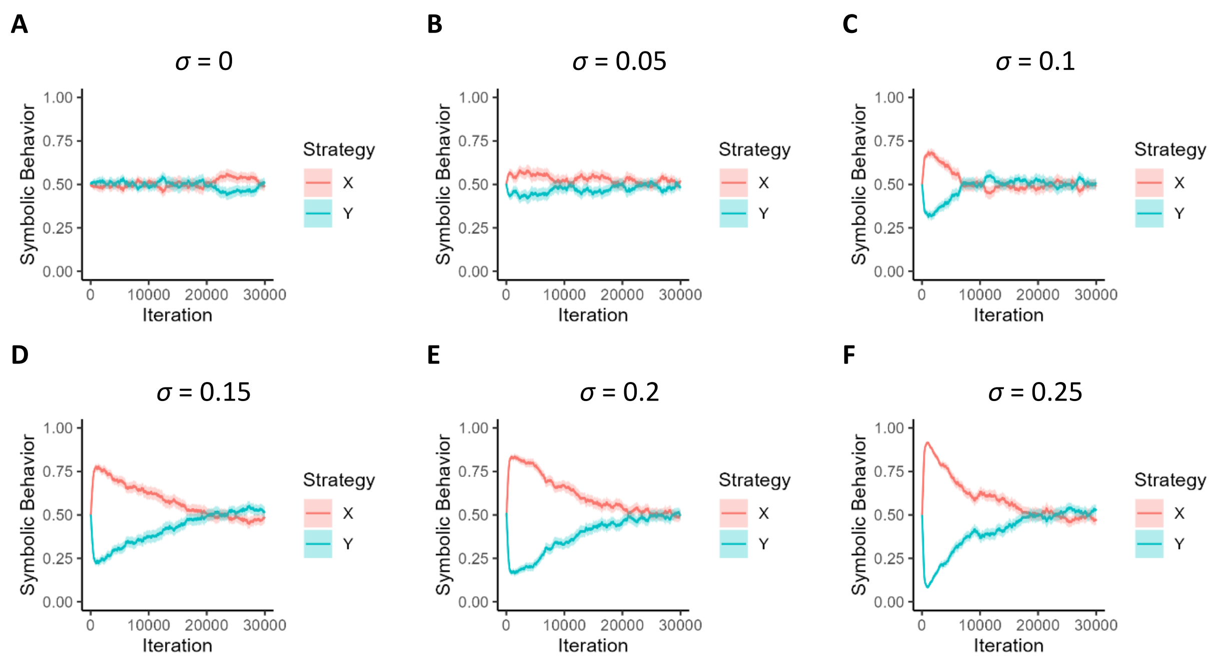
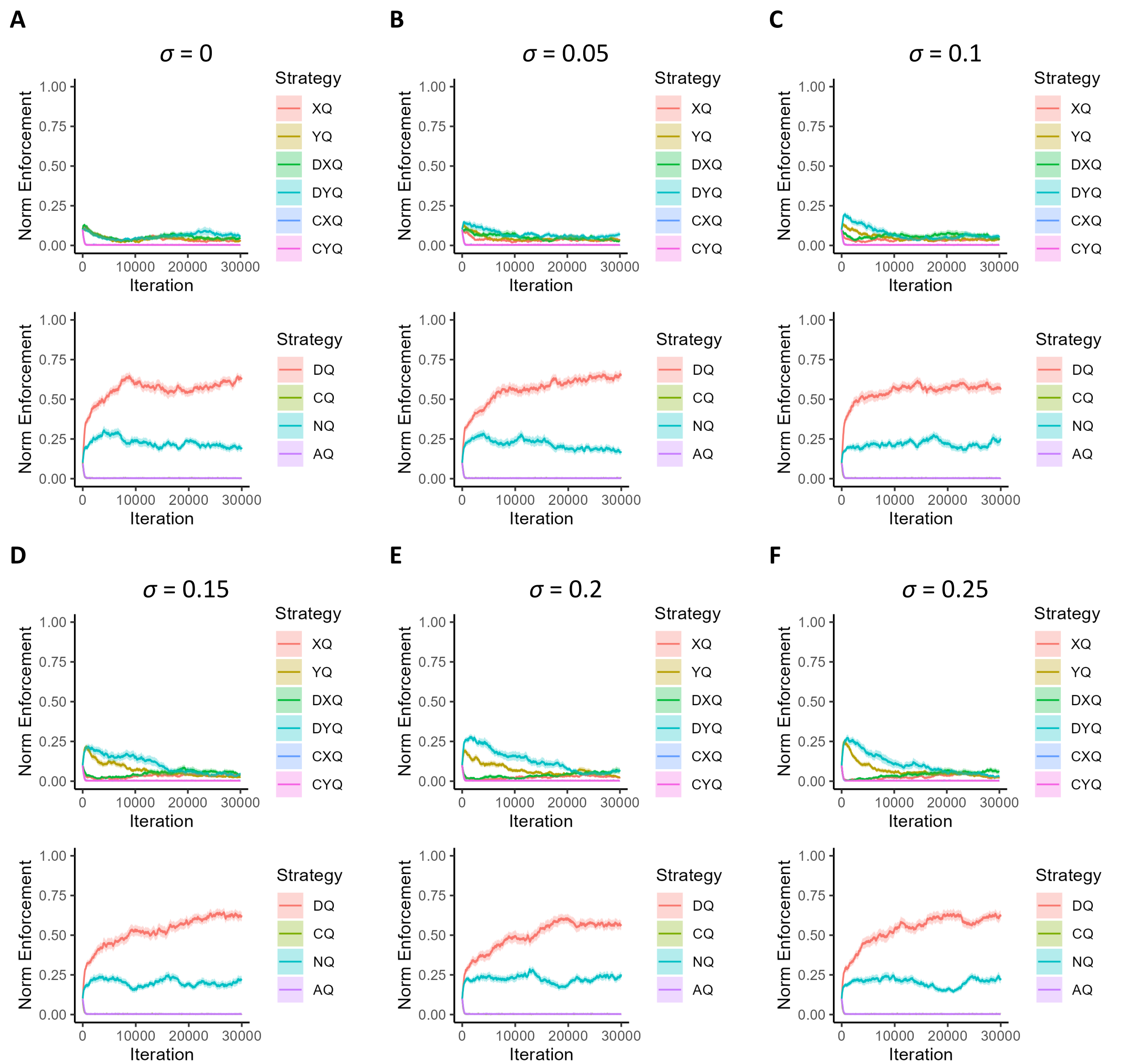


Figure 12 Norm Enforcement When Both Cooperation and Symbolic Strategies are Transparent



What if agents' symbolic strategies are as opaque as their cooperation strategies? This section examines the condition when agents' cooperation and symbolic strategies are both non-transparent. In this condition, the symbolic norm becomes much weaker (see **Figure 13**). Agents ostracize neither symbolic norm violators (i.e., few YQ or DYQ) nor practical norms violators (i.e., few DQ or DYQ). Instead, most agents become NQs who never enforce any norm (see **Figure 14**). This result indicates that for the norm to spillover to a symbolic domain, the behaviors in the symbolic domain have to be easily observable. In other words, a private behavior that cannot be easily observed is not susceptible to the norm spillover effect.

Figure 13 Symbolic Strategy When Both Cooperation and Symbolic Strategies are Non-transparent

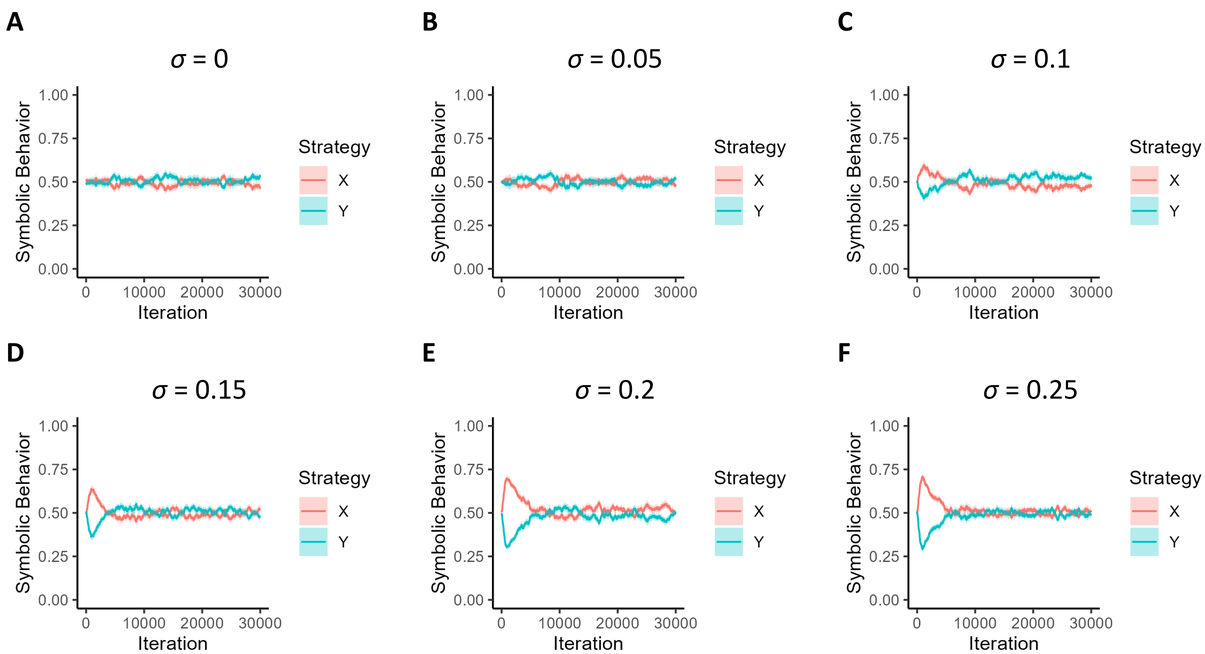
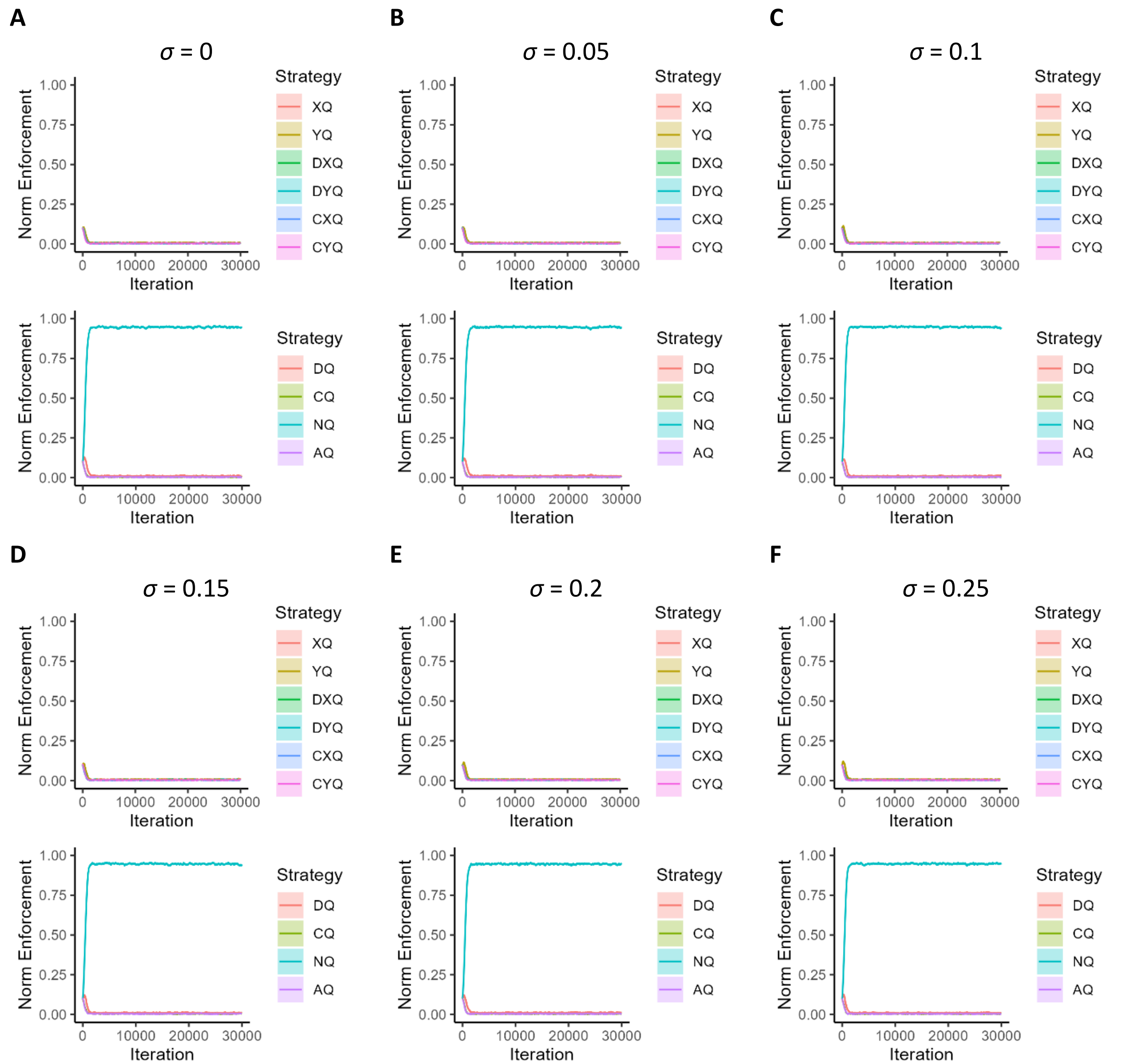


Figure 14 Norm Enforcement When Both Cooperation and Symbolic Strategies are Non-transparent



Next, I examine the condition when the cooperation strategies are transparent but the symbolic strategies are non-transparent. This is the condition in which agents should rely least on symbolic strategies. **Figure 15** supports this hypothesis. **Figure 15** shows that among all the strategies, the most prevalent is DQ—agents who only ostracize practical norm violators. On the other hand, few YQ or DYQ evolves. Interestingly, **Figure 16** shows that even when symbolic strategies are non-transparent, the proportion of X-performs still increases at the beginning the simulation. This is because many cooperators evolve in the early phase of the simulation and these cooperators are also more likely to be X-performers. **Figure 17** shows that X-performing cooperators (i.e., CX) take the major part of the population at the beginning of the population.

Figure 15 Norm Enforcement Under Transparent Cooperation and Non-transparent Symbolic Strategies

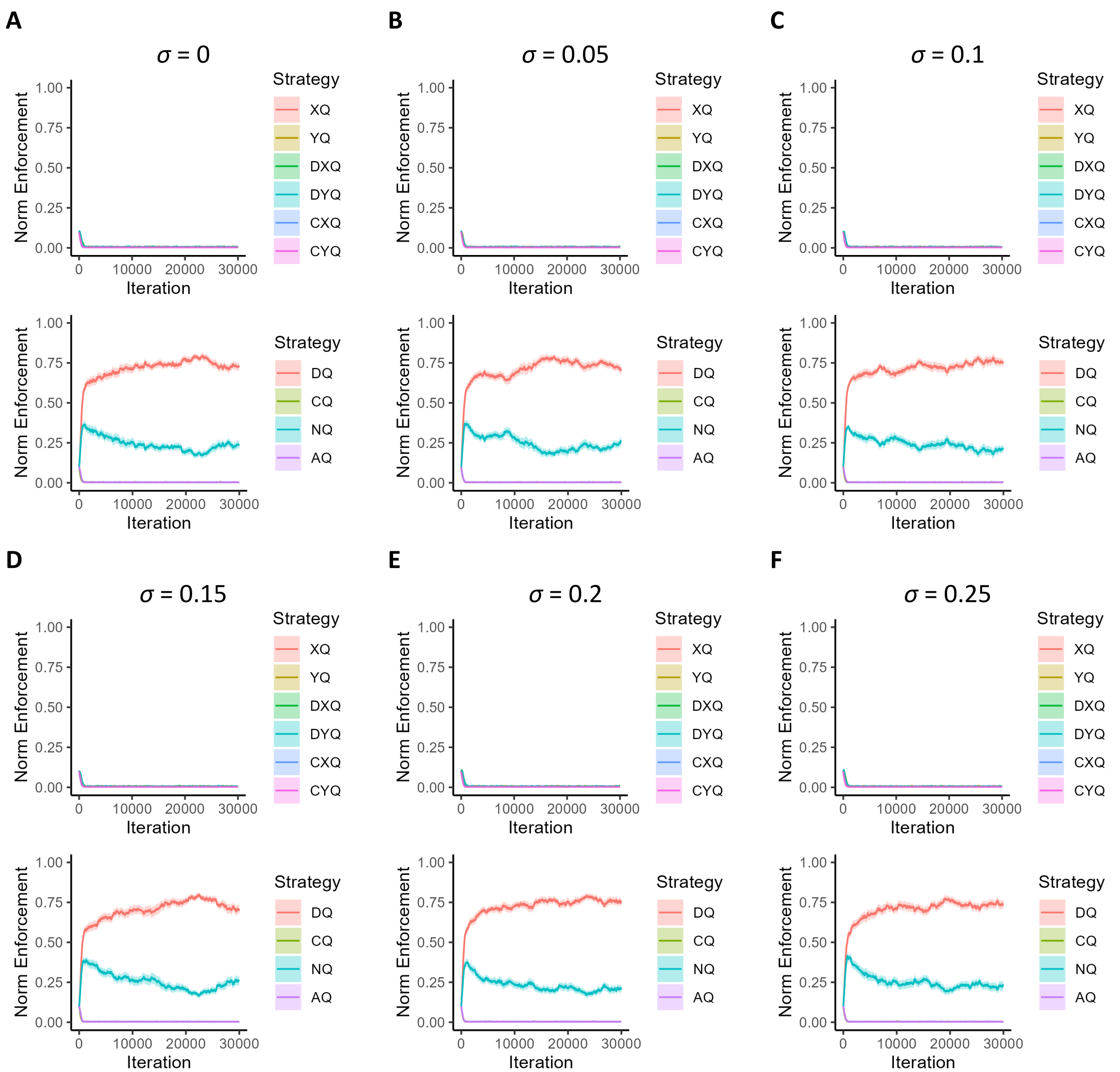


Figure 16 Symbolic Strategy Under Transparent Cooperation and Non-transparent Symbolic Strategies

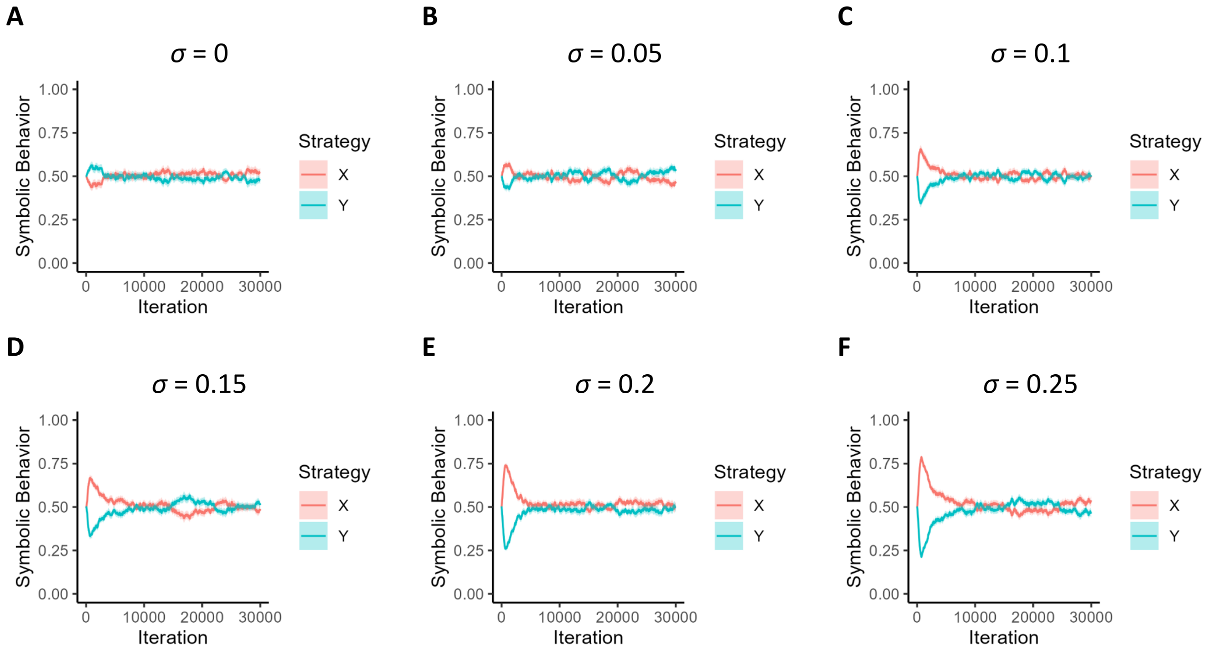
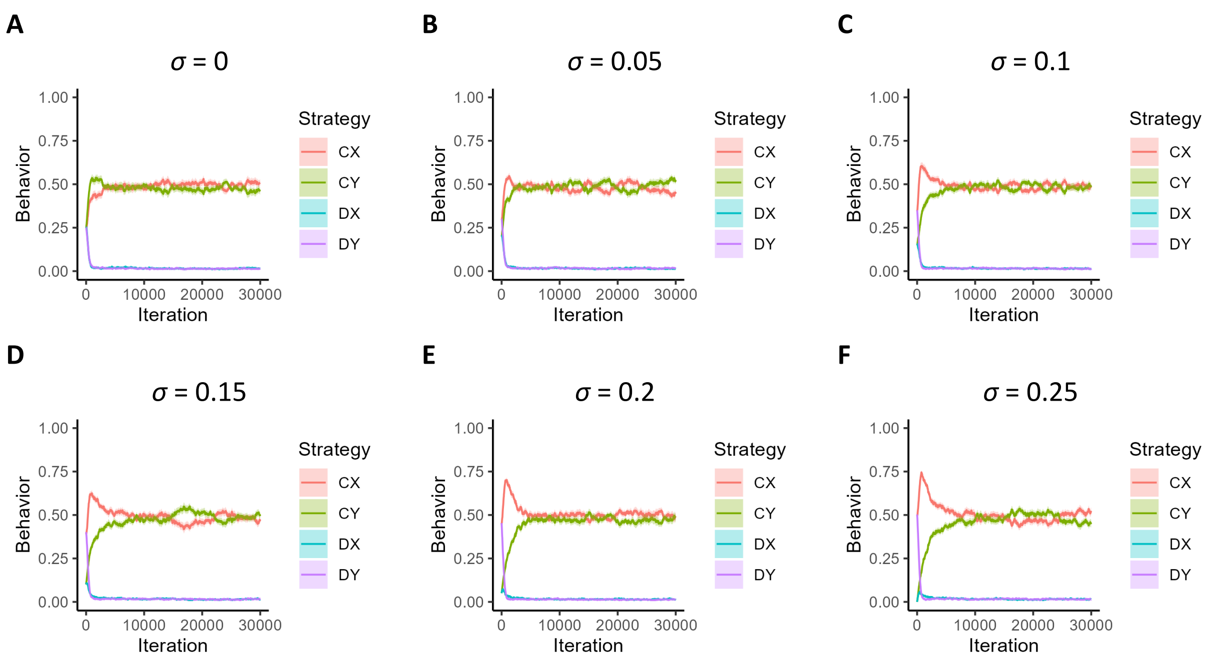


Figure 17 Behaviors Under Transparent Cooperation and Non-transparent Symbolic Strategies



To summarize, this section shows that the norm spillover effect largely relies on the transparency of the symbolic strategy. Only when the symbolic strategies are easily accessible, agents can decide whether to ostracize others based on their symbolic strategies.

In the current section, I have only examined the situations where a strategy is either completely transparent or completely non-transparent. In Supplementary Section ***2.1 Partial Strategy Transparency***, I also manipulate the transparency of a strategy as a continuous variable. The results show that the norm spillover effect remains robust when the symbolic and practical strategies are partially transparent, a point that will be elaborated later.

The Impact of Practical Behavior on the Norm Spillover Effect

The norm spillover effect can happen even without a prominent practical norm. In other words, the symbolic norm can evolve even when the correlated practical behavior is not the majority in the population. To examine the impact of the prevalence of practical behavior on the norm spillover effect, another set of simulations are run. Previous literature shows that the ratio of the benefit to the cost of cooperation is a key factor that decides cooperation rate (Nowak, 2006). Thus, the payoff matrix of the cooperation game (see **Table 1**) is manipulated this time, in order to achieve different levels of cooperation rates in the population. The higher the benefit of cooperation compared to the cost of cooperation, the higher cooperation rate a population will achieve.

In the default model, the cost of cooperation was set at $c = 1$ and the benefit of cooperation was set at $b = 5$. In this section, I first hold the cost of cooperation at $c = 1$ and used different levels of $b \in [2, 2.5, 3, 4, 6]$. Under these five levels of b, the ratio of b/c equals 2, 2.5, 3, 4, 6, respectively. Next, I hold the benefit of cooperation at $b = 5$ and use different levels of $c \in [2.5, 2, 1.67, 1.25, 0.83]$. Under these five levels of c, the ratio of b/c also equals 2, 2.5, 3, 4, 6, respectively. Under each of these ten conditions, I run 50 simulation runs under $\sigma = 0.15$ and compare the results with the original model where $c = 1$ and $b = 5$.

Figure 18 shows that as the benefit of cooperation gets higher, the cooperation rate increases, which is consistent with previous literature on cooperation (Nowak, 2006). Moreover, Plot A shows that when the ratio of *b*/*c* is 2, the major population become practical norm violators (i.e., defectors) because the benefit of cooperation is too small.

Nevertheless, Plot A in **Figure 19** shows that even when most of the population evolve to be defectors, the symbolic strategy that co-occurs more often with cooperation (i.e., X-performers) still evolves at the beginning of the simulation. Moreover, there are a substantial proportion of agents enforcing the symbolic norm *X* (see the evolution of YQs in Plot A in **Figure 20**). These results indicate that when a symbolic choice is correlated with a beneficial practical behavior, the population will adopt and enforce the symbolic norm even when they do not adopt the practically beneficial norm.

In addition, **Figure 19** shows that as the benefit of the practical norm increases, the symbolic norm becomes stronger. **Figure 20** shows that as the benefit of the practical norm increases, the enforcement of the symbolic norm also gets stronger.

Figure 18 Cooperation Rate Under Different Benefit of Cooperation

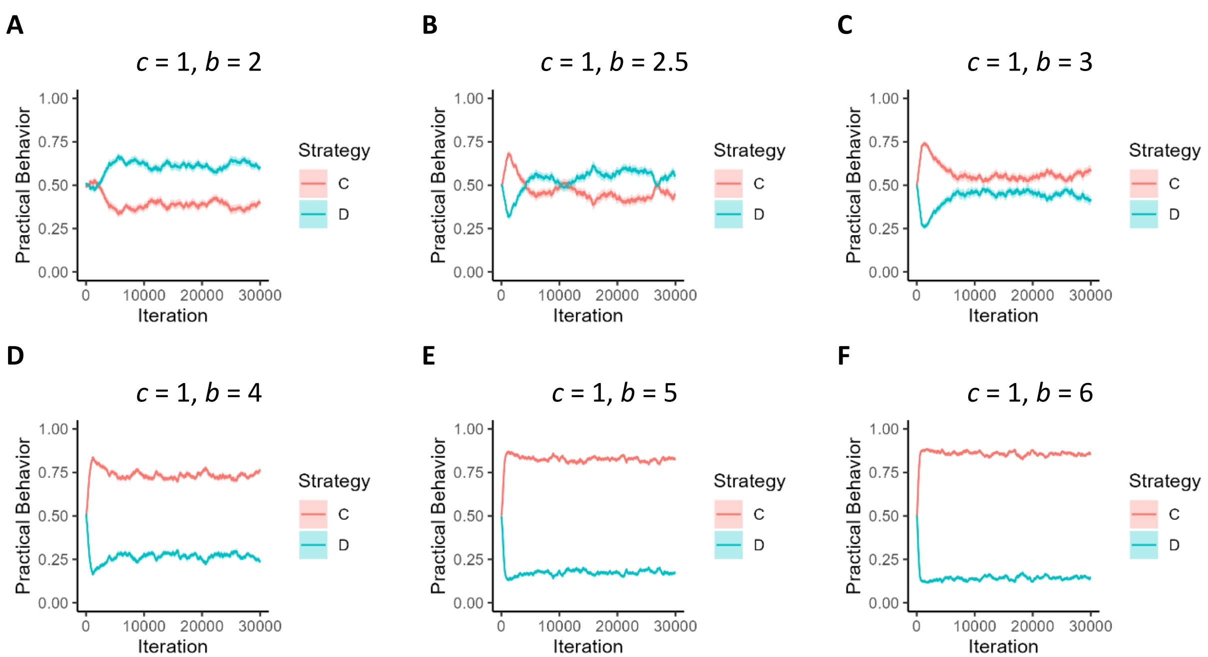


Figure 19 Symbolic Strategy Under Different Benefit of Cooperation

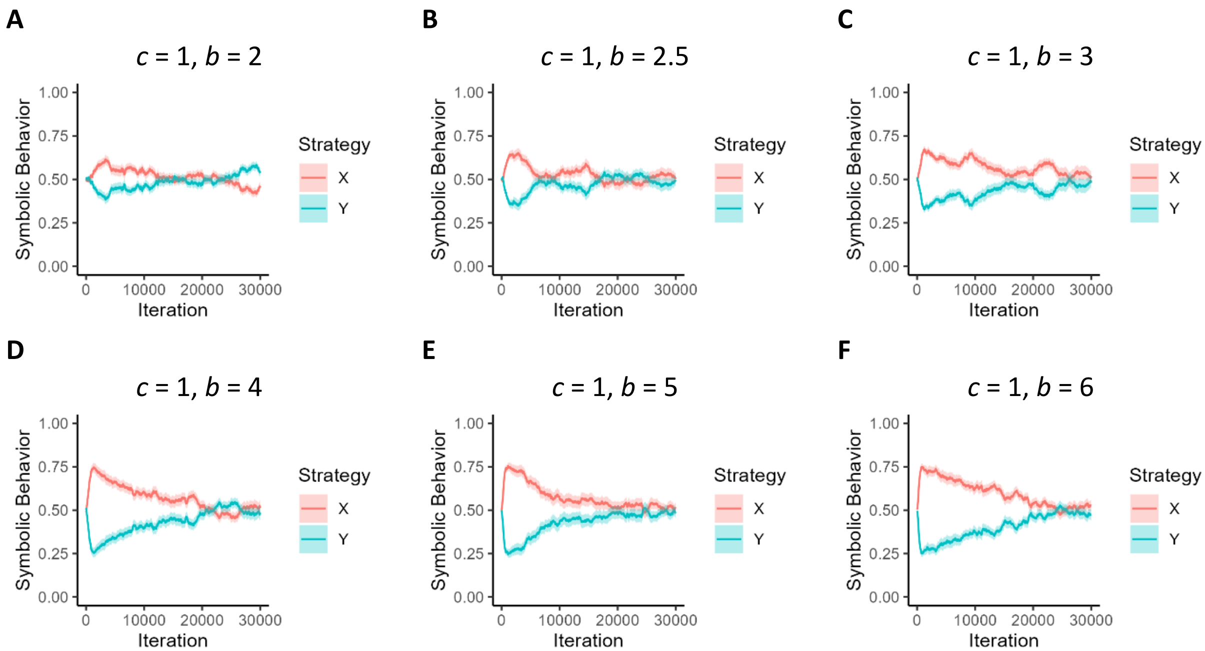


Figure 20 Norm Enforcement Under Different Benefit of Cooperation

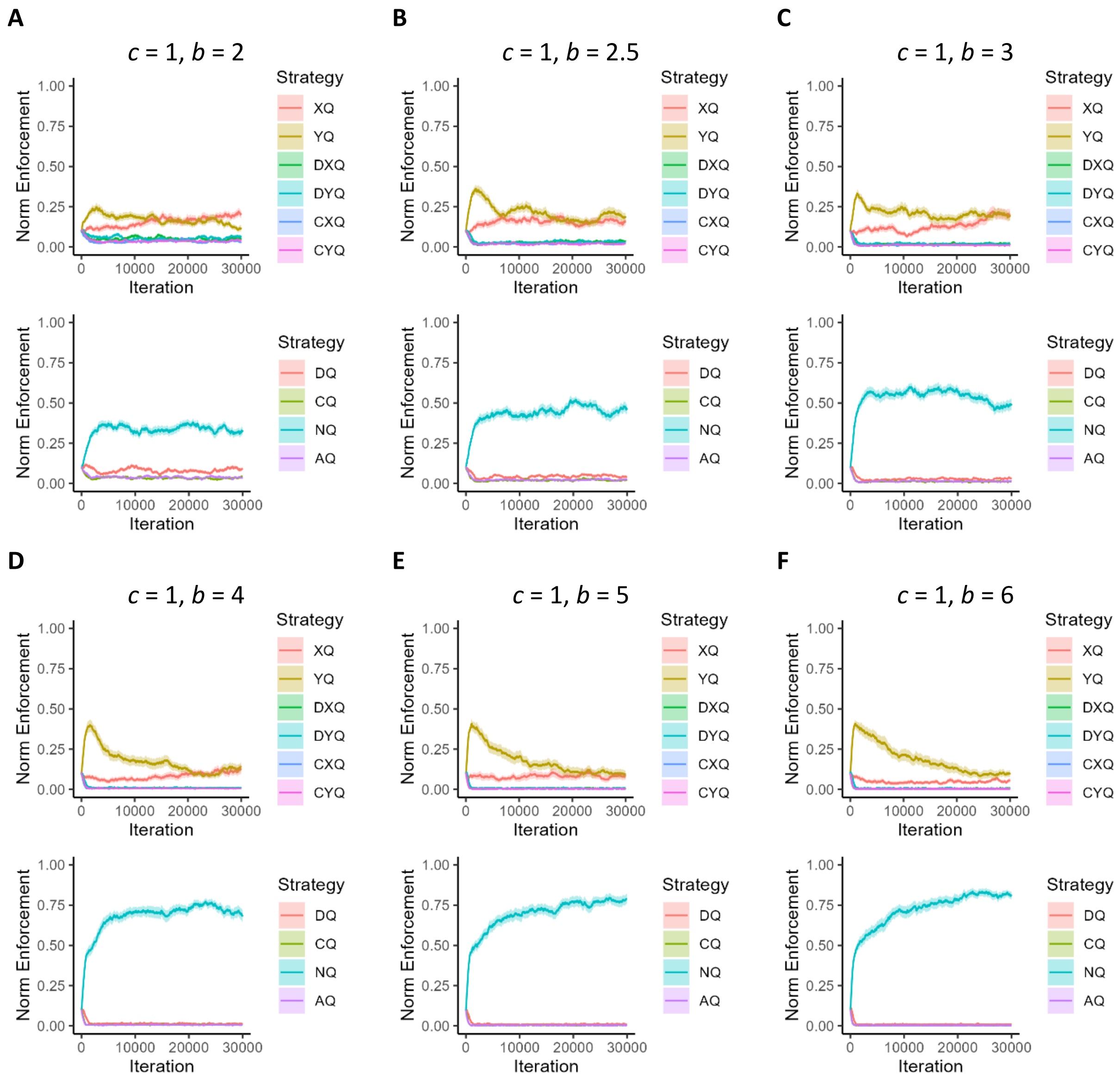


Figure 21 shows that as the cost of cooperation gets lower, cooperation rate increases, which is also consistent with previous literature (Nowak, 2006). Plots A and B in **Figure 21** show that when the ratio of *b*/*c* is lower or equal to 2.5, the major population become practical norm violators (i.e., defectors) because the cost of cooperation is too high.

Nevertheless, once again, Plots A and B in **Figure 22** show that even when the major population are defectors, the symbolic strategy that co-occurs more often with cooperation (i.e., X-performers) still evolves. Plots A and B in **Figure 23** show that there are a substantial proportion of agents enforcing the symbolic norm *X*. These results indicate that when a symbolic choice is correlated with a beneficial practical behavior, the population will adopt and enforce the symbolic norm even when they do not adopt the practically beneficial norm.

In addition, **Figure 22** shows that as the cost of cooperation decreases, the symbolic norm becomes stronger. **Figure 23** shows that as the cost of cooperation decreases, the enforcement of the symbolic norm also gets stronger.

To summarize, results in this section show that the norm spillover effect can happen as long as the symbolic choice correlates with a practically beneficial behavior, even when this practically beneficial behavior itself is not prevalent. Moreover, as the practical behavior becomes more prevalent, the symbolic norm becomes stronger.

Figure 21 Cooperation Rate Under Different Cost of Cooperation

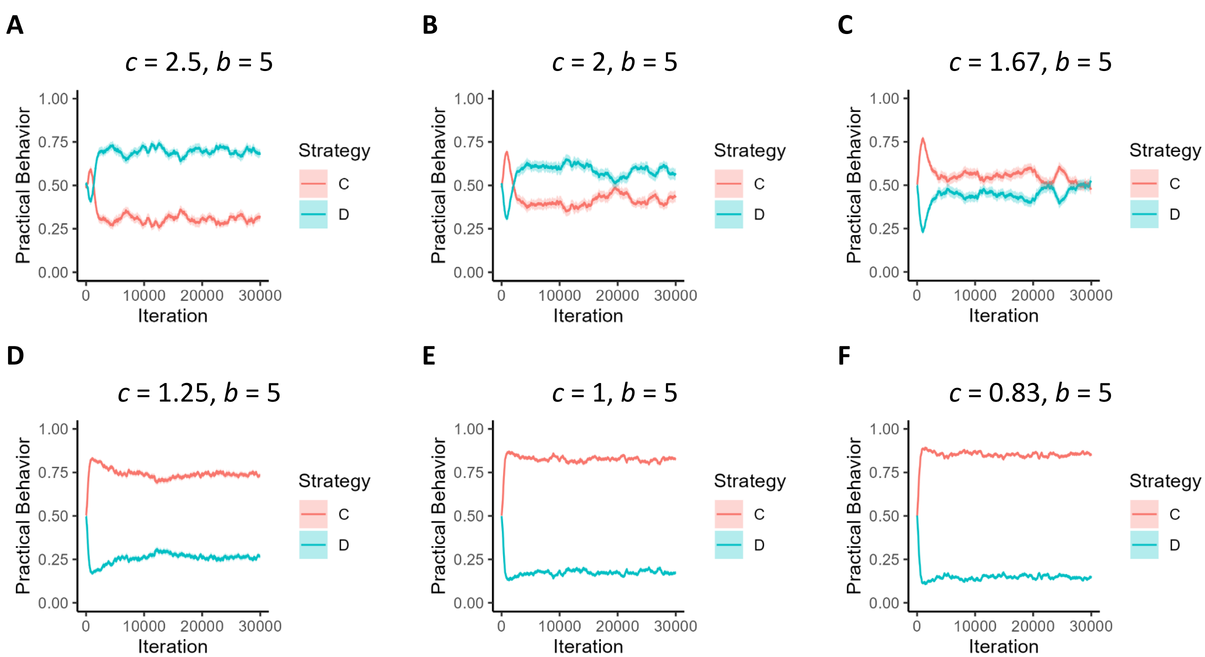


Figure 22 Symbolic Strategy Under Different Cost of Cooperation

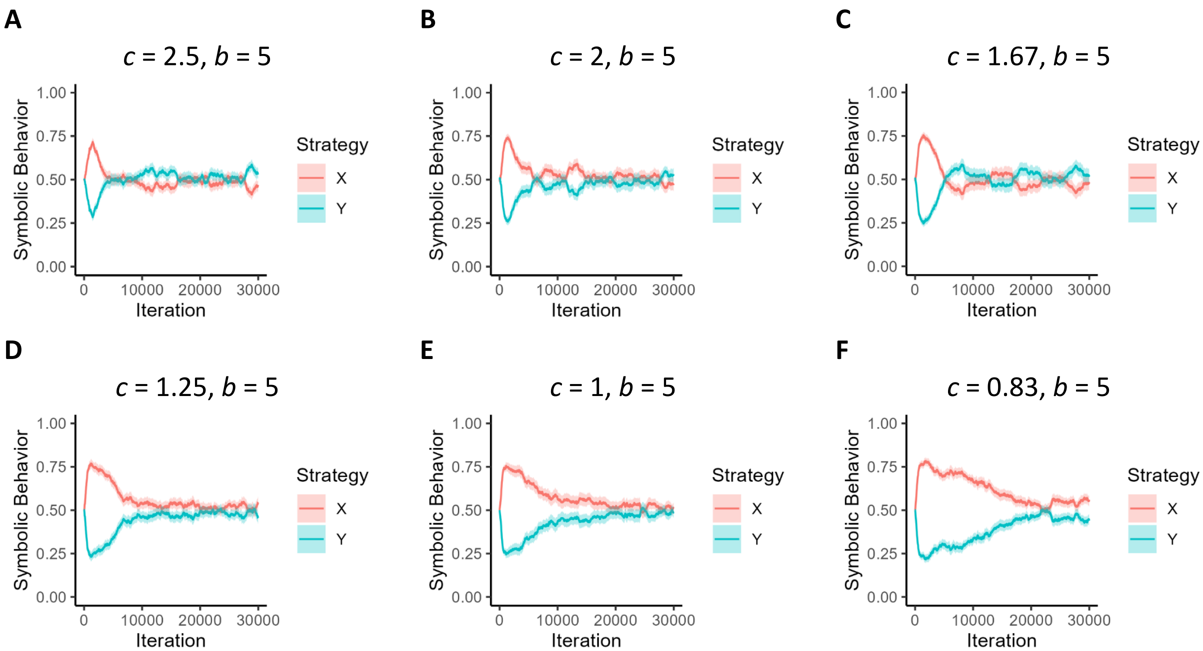
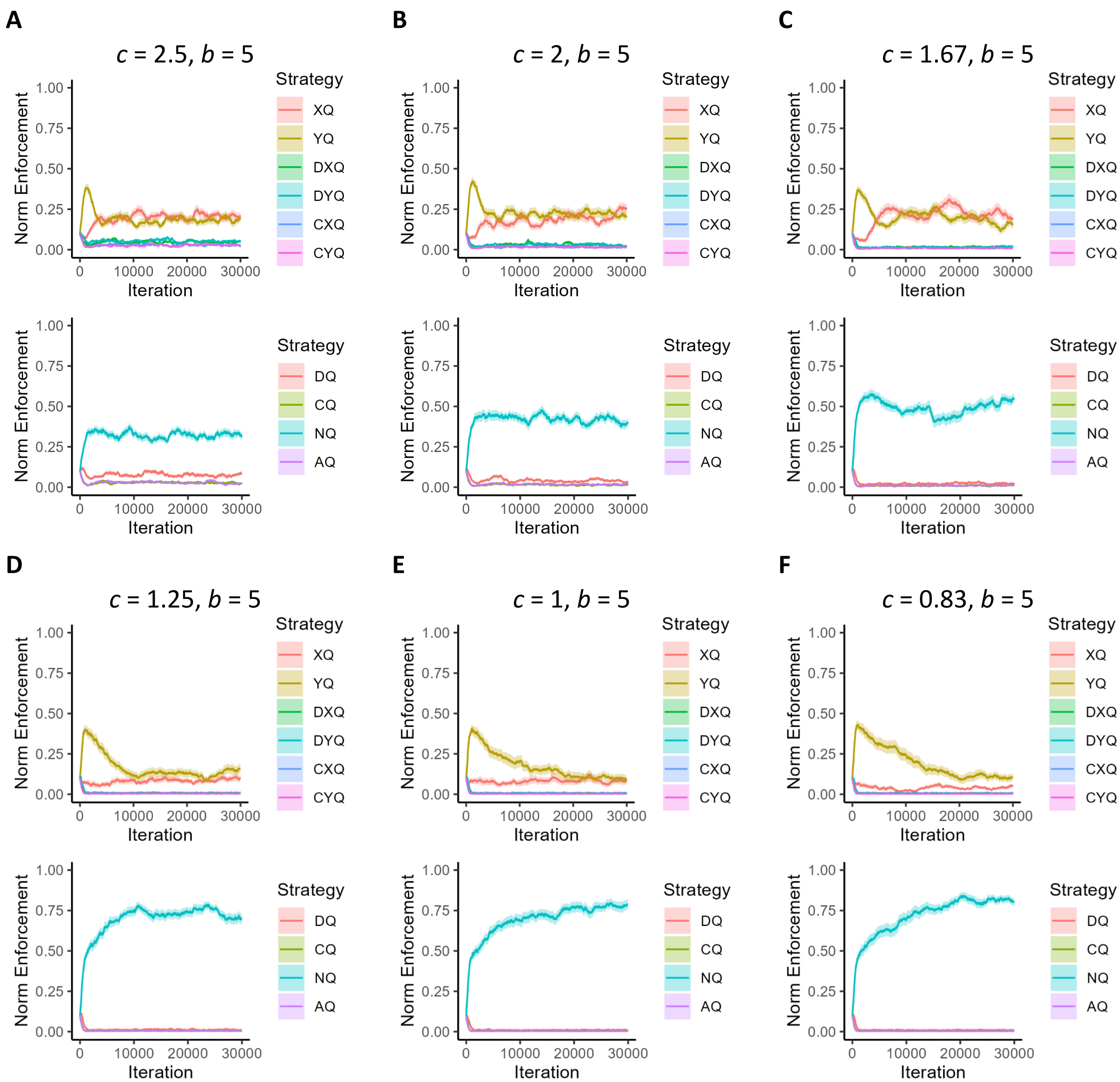


Figure 23 Norm Enforcement Under Different Cost of Cooperation



The Impact of Symbolic Norm on Practical Norm

Does the enforcement of symbolic norm benefit or compromise the practical norm? To examine the impact of the symbolic norm on practical norm, I did another set of simulations. The goal of this new set of simulations is to remove symbolic norm enforcers and compare the results with the original model.

This new model differs from the original model in the following ways: 1) when an XQ interacts with another agent, instead of ostracizing X-performers, an XQ ostracizes any agent with a probability of

0.5, 2) when a YQ interacts with another agent, instead of ostracizing Y-performers, a YQ also ostracizes any agent with a probability of 0.5, 3) when an DXQ/DYQ interacts with another agent, they ostracize all the defectors and ostracize cooperators with a probability of 0.5, and 4) when an CXQ/CYQ interacts with another agent, they ostracize all the cooperators and ostracize defectors with a probability of 0.5.

The rationale of this setup is to make the proportion of ostracizers in this new model match the proportion of ostracizers in the original model. If XQs and YQs (and DXQs, DYQs, CXQs, and CYQs) are directly removed from the population, the total amount of ostracism will change. Such change may influence cooperation rate and bring confounding variables to the contrast. Thus, to solve this problem, instead of removing these strategies directly, I keep the original strategy set, but make symbolic-based ostracism random. Instead of ostracizing X-performers or Y-performers specifically, XQs and YQs randomly ostracize half of the partners. Similarly, DXQs and DYQs ostracize all the defectors and half of the remaining cooperators; CXQs and CYQs ostracize all the cooperators and half of the remaining defectors. This way, no agent ostracizes others based on their symbolic strategy, but the total amount of ostracism can remain roughly the same.

Figure 24 compares the cooperation rate with vs. without symbolic enforcement in the early, middle, and late phases of the simulation. In all these three phases, the enforcement of symbolic norm boosts cooperation (see Supplementary Section **1.2.1 The Impact of Symbolic Enforcement When Symbolic Strategies Are Transparent** for the statistical tests). This result indicates that although it seems irrational to enforce a symbolic norm, being able to ostracize others based on their symbolic strategy helps agents avoid defectors and build mutual cooperation.

Interestingly, the cooperation rate is not influenced by the initial correlation between the symbolic and practical behaviors (i.e., σ value). The cooperation rate is higher with symbolic enforcement than without symbolic enforcement even when the symbolic strategy has no correlation with the practical behavior (i.e., $\sigma = 0$). This important result indicates that what benefits cooperation is *not* the symbolic norm, but the ability to categorize others based on a visible marker. Previous computer simulations on cooperation also showed that when agents are categorized by tags and can cooperate only with same-tag others, cooperation will emerge (Riolo et al., 2001). Based on this mechanism, the symbolic strategy in the current model can serve as such a tag and promote tag-based cooperation.

Notably, for a symbolic strategy to serve as a tag, it needs to be easily observable. If agents' symbolic strategies are non-transparent, a symbolic strategy cannot serve as a tag and thus cannot promote cooperation. **Figure 25** compares the cooperation rate with vs. without symbolic enforcement when symbolic strategies are non-transparent. When agents' symbolic strategies are not easily visible, their effect on promoting cooperation is minimal (see Supplementary Section **1.2.2 The Impact of Symbolic Enforcement When Symbolic Strategies Are Non-transparent** for the statistical tests).

In Supplementary Section **1.2.3 Another Model Option to Remove Symbolic Enforcement** in **Supplementary Materials**, I also tried another model setup to remove symbolic enforcement and the results remain the same.

Figure 24 The Impact of Symbolic Norm Enforcement on Cooperation

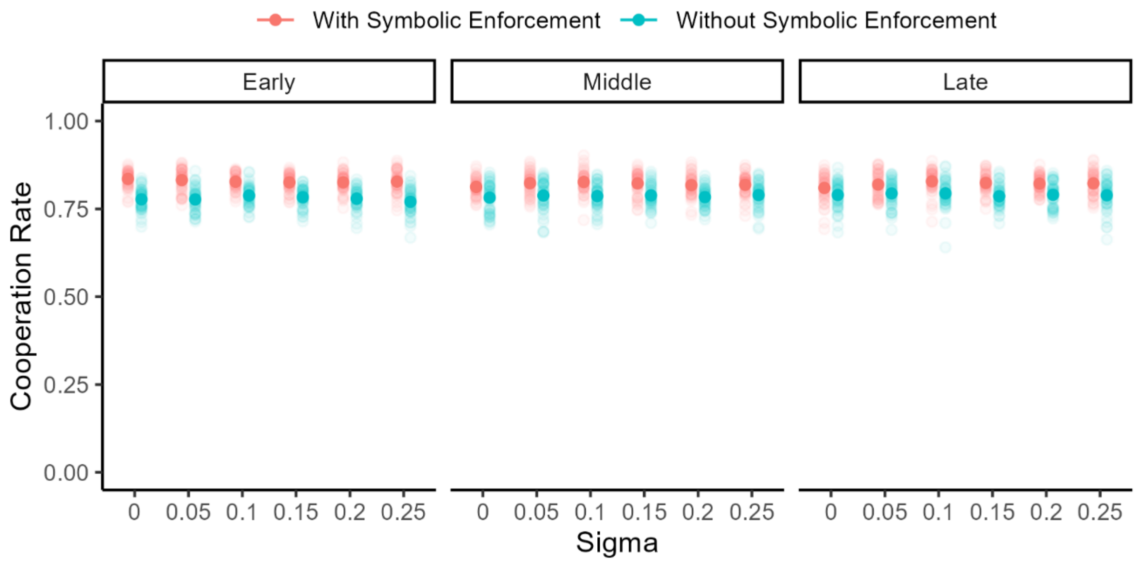
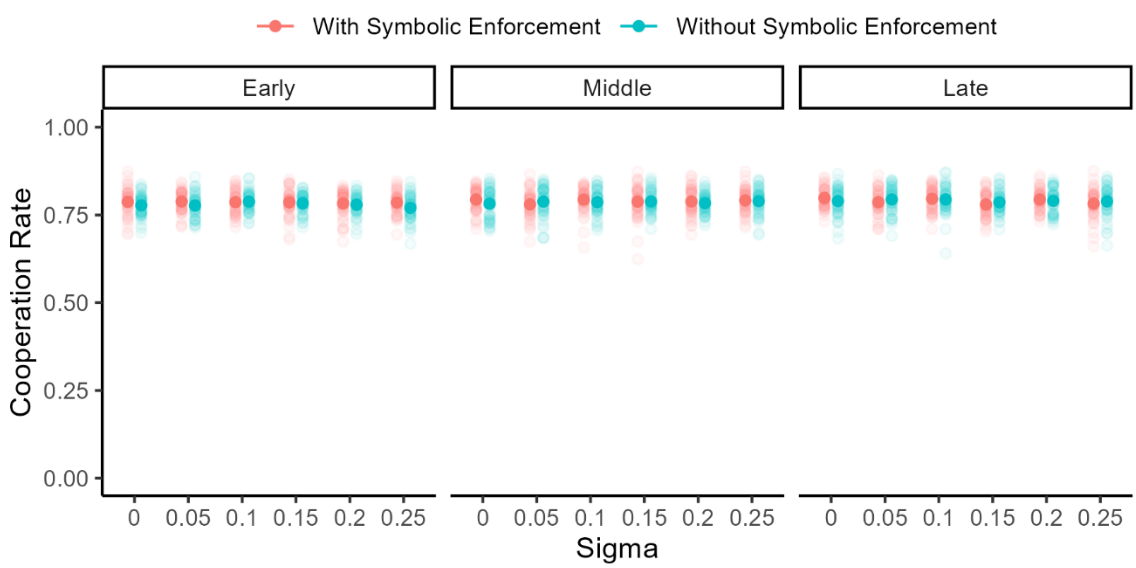


Figure 25 The Impact of Symbolic Norm Enforcement When Symbolic Strategies are Non-transparent



The Norm Spillover Effect Is Different From Tag-based Cooperation

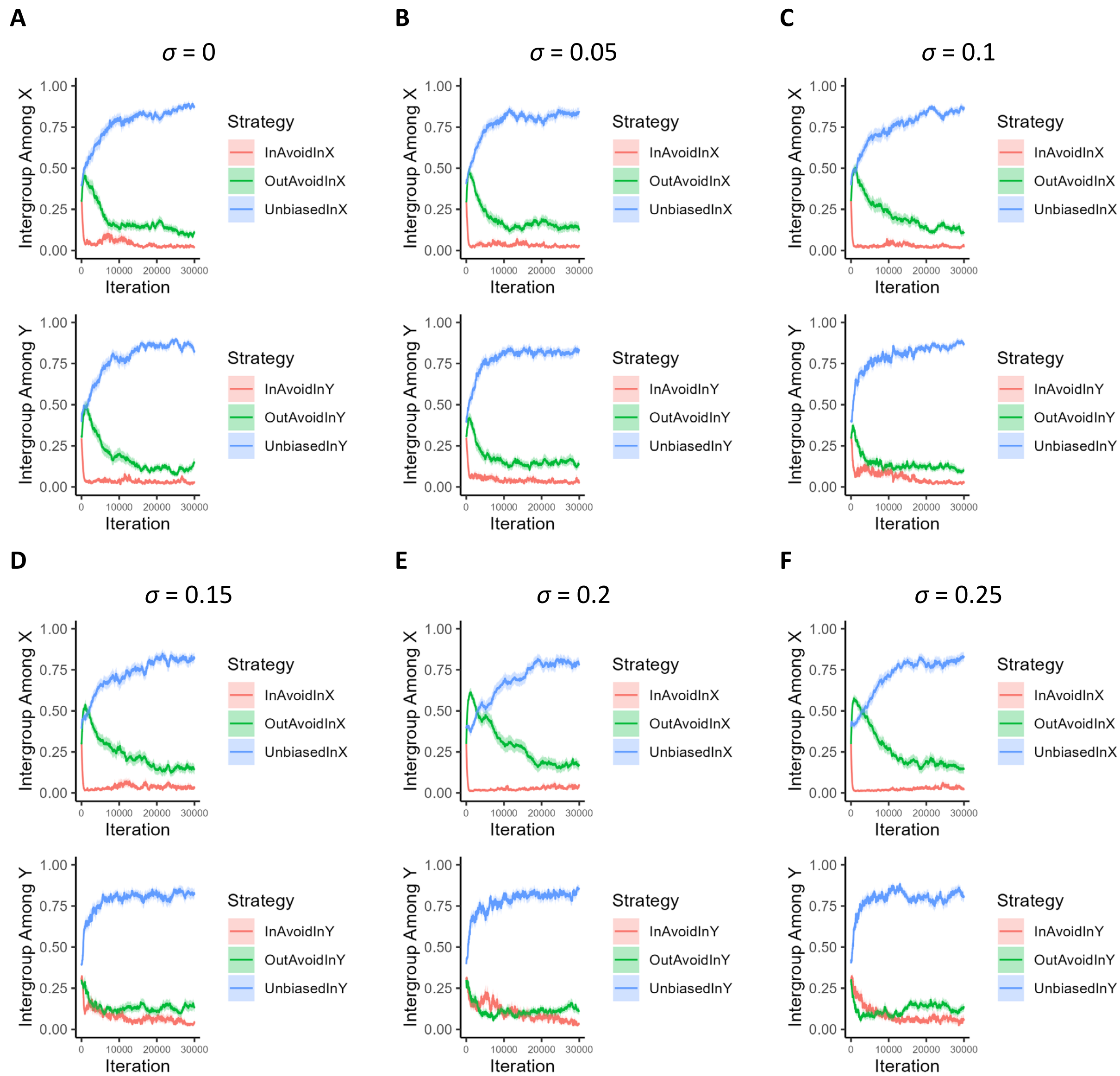
When a symbolic strategy correlates with a practical behavior, the mechanism behind the evolution of symbolic norm is different from tag-based cooperation. This section explains the difference between the norm spillover effect and tag-based cooperation.

In the models of tag-based cooperation, when agents are characterized with different tags and have the ability to condition their behaviors on others' group tags, agents will end up cooperating with same-tag others. Such mechanism has been widely studied as tag-based cooperation, which promotes the evolution of cooperation even among agents who have only rudimentary ability to detect environmental signals (Hammond & Axelrod, 2006; Riolo et al., 2001).

If my model is just a variation of previous models on tag-based cooperation, we should observe that X-performers prefer to interact with X-performers while Y-performers prefer to interact with Y-performers. However, **Figure 26** shows that this is *not* the case. Plots A-C in **Figure 26** show that when σ is low, both X-performers and Y-performers ostracize others who perform different symbolic strategies from themselves (i.e., outgroups). However, Plots D-F show that as σ gets higher, while X-performers still ostracize Y-performers, Y-performers do not ostracize X-performers. In fact, when σ is very high (i.e., $\sigma = 0.25$), more Y-performers ostracize their ingroup members.

This result suggests that tag-based cooperation and the norm spillover effect both exist in this model. When σ is small, tag-based cooperation plays a more prominent role. Under this circumstance, agents prefer to interact with ingroup members and ostracize outgroup members. However, as the correlation between the symbolic and the practically beneficial behavior becomes stronger, even symbolic norm violators prefer to interact with symbolic norm followers and ostracize other symbolic norm violators. Under this circumstance, the norm spillover effect becomes more prominent than ingroup favoritism.

Figure 26 The Trajectory of Ingroup and Outgroup Ostracism



Note. The trajectory of different ostracism strategies among X-performers and Y performers. In each plot, the upper part shows the strategies among X-performers and the lower part shows the strategies among Y-performers. In each plot, the red line, InAvoid, refers to the proportion of agents who ostracize others with the same symbolic strategy (i.e., ingroup). For X-performers, InAvoid are agents who are XQ, DXQ, and CXQ (i.e., ostracizing other X-performers). For Y-performers, InAvoid are agents who are YQ, DYQ, and CYQ. The green line, OutAvoid, refers to the proportion of agents who ostracize others with a different symbolic strategy (i.e., outgroup). For X-performers, OutAvoid are agents who are YQ, DYQ, and CYQ (i.e., ostracizing Y-performers). For Y-performers, OutAvoid are agents who are XQ, DXQ, and CXQ. The blue line, Unbiased, refers to the agents whose ostracism is not based on other's symbolic strategy. They are DQ, CQ, NQ, and AQ. When σ is low, there is more outgroup ostracism among both X-performers and Y-performers. However, as σ gets higher, more Y-performers start to ostracize ingroup. The norm spillover effect is stronger than ingroup favoritism under this circumstance.

Summary of the Mechanism Behind the Norm Spillover Effect

In this section, I discuss the mechanisms behind the norm spillover effect. First, the transparency of the symbolic strategy is the key to the norm spillover effect. Only when agents' symbolic strategies are easily visible, will agents enforce the symbolic norm. Second, the emergence of the symbolic norm does not require the emergence of a cooperative norm. Agents can enforce the symbolic choice that correlates with cooperative behavior even when the major population do not cooperate.

Allowing agents to ostracize others based on symbolic strategies promotes cooperation in the population. This is because agents' symbolic strategies can serve as an easily visible tag that categorizes individuals. When agents can interact only with similar others, tag-based cooperation will happen and promote cooperation. However, the norm spillover effect is more than just tag-based cooperation. When the correlation between the symbolic and practical behaviors is small, tag-based cooperation drives the results. However, as the correlation between the symbolic and practical behaviors gets stronger, the force of the norm spillover effect becomes stronger than tag-based cooperation and generates opposite results from tag-based cooperation.

Robustness Tests

To test the robustness of the results and make sure that the results are not due to one particular model setup, I perform robustness tests using a variety of different model options and parameter values. The goal of these tests is to find the boundary conditions of the norm spillover effect and explore how these parameters influence the strength of the symbolic norm.

These model options include 1) partial transparent cooperation and symbolic strategies, 2) noise in norm enforcement, 3) different levels of ecological threat, 4) different death rates, 5) different exploration rates and mutation methods, 6) two versions of pairwise comparison as the evolutionary dynamic, 7) well-mixed, random regular, and small world networks with different network degrees, 8) different levels of network mobility, and 9) punishment as the norm enforcement method.

The norm spillover effect remains robust across a broad range of model options. The symbolic norm is stronger when 1) individuals' symbolic strategies are more easily visible, 2) when individuals ostracize others strictly based on their strategies, 3) when ecological threat is higher, 4) when evolution happens more slowly, 5) when random mutation happens less frequently and when agents mutate their cooperation and symbolic strategies together, 6) when individuals have less connections on a random regular network, and 7) when network mobility is low. The norm spillover effect happens regardless of whether individuals' practical behaviors are visible or not. The norm spillover effect also happens in a well-mixed population. The norm spillover effect happens when individuals are embedded in a small-world network, although the effect of network degree is unclear for a small-world network. The norm spillover effect also happens when pairwise comparison is used as the evolutionary dynamic, regardless of whether agents keep or forget their memories after updating their strategies. The norm spillover effect can happen when punishment is used as the way to enforce norms, but the direction of norm spillover depends on whether punishment itself is benefiting or harming cooperation. When the existence of punishment promotes cooperation, the normal norm spillover effect will happen. However, when anti-social punishment evolves and punishment harms cooperation, the reverse norm spillover effect will happen, in which the symbolic strategy of the defectors will evolve as the norm. The detailed methods and results for these robustness tests can be found in Supplementary Section **2 Supplementary Methods and Results for Robustness Tests**.

Discussion

Summary of the Evolutionary Game Theoretic Model

In this chapter, I describe the results obtained from a series of agent-based models to examine the evolutionary basis of the norm spillover effect. These simulations show that when a neutral behavior without any practical benefit is correlated with a practically beneficial behavior on the population level, individuals will enforce this neutral behavior and a symbolic norm will emerge through a bottom-up process. The stronger the initial correlation is, the stronger the symbolic norm will be. However, as agents with different practical behaviors all start to follow the symbolic norm, the correlation between symbolic and practical behaviors decreases. A symbolic norm will eventually fade away when the correlation no longer exists in the population. Since the symbolic norm does not maintain, the emergence of a behavior under a certain condition can be quantified as the difference between that condition and the no correlation control ($\sigma = 0$) condition during a period of simulation. In other words, if the proportion of some behavior is higher under a certain level of σ compared with the no correlation condition during a certain period, I argue that this behavior has emerged under that condition in that period.

The simulations show that behaviors that are more easily observable are more susceptible to the norm spillover effect. This echoes previous findings that visible characteristics are more likely to be stigmatized (Stutterheim et al., 2011; Summers et al., 2018). In real life, the transparency of a behavior describes the extent to which a behavior is discernible or perceptible by others even before a substantial interaction happens (Summers et al., 2018). For example, outfit, etiquette, and office neatness are more susceptible to the norm spillover effect because these behaviors are easily observable even by strangers. On the contrary, one's hidden hobbies, how an individual treats their families, and the neatness of one's lockers may be less susceptible to the norm spillover effect because these behaviors cannot be easily known by others.

A neutral behavior may evolve as the norm even when the practically beneficial behavior correlated to it is not the norm. This result suggests that the tradition of a noble few can be adopted by the wider population even if the population cannot adopt all aspects of the noble few.

The simulations also show that when the initial correlation between symbolic and practical behaviors is strong enough, the norm spillover effect can overcome ingroup favoritism. Even symbolic norm violators themselves may enforce the symbolic norm. Such result shows that the mechanism of the norm spillover effect is different from tag-based cooperation (Riolo et al., 2001). Additionally, when the symbolic norm is related to a power difference between its followers and violators, such result can be related to internalized oppression, in which the oppressed group members choose to emulate with oppressors and discriminate against once own group (David & Derthick, 2023).

The robustness tests show that these findings hold robust across a variety of model choices, although the strength of the symbolic norm depends on ecological factors. For example, higher ecological threat leads to a stronger symbolic norm. Previous literature has found that higher ecological threat leads to stronger social norms in the service of social cooperation and coordination for survival (Gelfand et al., 2011; Roos et al., 2015). The robustness test is consistent and has extended this literature by showing that threat can lead to stronger norms even in the social domains that are not critical for survival. Cultures under higher ecological threat not only cooperate and coordinate more tightly to fight against threat, but also bestow symbolic importance on following normative behaviors in general (J. Jackson & Gelfand, 2016), through the norm spillover effect (see Supplementary Section ***2.3 Ecological Threat***).

Future Directions

Although the robustness tests have covered a variety of model options, some other variations of the model may be tried in the future to extend the scope of the current work.

First, the practical behavior chosen in the current model is cooperation behavior. Cooperation behavior has been widely used to study norm compliance and violation. A cooperative move can be considered as a norm following behavior because it is costly to perform but benefits others. On the other hand, a defective behavior is norm violation because it benefits the performer at the cost of others (Axelrod, 1986). Nevertheless, a norm following behavior may be described by other games, too. For example, a coordination game has also been used to capture the coordination component in social norms (Roos et al., 2015). Social norms such as deciding which side to drive on and which time zone to use are more of a coordination game than a cooperation game. Future models may test whether a symbolic norm will evolve if it is correlated with another behavior that people need to coordinate on. It is worth noting that if a coordination game is used, the population may converge to either behavioral option in the game (e.g., to drive on either the left or the right). Thus, if the norm spillover effect exists, which symbolic choice emerges will depend on which practical behavior evolves in the coordination game.

Second, following the rule of parsimony, the current model uses a single symbolic strategy that has only two options (i.e., X or Y). In real life, people choose different hair colors, wear different outfits, and decorate office spaces in various ways at the same time. There are multiple symbolic domains. One symbolic choice may correlate with multiple practical behaviors even in opposite directions. How do people navigate all these correlations, and which are the key factors that determine which symbolic norm stands out? According to the robustness tests, I hypothesize that the norm spillover effect is more likely to happen when the correlated practical behavior is more critical (i.e., higher benefit of following the norm), when the symbolic strategy is more easily visible, and when the correlation is stronger. In other words, if a behavior can serve as a salient and reliable marker of a critical practical behavior, that behavior is most likely to emerge as a symbolic norm. An example of such norm is the stigmatization of tattoos in Japan, due to tattoos being traditionally associated with yakuza organized crime gangs (Demetriou, 2015). Future research may design a model with multiple symbolic and practical behaviors to test these hypotheses.

Third, the current model assumes no cost or consequence for the symbolic choice. However, in real life, many symbolic norms can be very costly or induce negative outcomes, such as extreme rituals (Sterelny, 2014) and excessive drinking (Robertson & Tustin, 2018). Future research can examine whether or how the norm spillover effect can be strong enough to induce harmful symbolic norms.

Chapter 3: Empirical Experiment

While the norm spillover effect accounts for the emergence of symbolic norms on the population level, such effect needs to be supported by concurrent psychological processes so that individual decisions can make the norm spillover effect happen. In other words, a psychological process should exist that makes people adopt a symbolic choice when they perceive a correlation between this choice and a practically beneficial behavior. In this part of this dissertation, I describe an online experiment that examines the psychological processes behind the norm spillover effect on the individual level.

I hypothesize that such a psychological process is realized through two pathways. First is through the perceived direct function of the symbolic norm. I hypothesize that the perceived correlation increases the perceived direct function of the symbolic norm and thus increases its adoption. Second is through norm enforcement. I hypothesize that when a symbolic choice correlates with a practically beneficial behavior on the population level, performing the symbolic choice signals good traits in the performer so that people are more likely to select symbolic norm followers as partners. Such enforcement from others motivates people to adopt the symbolic norm themselves. In the section below, I introduce theories on these two pathways.

Psychological Processes in the Norm Spillover Effect

Adopting a Symbolic Norm When Perceiving Its Direct Function

When a neutral behavior correlates with a practically beneficial behavior on the population level, people may perceive direct functions in the neutral behavior. Humans see meanings and functions in behaviors even when they may not exist. When people see others engaging in a detailed course of actions, they tend to think that particular details of the action sequence lead to some desired, intended outcome (Legare & Souza, 2012). Ritual behaviors are perceived to be causally efficacious because of their frequency of repetition, number of procedural steps, specificity, and involvement of supernatural agency (Legare et al., 2016). People even see meanings from empty claims (Lin et al., 2022). It's not uncommon for people to see functions in ambiguously functional behaviors. Instead, they are prone to do so.

Moreover, people are more likely to see functions in a neutral behavior when that neutral behavior co-occurs with a practical behavior. Theories of illusionary causality show that people are attempted to infer a causal relationship when events occur in close temporal succession (Matute et al., 2015). If the best colleagues in a company happen to like wearing a fleece vest, then every time when working with someone in a fleece vest, the outcome is likely to be good. Though it seems irrational to infer that a fleece vest has anything to do with work performance, the more frequently the two events coincide, the more likely people perceive an illusory causality (Matute et al., 2015). In fact, illusions of causality have been widely used to explain the existence of pseudoscience, superstitions, and quackery (Matute et al., 2011). Combined with the fact that people are motivated to make sense of social norms, the co-occurrence of a practical and a neutral behavior on the population level may increase the perceived direct function of the neutral behavior.

Furthermore, when a neutral behavior becomes the majority, people are even more likely to see functions in it. The theory of informational social influence suggests that people conform with others because others' behavior provides information about the reality (Deutsch & Gerard, 1955). When people see many others eat certified functional food, they are more likely to believe that certified functional food is effective (Wang & Chu, 2021). Similarly, if people see a neutral behavior becomes normative, they are more likely to see functions in it. Such an inference of function will give rise to more conformity and a stronger motivation to see direct functions in the symbolic norm. As a result, a *positive feedback loop* will emerge and maintain the symbolic norm.

Perceiving Good Traits in Symbolic Norm Followers

When a neutral behavior correlates with a practically beneficial behavior on the population level, people may believe that the neutral behavior indicates some underlying trait of the performer that is related to a practical behavior. In this case, the behavior works either as a signal of a specific trait or a signal of the general trait of norm conformity, which will be elaborated below.

A neutral behavior works as a signal of a specific trait when people believe that there is some underlying feature that leads to both the symbolic and the practical behavior. People are prone to draw conclusions of others personalities (Tetlock, 1985). In Iran, tattoos, spiky haircuts, and the use of sunbeds on the grounds imply "devil worship" (*The World's Least Tattoo-Friendly Countries*, 2016). In Japan, having a tattoo is associated with yakuza organized crime gangs (Demetriou, 2015). A lawyer who has pink hair may be considered unprofessional while an artist with pink hair may be perceived creative. In these cases, performing a certain behavior indicates that this person also processes another trait that indicates their behavior in a practical domain. Note that this trait may or may not actually exist. For example, a person who keeps their office neat may indeed be more conscientious (Gosling et al., 2002), but it is questionable whether a lawyer without a suit is less competent (Cumberbatch, 2021). However, what is important is whether people believe in such a link. When people perceive the correlation on the population level, they are more likely to perceive an underlying trait that leads to both behaviors. Such perception makes the behavior a cue for predicting one's practical behavior.

In a special case, the symbolic choice may work not as a signal of a specific trait, but a signal of general norm conformity or commitment of the group. In an empirical study, researchers measured people's intention to conform to a non-beneficial ingroup norm. They found that peripheral group members are more likely to support a non-beneficial ingroup norm so that they can secure their membership status (Masson & Fritsche, 2019). Literature on ritual behavior also suggested that people perform ritual behaviors because these behaviors provide credible, difficult-to-fake cues of group loyalty, which helps promote trustworthiness and affiliation (Hobson et al., 2018). In this sense, following a symbolic norm signals one's general norm conformity and loyalty to the group. For a norm enforcer's perspective, whether an individual follows the symbolic norm signals whether they are a reliable ingroup member and interaction partner.

Partner Selection Based on the Symbolic Choice

When a symbolic choice is perceived to reflect the actor's trait, people will select their interaction partners based on the partners' symbolic choices. If someone believes that a good lawyer does not dye their hair pink, they will likely avoid working with a pink-hair lawyer. Whether they realize it or not, this is an enforcement of the hair-color norm, because the norm followers are rewarded by having a client while violators are punished by losing a client.

Consequently, if a person believes that others select their partners based on their symbolic choice, the person is also more likely to adopt the symbolic norm themself. If a lawyer believes that their potential clients are linking hair color to a lawyer's competency, the lawyer will avoid dying their hair. In this case, adopting a symbolic norm is the result of perceived norm enforcement from others.

Notably, people can be quite bad at guessing other's opinions. When participants were asked about their own and their friends' political attitudes, friends disagree considerably more than they think they do (Goel et al., 2010). In the worst case scenario, pluralistic ignorance may happen, in which most people privately dislike a norm but still follow it and publicly claim liking it (Prentice & Miller, 1996). In this sense, a symbolic signaling behavior will maintain as a norm as long as people believe that others select their partners based on the symbolic norm, even when such norm enforcement does not really exist.

Norm Dynamics After the Symbolic Norm Emerges

Finally, once a symbolic choice becomes normative, conformity is likely to happen and further facilitates the symbolic norm. On the one hand, conformity may happen because people believe in the function of a norm, when they think others' behavior provides information about the reality (Deutsch & Gerard, 1955). On the other hand, conformity may also happen due to the goal of affiliation (Cialdini & Goldstein, 2004; Sterelny, 2014). Moreover, people conform to social norms even when they understand that the norms are arbitrary and do not reflect the actual preferences of others (Pryor et al., 2018). In these cases, people follow a symbolic norm just because it is the norm. This dissertation will not discuss the conformity dynamics after the symbolic choice has become a norm.

To summarize, I hypothesize that there are two pathways that explain why perceived correlation between a neutral and a practically functional behavior induces the adoption of the symbolic norm. The first is through the perceived direct function of the symbolic norm and the second is through signaling and norm enforcement. It is worth mentioning that people may not realize the reason why they adopt a symbolic norm. Their reason may change over time, too. Self-perception theory suggests that people may perform a behavior and then observe and interpret their own behavior as they interpret someone else's (Bem, 1967, 1972). In this sense, it is possible that people may start adopting a symbolic norm unconsciously or under conformity pressure, but see its direct functions afterwards. In fact, multiple reasons may exist within a population and within an individual (Gross & Vostroknutov, 2022). The individual differences and the dynamic changes in perceived reasons can be explored in the future.

In the rest of this Chapter, I describe the empirical experiment that examines psychological processes behind the norm spillover effect.

Methods

Overview

In this study, participants are asked to imagine that they work in an organization and view the profiles of 28 colleagues. Each colleague performs behaviors in a practical and a symbolic domain. Same as the model, in the practical domain, the beneficial behavior is to follow a cooperative norm. In the symbolic domain, there are two options of behaviors: one is to meet at the start of work hours and the other is to meet at the end of work hours. The correlation between the symbolic and practical behaviors is manipulated. The ultimate dependent variable is participants' intention to adopt the symbolic choice. The main hypothesis is that, the correlation between the symbolic and practical behaviors leads people to adopt the symbolic choice that co-occurs more often with the practically beneficial behavior.

I also examine the psychological mechanisms that mediate the relationship between perceived correlation and behavioral intention. As described above, I hypothesize two pathways (see **Figure 27**). The first pathway is that the perceived correlation increases the perceived direct function of the symbolic norm and thus increases its adoption (H1 and H2). The second is that following the symbolic norm signals good traits in the performer so people are more likely to select symbolic norm followers as partners. Such enforcement from others motivates people to adopt the symbolic norm themselves (H3, H4, and H5). Specifically, the following variables are measured, including 1) the perceived correlation between the symbolic and practical behaviors, 2) the perceived direct function of the symbolic norm, 3) the perceived good traits in symbolic norm followers (signaling efficacy), 4) the intention to work with symbolic norm followers, 5) the perceived norm enforcement from others, and 6) the general perception of the symbolic norm. The following hypotheses are tested:

H1: The perceived correlation between a symbolic and a practical behavior increases the perceived direct function of the symbolic norm.

H2: The perceived direct function of a symbolic norm increases its adoption.

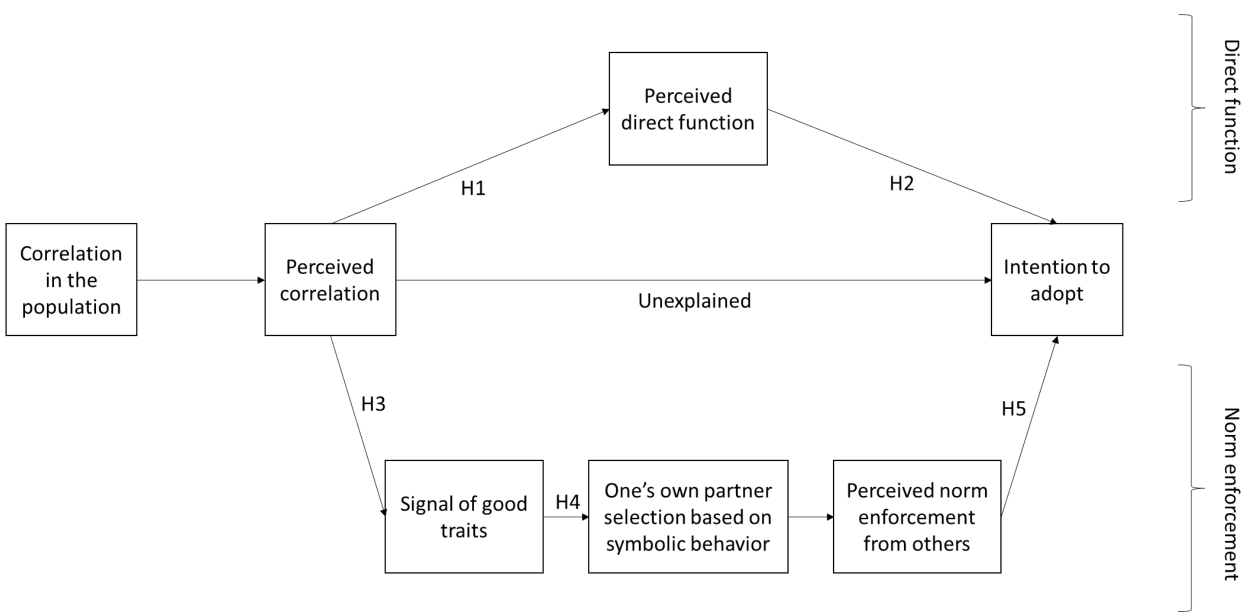
H3: The perceived correlation increases the likelihood that people believe the symbolic choice signals positive traits of its performer (i.e., signaling efficacy of the symbolic norm).

H4: The signaling efficacy leads to partner selection based on symbolic norm.

H5: Perceived enforcement of the symbolic norm increases norm adoption.

There is also a tentative hypothesis that one's own partner selection based on symbolic norm leads them to believe that others will also select partners based on symbolic norm. This was not included in the proposal but will be tested as an exploratory analysis.

Figure 27 Psychological Processes That Mediates the Norm Spillover Effect



Materials

After informed consent, participants are asked to imagine that "you just took a job at a new company and you want to get to know the people you will be working with." Participants are then presented with the profiles of 28 colleagues randomly sampled from this fictional company. They are instructed to "read each profile carefully and try to remember as much information as you can."

For each colleague, the profile includes their cooperative behavior (i.e., the practical behavior) and whether they prefer to schedule group meetings at the start or the end of work hours (i.e., the symbolic choice) (see **Figure 28**). The colleague's cooperative behavior is indicated by their internal ratings on four behaviors: 1) assist others during high workload, 2) caution others for potential problems, 3) take time to mentor new recruits, and 4) volunteer for company events. These items are selected from the Teamwork Perceptions Questionnaire (*Teamwork Perceptions Questionnaire (T-TPQ)*, 2017) and the Organizational Citizenship Behavior Checklist (Fox & Spector, 2009) to cover the nature of cooperative behaviors in an organizational setting. Notably, participants are not explicitly told that these behaviors are about cooperation. When a colleague has high cooperative ratings, their ratings on all the four items are between 4.4 and 5. When a colleague has low cooperative ratings, their ratings on all the four items are between 2 and 2.6.

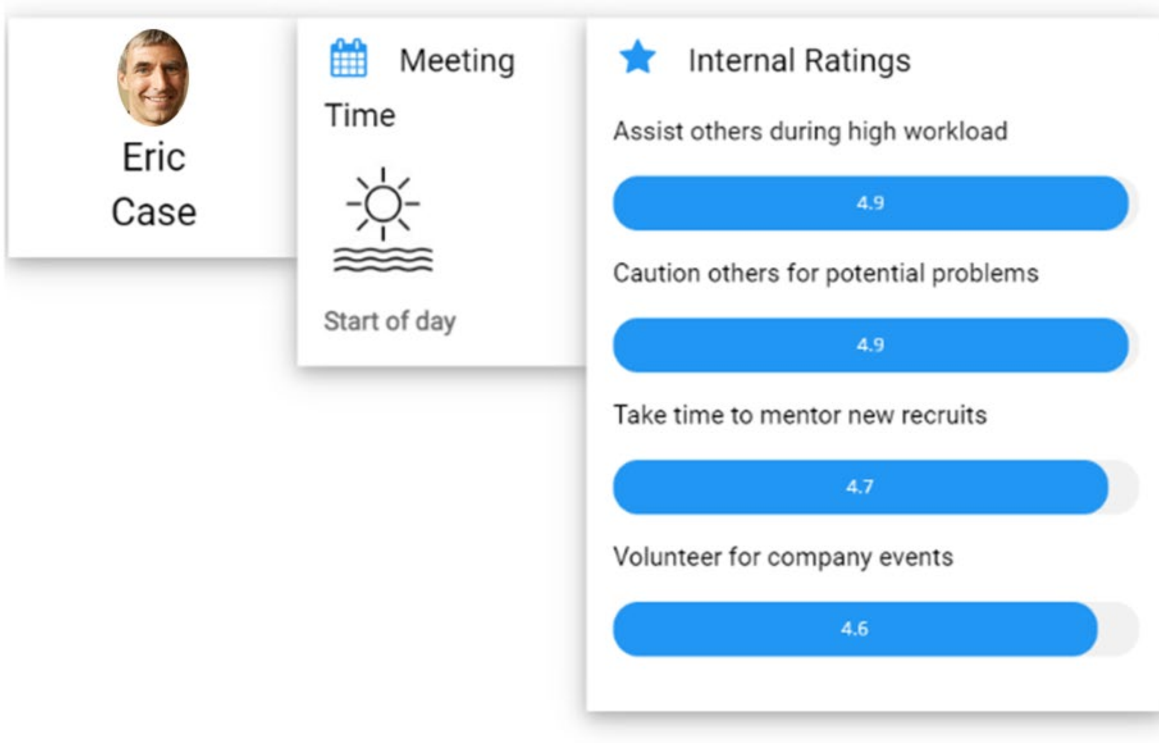
I chose meeting time preference as the symbolic norm for the following reasons. First, whether one prefers to meet at the start or the end of the day is an arbitrary behavior that can be interpreted as either good or bad. For example, preferring to meet at the start of the day can be interpreted as hardworking while preferring to meet at the end of the day can also be interpreted as hardworking if a colleague wants to save morning time for more important work. Second, meeting time preference is a work-related behavior, so it makes the experimental scenario more realistic. Third, meeting time preference does not require much coordination because participants do not necessarily need to meet at the same time with these colleagues. Fourth, meeting time preference is a relatively public behavior which should be susceptible to the norm spillover effect.

In addition, to make the scenario more realistic, a photo and a name of each colleague is provided. The photos include 22 White, 8 Hispanic or Latino, 6 Black or African American, and 4 Asian faces, based on the breakdown in The 2020 U.S. Census (Jones et al., 2021). Half of the faces are males. The faces are randomly selected from the 10k US Adult Faces Database by an R program (Bainbridge et al., 2013; RStudio Team, 2020). All the faces have been rated on multiple dimensions by Bainbridge et al. (2013). I select from the pictures that meet the following requirements: 1) the eyes are gazing the front, 2) the emotion is happy, 3) the picture is a good profile picture, 4) the image quality is larger than 3.5 on a 1-5 scale, 5) the face is facing the front, 6) the person is not famous, 7) the age of the figure is at least 20, and 8) the participants passed the attention check when giving these ratings.

As for the colleagues' names, their first names are randomly selected from the 21-100 most popular names by gender during the 1980s and 1990s in the U.S. (Social Security Administration, n.d.). Their last names are randomly selected from the 21-100 most popular surnames by race from the 2010 Census (Bureau, n.d.-b). The names are randomly selected by an R program (RStudio Team, 2020). The rationale is to make the names common enough so that they are not perceived as distinctive, but not so common in order to avoid acquaintances' names.

When presenting the profiles of these colleagues, the order of the 28 faces is randomized. The name shown below each face is randomly selected without replacement from the names that match the gender and race of the face. While the composition of these colleagues' behaviors is manipulated at the population level, which behavior is assigned to whom is randomized. Such randomization is implemented by a JavaScript code embedded in Qualtrics. The purpose of this randomization is to minimize the effects of faces and names on the results.

Figure 28 Example Profile of a Colleague



Note. The profile of a colleague who prefers to meet at the start of the day and has high cooperative ratings. The profile picture is not one of the actual pictures used in this experiment because the actual pictures are not publishable according to the license agreement.

Design

A between-group design is used. The independent variable (IV) is the correlation between the symbolic and practical behaviors in the population. The IV is operationalized by manipulating the proportion of different combinations of symbolic and practical behaviors (see **Table 4**). Five conditions are implemented. In the “start of day - strong correlation” (start-strong) condition, cooperation co-occurs more often with meeting at the start of the day and the correlation is strong. Thus, among the 28 colleagues, 12 of them have high cooperative ratings and prefer to meet at the start of the day; 12 of them have low cooperative ratings and prefer to meet at the end of the day; only 4 of them show the opposite pattern (see **Table 4**). In the “start of day - weak correlation” (start-weak) condition, cooperation co-occurs more often with meeting at the start of the day, but the correlation is weaker. In this condition, among the 28 colleagues, 10 of them have high cooperative ratings and prefer to meet at the start of the day; 10 of them have low cooperative ratings and prefer to meet at the end of the day; 8 of them show the opposite pattern. In the “no correlation” condition, cooperative and uncooperative colleagues are equally likely to prefer meeting at the start and the end of the day. In a similar fashion, in the “end of day - weak correlation” (end-weak) condition, cooperation co-occurs more often with meeting at the end of the day and the correlation is weak. In the “end of day - strong correlation” (end-strong) condition, cooperation co-occurs more often with meeting at the end of the day and the correlation is strong. Participants were randomly assigned to one of these five conditions by a JavaScript code in Qualtrics.

Table 4 Manipulating the Correlation Between Symbolic and Practical Behaviors

Behavior \ Condition	Start of day – strong correlation	Start of day – weak correlation	No correlation	End of day – weak correlation	End of day – strong correlation
High cooperation & meeting at the start of day	12 (42.9%)	10 (35.7%)	7 (25.0%)	4 (14.3%)	2 (7.1%)
High cooperation & meeting at the end of day	2 (7.1%)	4 (14.3%)	7 (25.0%)	10 (35.7%)	12 (42.9%)
Low cooperation & meeting at the start of day	2 (7.1%)	4 (14.3%)	7 (25.0%)	10 (35.7%)	12 (42.9%)
Low cooperation & meeting at the end of day	12 (42.9%)	10 (35.7%)	7 (25.0%)	4 (14.3%)	2 (7.1%)
Total	28 (100%)	28 (100%)	28 (100%)	28 (100%)	28 (100%)

Preregistration and Participants

Unless specified, the hypotheses, design plan, sampling plan, variables, and analysis plan of this experiment were all preregistered through the Open Science Framework (https://doi.org/10.17605/OSF.IO/H42ZY). The preregistration happened after pilot tests but before the data collection of the formal experiment.

A power analysis is done on the pilot data to decide the sample size using G*Power 3.1.9.7 (Faul et al., 2009). In the pilot study, the effect size f for the five-level one-way ANOVA on the main dependent variable (DV) – behavioral intention – is 0.538. Given this effect size, a sample of $N = 70$ is enough to get a significant ($\alpha = 0.05$) one-way ANOVA at the power (i.e., $1 - \beta$) of 0.95. However, it requires a much larger sample to detect the differences between conditions in post-hoc pairwise comparisons. For example, $N = 17$ per condition is needed to detect the difference between end-strong and no correlation conditions at a power of 0.95 (effect size for independent t-test $d = 1.300$, given $\alpha = 0.05$); $N = 51$ per condition is needed to detect the difference between end-weak and no correlation conditions at a power of 0.95 ($d = 0.724$); $N = 80$ per condition is needed to detect the difference between end-strong and end-weak conditions at a power of 0.95 ($d = 0.575$); $N = 254$ per condition is needed to detect the difference between start-strong and no correlation conditions at a power of 0.7 ($d = 0.221$); $N = 852$ per condition is needed to detect the difference between start-strong and start-weak conditions at a power of 0.7 ($d = 0.120$). Given the practical constraints, I *preregistered* to collect 625 participants (i.e., 125 per condition). Such sample size is enough to detect the differences between conditions with relatively large distances.

In addition, to test the structural equation model in **Figure 47**, the *N:q* rule suggests that the minimum sample size is 20 times the number of parameters require estimates (D. L. Jackson, 2003; Kline, 2015, p. 16). In this model, there are 15 parameters including pathways and variances. Thus, a minimum of 300 participants are needed. A sample of $N = 625$ satisfies this requirement.

625 participants over the age of 18 who are fluent in English and located in the U.S. were recruited through Prolific (https://www.prolific.co/). Participants are compensated $5.00 based on Prolific recommended rate. This research has been reviewed according to the University of Maryland, College Park IRB procedures for research involving human subjects (IRBNet Package: 2072708-2). Since participants' own preference of working in the morning vs. evening may influence their responses, 155 participants are recruited at around 10 am EDT, 156 participants are recruited at around 2 pm EDT, 155

participants are recruited at around 7 pm EDT, 4 participants are recruited at around 9 pm EDT, and 155 participants are recruited at around 11 pm EDT to cover a full range of time. Note that the local time is earlier for participants from later time zones.

Among the recruited participants, 609 participants completed all the blocks (i.e., Progress = 100). In the survey, there are two attention check questions. One asks the participants to select "extremely unlikely" and the other asks the participants to select "extremely likely." Participants who did not finish the study, who did not pass either attention check question, and who spent less than five minutes on the experiment were excluded. The data exclusion protocol has been preregistered. After this exclusion, there are 551 participants. 271 are self-identified as males and 258 are self-identified as females. Others are self-identified as non-binary, prefer to self-describe, or prefer not to say. The age composition can be found in **Table 10** in Supplementary Section ***3.1 The Age Composition of Participants***. 106, 90, 116, 118, and 121 participants were randomly assigned to the start-strong, start-weak, no correlation, end-weak, and end-strong conditions, respectively.

Measures

After reading the profiles of the 28 colleagues, participants are given the following tasks to measure the DVs, including 1) the perceived correlation between the symbolic and practical behaviors, 2) the intention to adopt the symbolic norm, 3) the perceived direct function of the symbolic norm, 4) the perceived good traits in symbolic norm followers, 5) one's own intention to work with symbolic norm followers, 6) the perceived norm enforcement of others, and 7) general perception of the symbolic norm.

To measure the perceived correlation between the symbolic and practical behaviors, participants are asked to estimate the following four frequencies in order:

1. Among the 28 colleagues you have seen, how many of them prefer to schedule meetings at the start of the day?
2. Among the colleagues who prefer to meet at the start of the day, how many of them have an average rating higher (lower) than 4?

3. Among the 28 colleagues you have seen, how many of them prefer to schedule meetings at the end of the day?
4. Among the colleagues who prefer to meet at the end of the day, how many of them have an average rating higher (lower) than 4?

For the second and the fourth questions, half of the participants are asked how many of the colleagues have an average rating *higher* than 4 and the other half are asked how many have an average rating *lower* than 4. Such counterbalance aims to decrease the anchoring effect (Furnham & Boo, 2011) – if the participants first see a question (i.e., the second question) that connects meeting at the start of the day with higher ratings, they may be anchored to believe that meeting at the start of the day is related to higher ratings.

This frequency estimation paradigm is adapted from a previous study on illusory correlation (D. L. Hamilton & Gifford, 1976). This task will be done immediately after the participants read the profiles. This task can also be considered as a manipulation check.

Each participant's estimations are put in a matrix as **Table *5***. The phi coefficient (ϕ) of this matrix is calculated as in Equation (4) to measure their perceived correlation between cooperative ratings and meeting at the end of the day. A positive ϕ (i.e., $\phi > 0$) means that the participant perceives that more cooperative colleagues prefer to meet at the *end* of the day. A negative ϕ (i.e., $\phi < 0$) means that the participant perceives that more cooperative colleagues prefer to meet at the *start* of the day. The larger the absolute value of ϕ (i.e., $|\ \phi\ |$) is, the stronger correlation the participant perceives. The theoretical ϕ values of the start-strong, start-weak, no correlation, end-weak, and end-strong conditions are -0.714, -0.429, 0, 0.429, and 0.714 respectively.

Note that, logically, the answer to second question should not exceed the answer to the first question and the answer to the fourth question should not exceed the answer to the third question. Thus, if a participant's answers do not meet these requirements, their phi coefficient will not be computed. Such missing data protocol has been preregistered. The phi coefficient was successfully computed among 484 participants.

Table 5 Matrix of Perceived Correlation

	Prefer to meet at the start of day	Prefer to meet at the end of day	Total
High cooperative rating	n_{CX}	n_{CY}	n_C
Low cooperative rating	n_{DX}	n_{DY}	n_D
Total	n_X	n_Y	n

$$\phi = -\frac{n_{CX}n_{DY} - n_{CY}n_{DX}}{\sqrt{n_C n_D n_X n_Y}} \quad (4)$$

To measure the main DV, the intention to adopt the symbolic norm, participants are asked, on a 1-5 scale, about "if you work in this company, will you schedule the meetings at the start or the end of the day" (1 – Definitely start of day; 5 – Definitely end of day). This question is asked immediately after the manipulation check to avoid priming effects from other questions.

To measure the perceived direct function of the symbolic norm, participants are asked, on a 1-5 scale, whether they think meeting at the start/end of the day directly increases task performance. The items are adapted from the Individual Workplace Performance Questionnaire (Ramos-Villagrasa et al., 2019). Specifically, the following items are used:

1. Scheduling meetings at the start of the day helps people in this organization finish their work on time.
2. Scheduling meetings at the start of the day helps people in this organization carry out work more efficiently.
3. Scheduling meetings at the end of the day helps people in this organization finish their work on time.
4. Scheduling meetings at the end of the day helps people in this organization carry out work more efficiently.

The average of the first two items measures the perceived direct function of meeting at the start of the day. The average of the last two items measures the perceived direct function of meeting at the end of the day. The difference between these two average values is computed as the final measure of the perceived direct function of meeting at the end (as against to the start) of the day. This is to control acquiescent responding and within-group variance. The order of the first two items is randomized. The order of the last two items is also randomized. The last two questions are always presented after the first two questions to avoid confusion.

To measure the signaling efficacy of the symbolic norm, participants are asked, on a 1-5 scale, "to what extent do you think a colleague in this organization who prefers to schedule meetings at the start/end of the day is likely to possess the following traits." The traits fall into four dimensions: warmth, competence, conformity, and commitment. The first two dimensions are considered universal dimensions of social cognition (Fiske et al., 2007). Conformity and commitment are included to test whether people perceive symbolic norm followers to be more likely to follow general social norms and committed to the group. On each dimension, five traits are selected. Warmth and competence traits are selected from Fiske et al. (2007). Conformity related traits are inspired by a scale on tendency to conform (Goldsmith et al., 2005). Commitment related items are adapted from the Organizational Commitment Questionnaire (Mowday et al., 1979). The traits are listed in **Table 6**.

Table 6 Traits That the Symbolic Choice Signals

Dimension	Traits
Warmth	Sociable, sincere, warm, good natured, helpful
Competence	Persistent, determined, industrious, skillful, intelligent
Conformity	Compliant, agreeing, cooperative, obedient, accommodating
Commitment	Loyal to this organization, glad to join this organization, stick with this organization, care about the fate of this organization, make efforts to help this organization succeed

These 20 traits are presented in a random order. Participants are first asked to rate for colleagues who prefer to meet at the start of the day and then asked to rate for colleagues who prefer to meet at the end of the day. For each trait, the difference between the two ratings is computed as the signaling efficacy of meeting at the end of the day (as against the start of the day).

To measure the intention to enforce the symbolic norm, participants are asked about the extent to which participants choose to work with the colleagues who meet at the start vs. end of the day. This section contains two blocks.

In one block, participants are presented with the profiles of two new White male faces. One prefers to meet at the start of the day and the other prefers to meet at the end of the day. They both have mediocre cooperation ratings between 3.4-3.6. In the other block, participants are also presented with the profiles of two new White male faces with one preferring meeting at the start and the other at the end of the day. However, this time, their cooperation ratings are not shown. After reading each profile, participants are asked to rate, on a 1-5 scale, on the following questions: 1) I want to work with him; 2) I want him to join my team; 3) I trust him; 4) I like him; and 5) I feel similar to him.

The faces are selected from all White male faces from the 10k US Adult Faces Database (Bainbridge et al., 2013) with the most similar traits (see Supplementary Section ***3.2 The Selection of the Four Faces in the Partner Selection Task*** for the method for selecting these four faces). The order of the faces and their names are randomized by a JavaScript Code in Qualtrics. The order of the two blocks is randomized. Within each block, the colleague who prefers to meet at the start of the day is always shown before that who prefers to meet at the end of the day to avoid confusion. The order of the five items is randomized.

To measure the perceived norm enforcement from others, participants are asked to what extent they believe others will select to work with the colleagues who meet at the start vs. end of the day. The same stimuli as for one's own partner selection are used. For each profile, after providing ratings for themselves, participants are asked about to what extent they think *others* in this company 1) want to work with him, 2) want him to join their teams, 3) trust him, 4) like him, and 5) feel similar to him.

To measure the general perception of the symbolic norm, participants are asked the following questions about their general perception of the symbolic norm in this company:

1. People in this company usually schedule meetings at the start/end of the day.
2. People in this company should schedule meetings at the start/end of the day.
3. People in this company will expect me to schedule meetings at the start/end of the day.
4. People in this company will disapprove if I schedule meetings at the end/start of the day.
5. I will schedule meetings at the start/end of the day because I expect that others prefer to meet at the start/end of the day.

Items 1-2 aim to capture the perceived descriptive and injunctive norms (Cialdini et al., 1990). Items 3-4 aim to capture the pressure from others. Item 5 aims to capture the conformity to others' behavior. Items 1-5 together capture the interactive perspective of a social norm, which is "everyone conforms, everyone is expected to conform, and everyone wants to conform when they expect everyone else to conform" (Young, 2015, p. 359).

The questions are divided into two blocks. One for the start of the day and the other for the end of the day. The difference between the start and the end of the day is computed as the perceived symbolic norm of meeting at the end of the day (as against the start of the day). The start-of-day block always appears before the end-of-day block. Within each block, the order of the questions is randomized.

Finally, some demographic and control variables are measured, including 1) age, 2) gender, 3) income, 4) the US state that the participant resides, 5) employment status, 6) whether the person is a morning or night person, and 7) the strategies that the participant used when doing this study, inspired by Haslam et al. (1996) (see Supplementary Section ***3.4 The Strategy that Participants Used When Answering the Questions***). If the participant is working full-time or part-time, questions about the participant's workplace will be asked, including 1) meeting time preference of the participant's colleagues, 2) the participant's own meeting time preference, 3) the descriptive norm for meeting time in the participant's workplace, 4) tightness in the participant's workplace, adapted from Gelfand et al.

(2011) (see Supplementary Section ***3.3 The Strength of Social Norms in the Participant's Workplace***), and 5) the industry of the participant.

The question on perceived correlation is presented immediately after reading the profiles since this is the manipulation check. The question on the intention to adopt the symbolic norm is asked immediately afterwards to avoid the priming effects from the following questions. The blocks on perceived direct function, signaling efficacy, partner selection, and the general perception of norm are presented in a random order. The control variables are measured at the end. Participants are debriefed on the purpose of the study after finishing the experiment.

Results

Manipulation Check and Perceived Correlation

When the participants were reading the 28 files, six questions were implemented between profiles asking participants to recall the information in the profile that they had just seen. These questions aim to keep the participants' attention. The participants answered the questions with an accuracy of 0.93 (SD = 0.13), showing that participants were reading the profiles carefully.

Participants successfully perceived the direction of the correlation between symbolic and practical behaviors. In the start-strong and start-weak conditions, participants perceived that meeting at the start of the day co-occurs more often with higher cooperative ratings (i.e., $\phi < 0$; t-test from 0 for the start-strong condition $t(98) = -12.46$, $p < 0.001$; t-test from 0 for the start-weak condition $t(76) = -8.16$, $p < 0.001$). In the no correlation condition, the participants perceived no correlation between meeting time and cooperative behavior (t-test from 0 for the no correlation condition $t(97) = -1.10$, $p = 0.272$). In the end-weak and end-strong conditions, participants perceived that meeting at the end of the day co-occurs more often with higher cooperative ratings (i.e., $\phi > 0$; t-test from 0 for the end-weak condition $t(102) = 4.85$, $p < 0.001$; t-test from 0 for the end-strong condition $t(106) = 9.68$, $p < 0.001$) (see **Table 7** for the statistics).

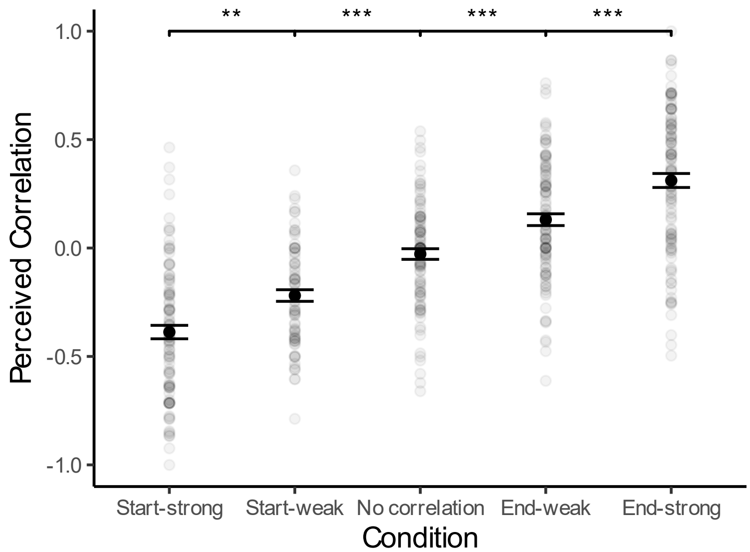
Participants also correctly perceived the relative strength of the correlations. A one-way ANOVA showed that the perceived correlation is significantly different across the five conditions ($F(4, 479)$ =

94.29, $p < 0.001$). Tukey's Honest Significant Difference (HSD) post hoc tests show that the ϕ in the start-strong condition is significantly smaller than the start-weak condition ($p = 0.001$). The ϕ in the start-weak condition is significantly smaller than the no correlation condition ($p < 0.001$). The ϕ in the no correlation condition is significantly smaller than the end-weak condition ($p < 0.001$). And the ϕ in the end-weak condition is significantly smaller than the end-strong condition ($p < 0.001$) (see **Figure 29**). These results show that the participants successfully perceived the correlation between the symbolic and practical behaviors in the population. The manipulation of DV is successful.

Table 7 Means and Standard Errors for the Variables in Different Conditions

Condition	*N*	Perceived correlation	Behavioral intention	Perceived direct function	Signaling efficacy	Partner selection (practical known)	Partner selection (practical unknown)	Perceived enforcement from others (practical known)	Perceived enforcement from others (practical unknown)	General perception of norm
Start-strong	106	-0.39	1.70	-1.95	-1.06	-0.24	-0.45	-0.23	-0.43	-1.42
		(0.031)	(0.085)	(0.149)	(0.097)	(0.058)	(0.077)	(0.064)	(0.090)	(0.122)
Start-weak	90	-0.22	2.18	-1.43	-0.66	-0.29	-0.23	-0.09	-0.14	-0.93
		(0.027)	(0.111)	(0.169)	(0.091)	(0.065)	(0.093)	(0.066)	(0.080)	(0.124)
No correlation	116	-0.03	2.50	-0.98	-0.24	-0.05	0.00	-0.03	-0.02	-0.48
		(0.025)	(0.115)	(0.157)	(0.075)	(0.070)	(0.078)	(0.054)	(0.066)	(0.122)
End-weak	118	0.13	3.13	-0.08	0.22	0.09	0.23	-0.02	0.21	0.01
		(0.027)	(0.124)	(0.189)	(0.098)	(0.065)	(0.076)	(0.058)	(0.073)	(0.135)
End-strong	121	0.31	3.56	0.80	0.82	0.21	0.38	0.04	0.34	0.69
		(0.032)	(0.127)	(0.191)	(0.100)	(0.067)	(0.075)	(0.054)	(0.078)	(0.136)

Figure 29 Perceived Correlation Across Conditions



Main DV: Intention to Adopt the Symbolic Norm

The main DV is the intention to adopt the symbolic norm. A one-way ANOVA shows that the behavioral intention is significantly different across the five conditions ($F(4, 546) = 41.83$, $p < 0.001$). As meeting at the end of the day co-occurs more often with cooperative behavior, participants are more likely to meet at the end of the day (**Figure 30**; see **Table 7** for the statistics). Post hoc tests show that the difference between start-strong and start-weak conditions is significant ($p = 0.047$). Although the difference between start-weak and no correlation conditions is not significant ($p = 0.323$), the trend is as expected. The difference between no correlation and end-weak conditions is significant ($p < 0.001$). And the difference between end-weak and end-strong conditions is significant ($p = 0.046$).

Figure 31 shows the correlation between perceived correlation and behavioral intention. When participants perceive that meeting at the end of the day co-occurs more often with cooperative behavior in the population, they are more likely to choose to meet at the end of the day ($r(482) = 0.587$, $p < 0.001$). The main hypothesis is supported.

Figure 30 Behavioral Intention Across Conditions

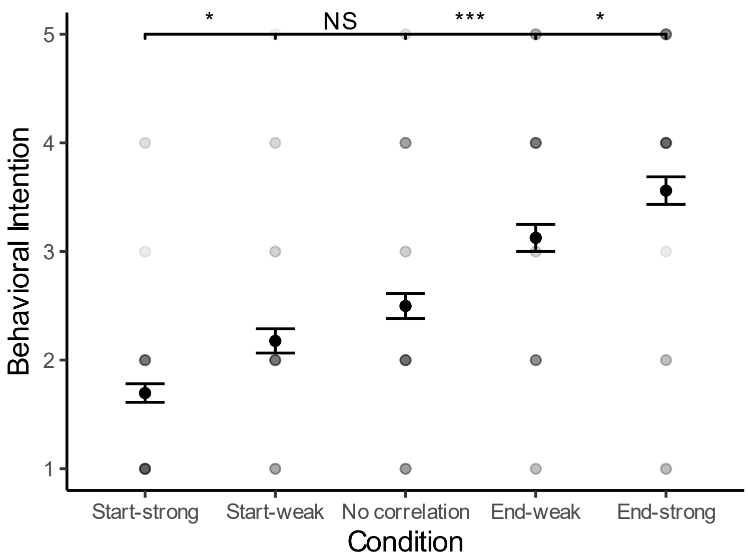
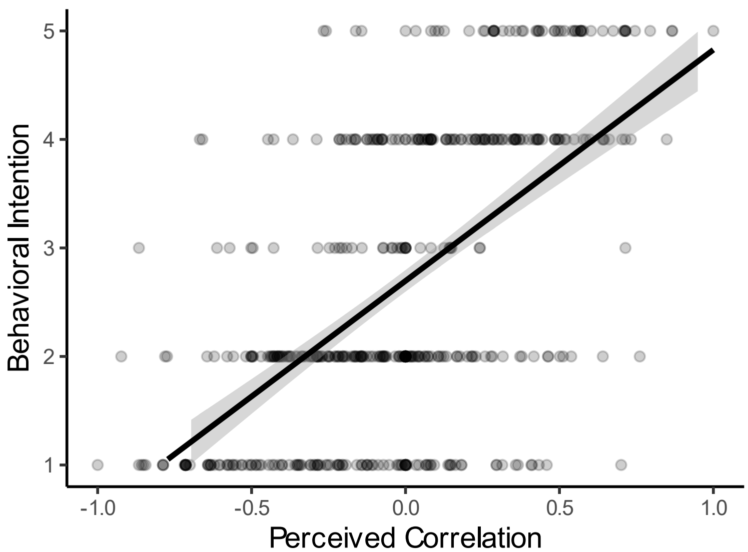


Figure 31 The Correlation Between Perceived Correlation and Meeting at the End of Day



Perceived Direct Function of the Symbolic Norm

The perceived direct function of the symbolic norm is measured as the extent to which participants think meeting at the end of the day directly increases task performance (as against meeting at the start of the day). I hypothesize that the perceived correlation between symbolic and practical behaviors increases the perceived direct function of that symbolic norm (H1). **Figure 32** supports this hypothesis. The correlation between perceived correlation (i.e., ϕ) and the perceived direct function of meeting at the end of the day (against the start of the day) is significant ($r(473) = 0.62$, $p < 0.001$). **Figure 33** also shows that the perceived direct function of the symbolic norm increases as the symbolic choice co-occurs more often with the practically functional behavior (one-way ANOVA $F(4, 535) = 40.34$, $p < 0.001$; see **Table 7** for the statistics across the five conditions; see Supplementary Section ***4.1 Perceived Direction Function*** for post hoc tests).

Figure 32 The Correlation Between Perceived Correlation and Perceived Direct Function

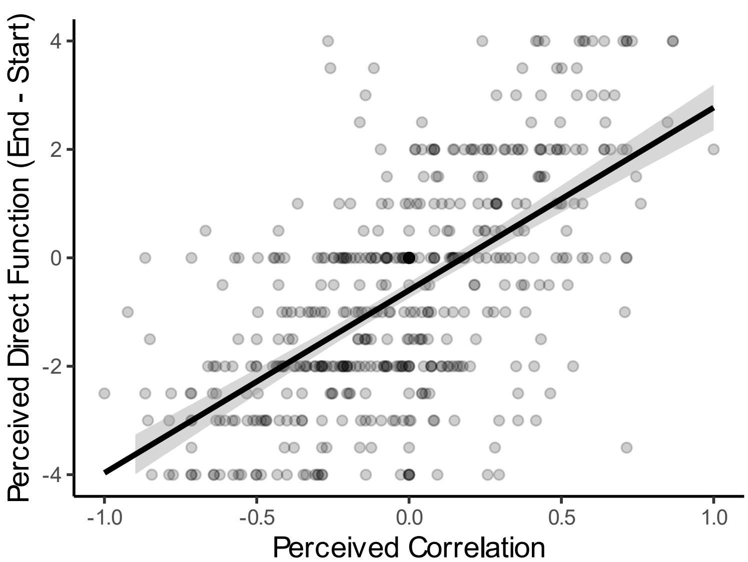
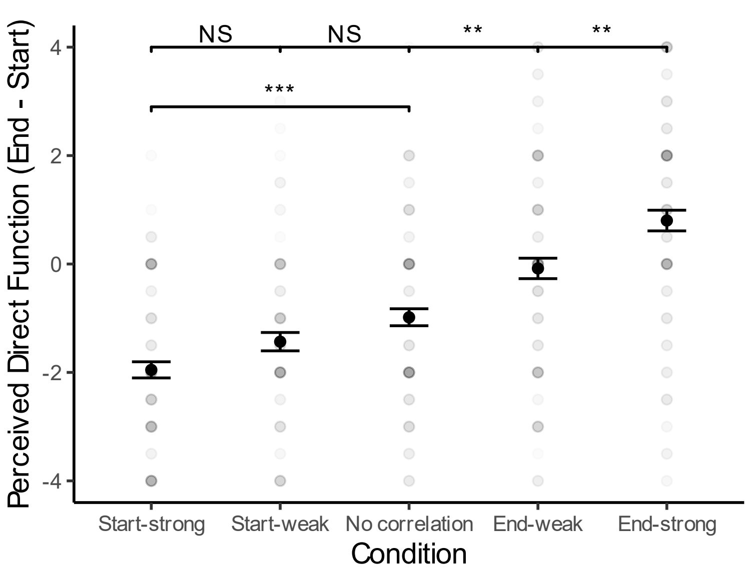
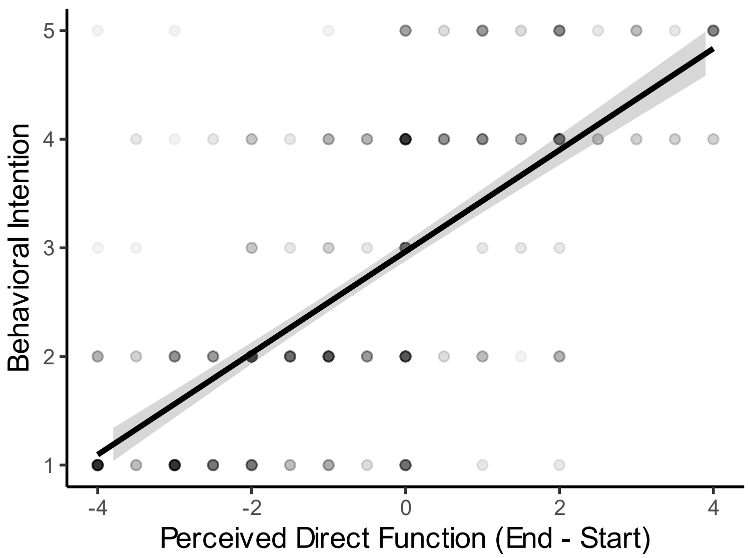


Figure 33 Perceived Direct Function of Meeting at the End (as Against Start) of Day



I also hypothesize that the perceived direct function of a symbolic norm increases its adoption (H2). **Figure 34** shows that as the perceived direct function increases, participants are more likely to meet at the end of the day ($r(538) = 0.70$, $p < 0.001$).

Figure 34 The Correlation Between Direct Function and Behavioral Intention



Signaling Efficacy of the Symbolic Norm

Participants are asked whether they think meeting at the start/end of the day signals good traits. 20 traits on four dimensions (warmth, competence, conformity, and commitment) are measured. An exploratory factor analysis shows that the 20 traits fall on a single factor and the loadings of all the items are high (see Supplementary Section ***4.2.1 Factor Analysis Across the 20 Traits***). Thus, the 20 traits are averaged as a single measure of perceived good traits on colleagues who meet at the end (against the start) of the day.

I hypothesize that when a symbolic choice is positively correlated with a practically beneficial behavior, performing the symbolic choice signals good traits of the performer (H3). **Figure 35** supports this hypothesis ($r(452) = 0.71$, $p < 0.001$). **Figure 36** also shows that when meeting at the end of the day co-occurs more often with cooperation, participants are more likely to believe that meeting at the end of the day signals good traits. The one-way ANOVA shows significant difference between the conditions ($F(4, 510) = 63.04$, $p < 0.001$). The post hoc tests show that all the pairwise comparisons are significant (start-strong vs. start-weak: $p = 0.038$; start-weak vs. no correlation: $p = 0.021$; no correlation vs. end-weak: $p = 0.005$; end-weak vs. end-strong: $p < 0.001$; see **Table 7** for the statistics).

I also analyzed the traits on warmth, competence, conformity, and commitment separately and the results remain the same (see Supplementary Section **4.2.2 Results When the Four Dimensions of Traits Are Analyzed Separately**).

Figure 35 The Correlation Between Perceived Correlation and Perceived Good Traits

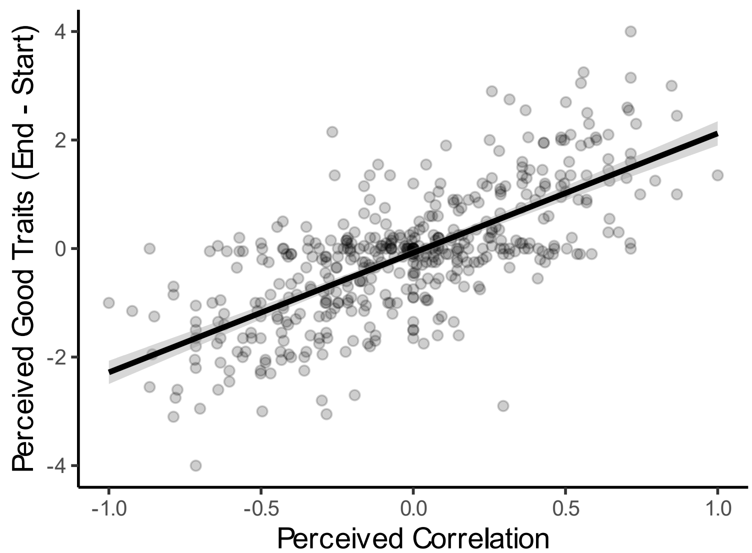
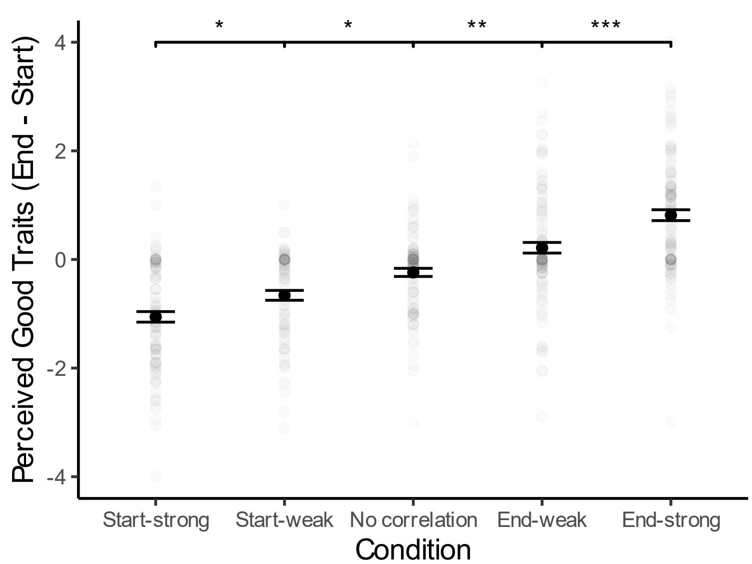


Figure 36 Perceived Good Traits in Those Who Meet at the End (as Against Start) of Day



Partner Selection Based on Symbolic Norm

Participants are asked about to which extent they want to work with a colleague who meets at the start vs. end of the day. Two tasks are used. In the first task, the cooperative behavior of the colleague is known to be mediocre. In the second task, their cooperative behavior is unknown. Five items are used to measure the intention to work with the colleague. Factor analyses show that the five items fall onto a single factor, so the items are averaged (see Supplementary Section **4.3.1 Factor Analysis Across the Five Partner Selection Items**). The final index of partner selection measures the extend to which the participants want to work with a colleague who meets at the end (as against the start) of the day.

I hypothesize that the signaling efficacy of a symbolic choice leads to partner selection based on that symbolic choice (H4). This hypothesis is supported by **Figure 37** and **Figure 38**. The more participants think that meeting at the end of the day signals good traits, the more they prefer to work with a colleague who meets at the end of the day. The correlation is significant both when the colleague's practical behavior is known ($r(503) = 0.34$, $p < 0.001$) and when the colleague's practical behavior is unknown ($r(505) = 0.48$, $p < 0.001$). Notably, the correlation is stronger when the practical behavior is unknown ($\Delta r = 0.14$, $z = 3.74$, $p < 0.001$). This is because when the practical behavior is known, people rely less on the symbolic norm.

Figure 37 Perceived Good Traits and Partner Selection With Known Practical Behavior

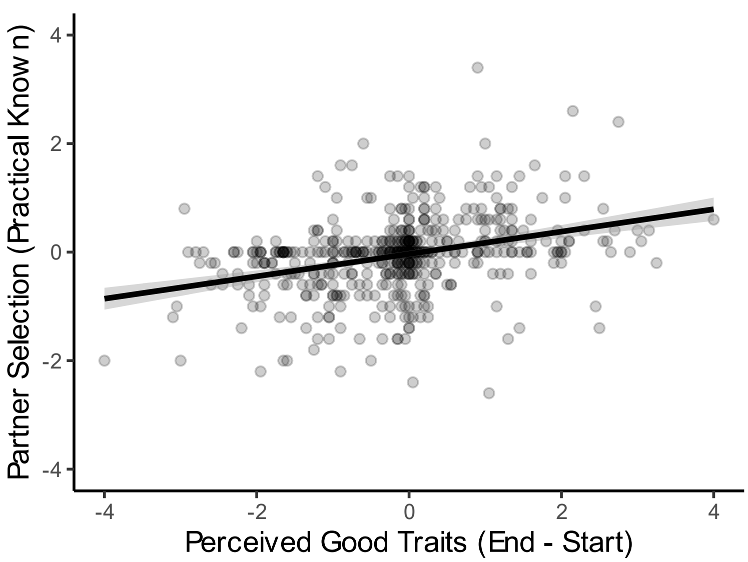
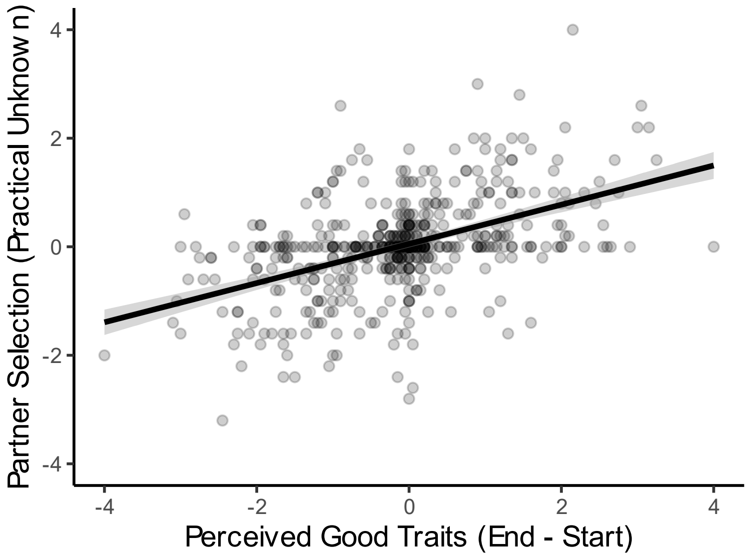


Figure 38 Perceived Good Traits and Partner Selection With Unknown Practical Behavior



Perceived Norm Enforcement of Others

Participants are also asked about to which extent they think *others* want to work with a colleague who meets at the start vs. end of the day. This index can be understood as the perceived norm enforcement from others. Similarly, two tasks are used. In the first task, the colleague's cooperative behavior is known to be mediocre. In the second task, their cooperative behavior is unknown. Five items are used to measure the intention to work with the colleague. Again, factor analyses show that the five items fall onto a single factor, so the items are averaged (see Supplementary Section **4.4.1 Factor Analysis Across the Five Perceived Norm Enforcement Items**).

I hypothesize that perceived enforcement of the symbolic norm from others increases norm adoption (H5). This hypothesis is supported by **Figure 39** and **Figure 40**. The more that participants think others prefer to work with colleagues who meet at the end of the day, the more likely they will meet at the end of the day themselves. The correlation is significant both when the colleague's practical behavior is known ($r(542) = 0.22$, $p < 0.001$) and when the practical behavior is unknown ($r(541) = 0.38$, $p < 0.001$). Notably, the correlation is stronger when the practical behavior is unknown ($\Delta r = 0.16$, $z = 3.81$, $p < 0.001$).

Figure 39 Perceived Enforcement and Behavioral Intention With Known Practical Behavior

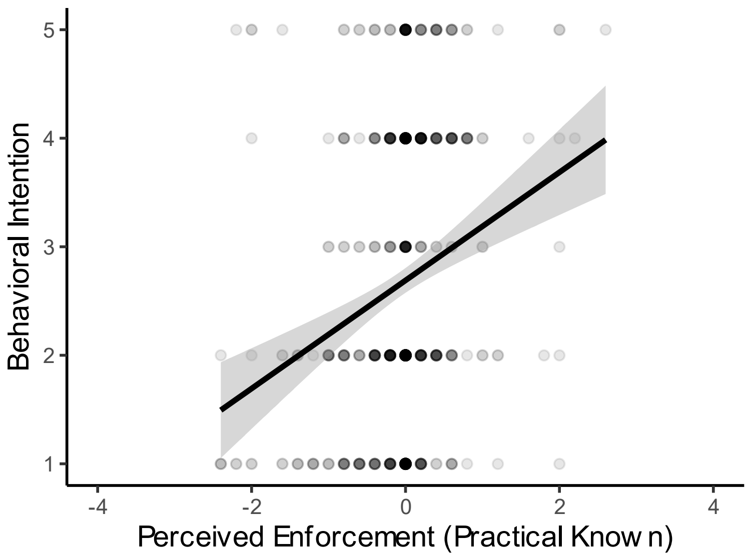
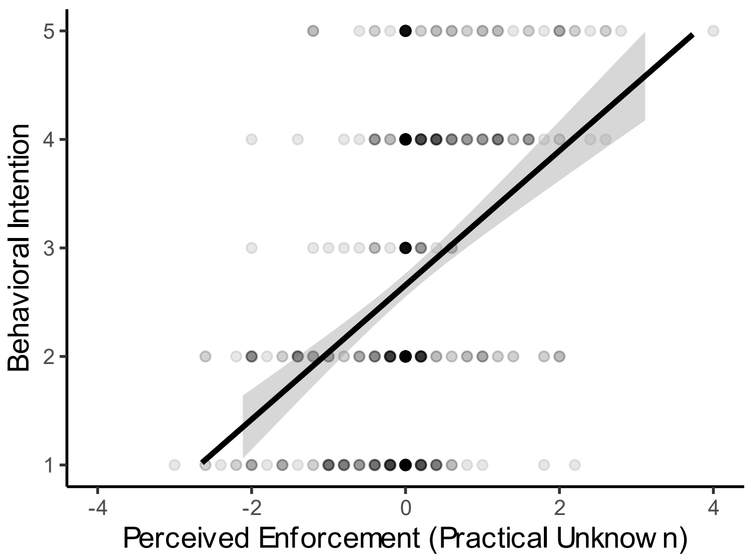


Figure 40 Perceived Enforcement and Behavioral Intention With Unknown Practical Behavior



I also tested a tentative hypothesis that one's own partner selection based on symbolic norm will predict their perceived norm enforcement from others. **Figure 41** and **Figure 42** show that these two variables are positively correlated both when the practical behavior is known ($r(532) = 0.39$, $p < 0.001$) and when the practical behavior is unknown ($r(533) = 0.66$, $p < 0.001$). The correlation is stronger when the practical behavior is unknown ($\Delta r = 0.28$, $z = 7.29$, $p < 0.001$).

It is noteworthy that this correlation might be overestimated because the items used to measure one's own partner selection and their perceived norm enforcement from others are similar. To solve this problem, I did an exploratory analysis in which I used a latent variable pathway model to control for the covariance between similar items. The results show that the path way between one's own partner selection and their perceived norm enforcement from others is strong even after considering the similarity between items (standardized pathway coefficient $b = 0.708$, $p < 0.001$) (see Supplementary Section ***4.4.2 Latent Variable Pathway Analysis Between Partner Selection and Perceived Norm Enforcement*** for more details).

Figure 41 Partner Selection and Perceived Norm Enforcement With Known Practical Behavior

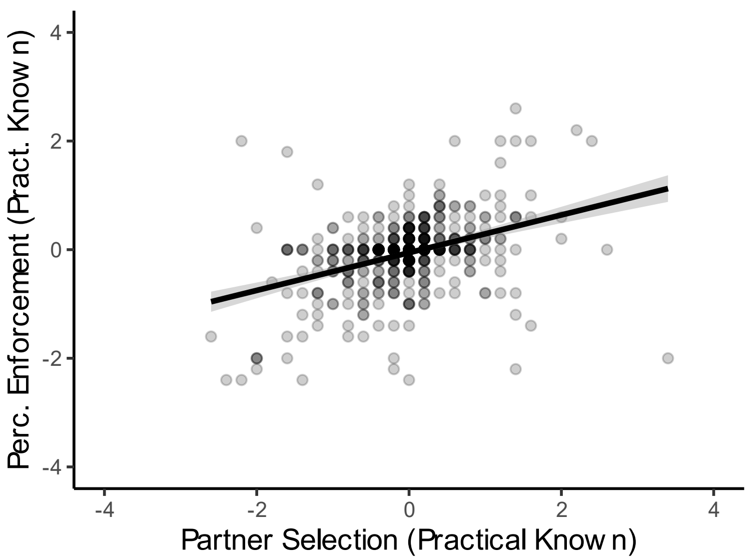
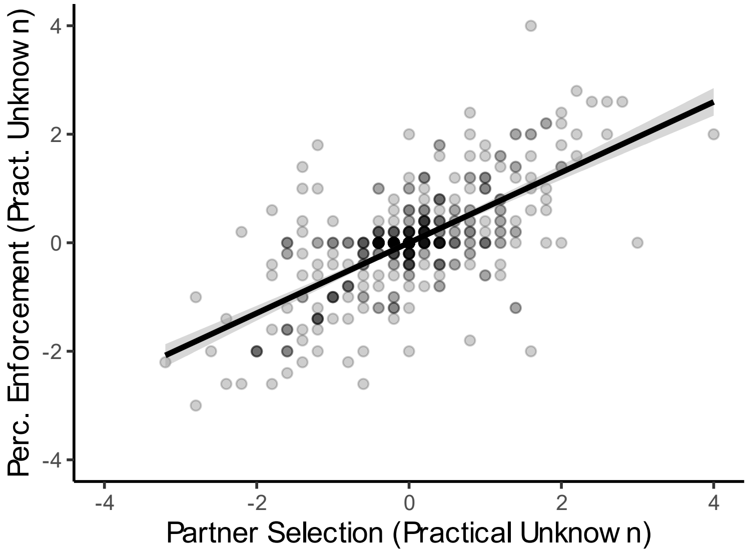


Figure 42 Partner Selection and Perceived Norm Enforcement With Unknown Practical Behavior

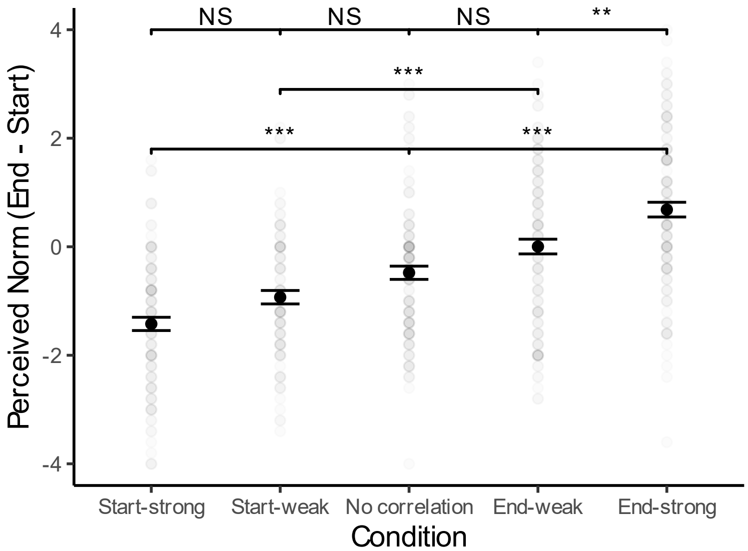


General Perception of the Symbolic Norm

Five items are used to measure the interactive perspective of the symbolic norm, which is "everyone conforms, everyone is expected to conform, and everyone wants to conform when they expect everyone else to conform" (Young, 2015, p. 359). Exploratory factor analysis shows that the five items fall on the same factor, so the average of them is computed as the final measure (see ***4.5.1 Factor Analysis Across the Five Perceived General Norm Items***).

One-way ANOVA shows that the perceived norm is significantly different across the five conditions ($F(4, 542) = 40.22$, $p < 0.001$). When meeting at the end of the day co-occurs more often with cooperative behavior, participants are more likely to perceive that the norm is to meet at the end of the day. Although not all the pairwise comparisons are significant, the pattern is as predicted. Post hoc tests show that the difference between start-strong and start-weak conditions is marginally significant ($p = 0.085$). Although the difference between start-weak and no correlation conditions is not significant ($p = 0.127$), the difference between start-weak and end-weak conditions is significant ($p < 0.001$). The difference between no correlation and end-weak conditions is marginally significant ($p = 0.051$). The difference between end-weak and end-strong conditions is significant ($p = 0.001$).

Figure 43 General Perception of the Symbolic Norm



Notably, Item 1 in this section is "People in this company usually schedule meetings at the end of the day." This is a question on the descriptive norm (Cialdini et al., 1990). Among the 28 profiles, there are always 14 colleagues (i.e., a half the population) meeting at the end of the day. From this sense, the descriptive norm should not vary across conditions. However, an ANOVA on Item 1 shows that participants do perceive a stronger descriptive norm when the symbolic choice is correlated with the practical behavior ($F(4, 545) = 12.08$, $p < 0.001$; see **Figure 44**).

Moreover, in the manipulation check, participants are asked to recall the number of colleagues meeting at the start vs. end of the day. An ANOVA on this recall task shows that when meeting at the end of the day co-occurs more often with cooperative behavior, participants recall more colleagues meeting at the end (as against the start) of the day ($F(4, 543) = 5.69$, $p < 0.001$; see **Figure 45**; see Supplementary Section ***4.5.2 Statistics on Perceived Descriptive Norm*** for statistics and ANOVA tests). This is interesting because the recall task is the first task of the experiment. It simply asks about the number of profiles and does not include any normative language. This result indicates that even when a symbolic choice is not objectively more prevalent in the population, the correlation between this symbolic and a practically beneficial behavior can make the symbolic choice seem more prevalent and thus create an *illusory descriptive norm*. These analyses on perceived descriptive norm were not preregistered and should be considered exploratory.

Figure 44 Perceived Descriptive Norm Across Conditions

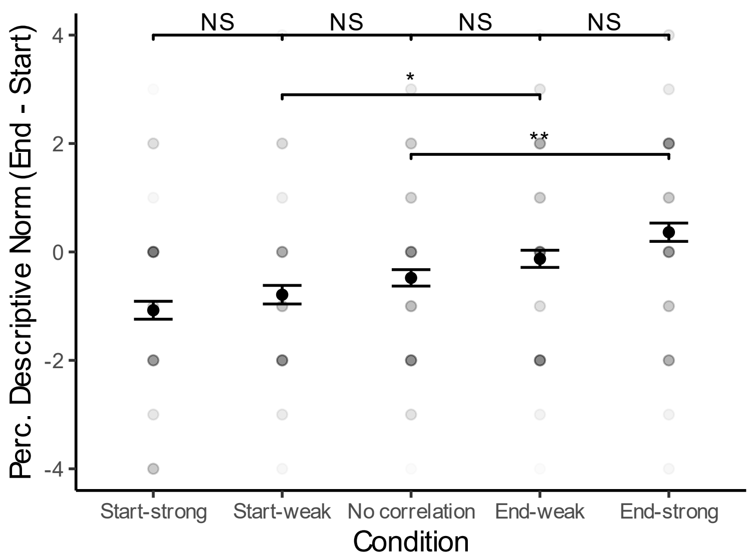
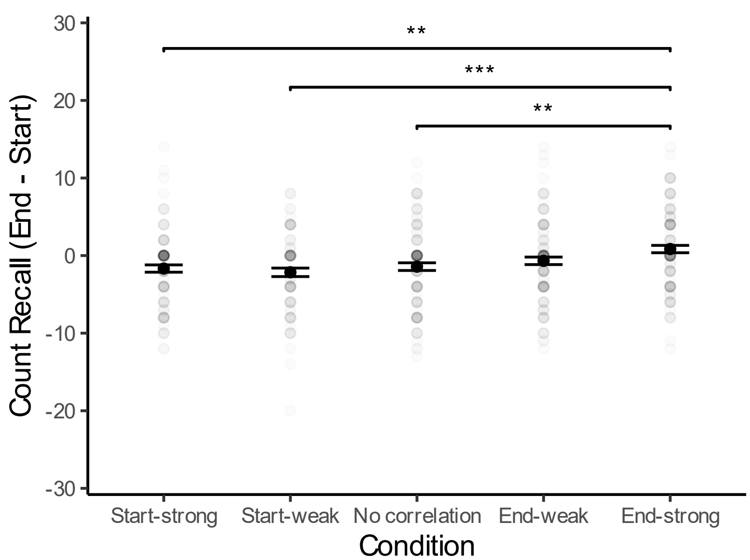
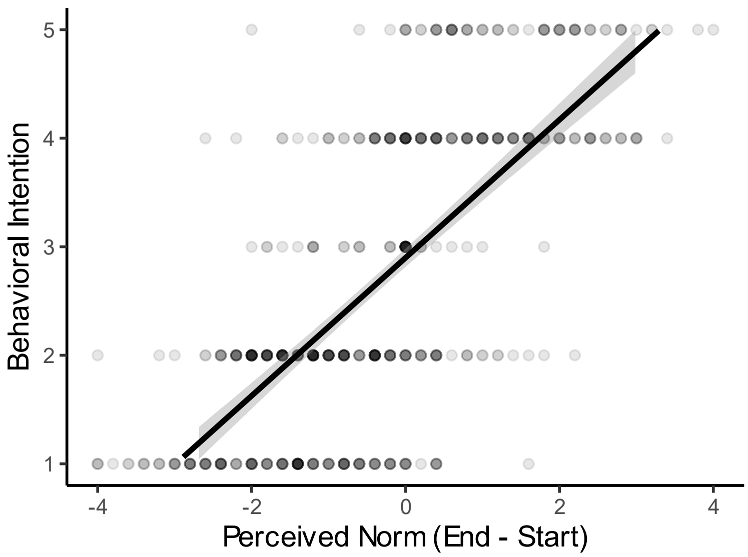


Figure 45 Recalled Number of Colleagues Meeting at the End (as Against Start) of the Day



The perceived symbolic norm also predicted the intention to adopt the symbolic norm ($r(545) = 0.71$, $p < 0.001$; see **Figure 46**).

Figure 46 The Correlation Between Perceived Symbolic Norm and Behavioral Intention



In the dissertation proposal, I also proposed that Items 3-4 could be used as a measure of perceived norm enforcement. A factor analysis was done across Items 3-4 and the five items in perceived norm enforcement. They do not fall on the same factor (see Supplementary Section ***4.5.3 Factor Analysis Among Item 3-4 and Perceived Norm Enforcement Items***). Thus, I did not analyze the data in this way.

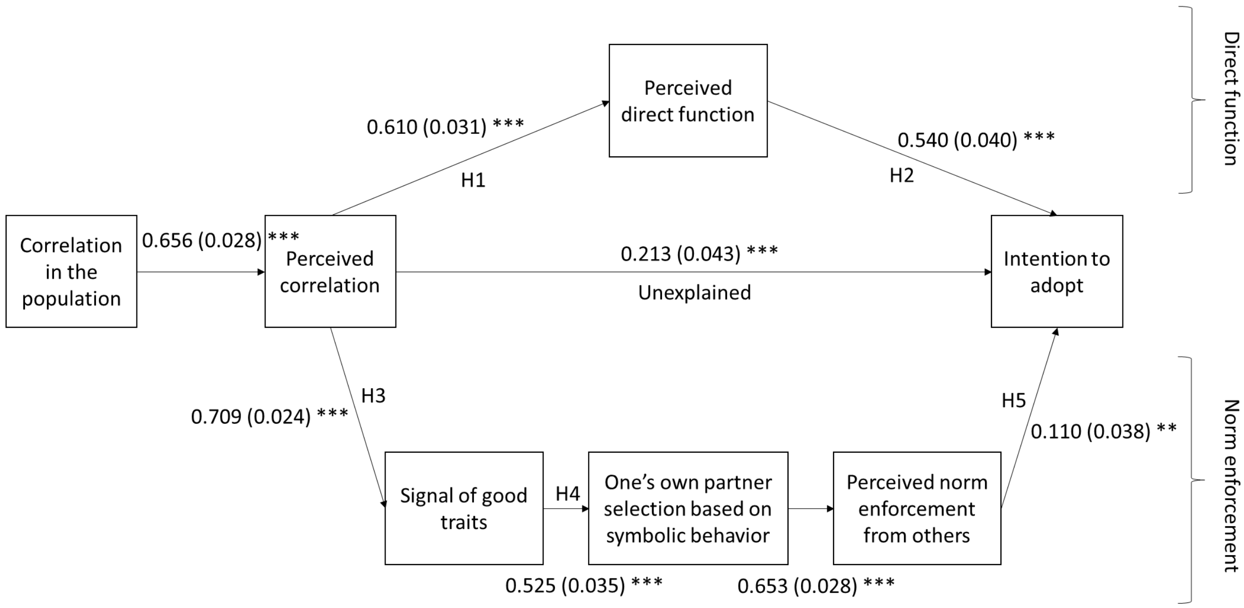
Structural Equation Model

Finally, a measured variable pathway model is used to test these hypotheses together. The goal of the model is to examine the psychological mechanisms that mediate the relationship between perceived correlation and behavioral intention. To summarize, I hypothesized two pathways. The first pathway is that the perceived correlation increases the perceived direct function of the symbolic norm and thus increases its adoption (H1 and H2). The second is that performing the symbolic choice signals good traits in the performer so that people are more likely to select symbolic norm followers as partners. Such enforcement from others motivates people to adopt the symbolic norm themselves (H3, H4, and H5).

Accordingly, the model in **Figure 47** is tested. For partner selection and perceived norm enforcement, the measures from the task with unknown practical behavior are used. Participants with incomplete data are excluded. This results in a sample size of $N = 535$. **Figure 47** shows that all the hypothesized pathways are significant, supporting H1-H5. However, the model fit is only mediocre

(RMSEA = 0.212; CFI = 0.851; SRMR = 0.122). This indicates that some important pathways may be missing in the hypothesized model.

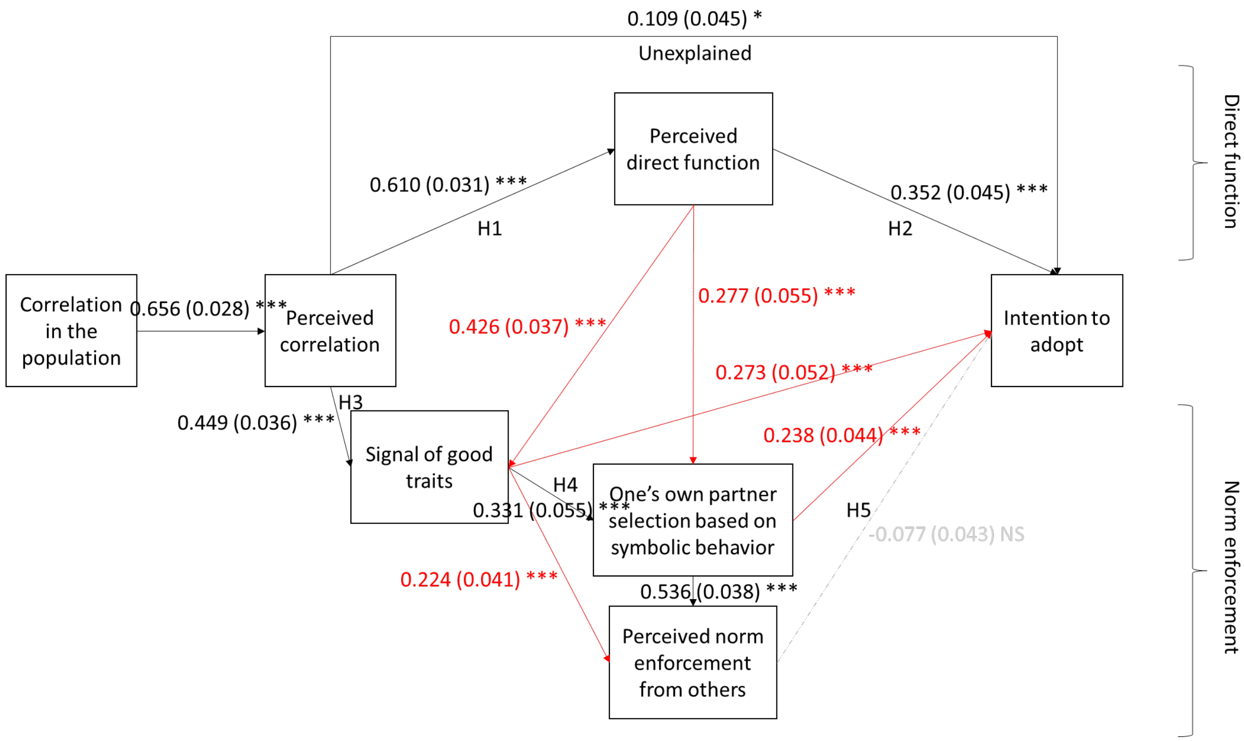
Figure 47 The Model of the Psychological Mechanisms Behind the Norm Spillover Effect



Note. All the coefficients are standardized. The values in brackets are standard errors.

As an exploratory process, a model re-specification is done. Pathways that lead to a significant increase of model fit (i.e., $\Delta\chi^2 > 3.841$) are added if they are theoretically reasonable. The final model is shown in **Figure 48**. The modification is highlighted in red. First, the perceived direct function of a symbolic norm increases the perceived good traits in its performers. This is reasonable because when people believe that meeting at the end of the day boosts productivity, they are likely to believe that colleagues meeting at the end of the day possess good traits, such as competence. Second, the perceived direct function in a symbolic norm predicts partner selection based on that behavior. This is also reasonable because people are likely to select high-performance colleagues as partners. Third, the new model indicates that the relationship between behavioral intention and the perceived good traits is partially mediated by one's own partner selection, but is *not* mediated by perceived partner selection from others. In fact, since one's own partner selection and their perceived other's partner selection are measured in the same task, they are highly correlated ($r(533) = 0.66$, $p < 0.001$). It is unclear whether they should be considered as a single latent construct or two separate constructs in the context of this experiment. It is also unclear whether one of them causes the other. The causal relationship between these two variables should be examined in future research. The fit of the new model is good after adding these pathways (RMSEA = 0.098; CFI = 0.980; SRMR = 0.043).

Figure 48 A Re-specified Model of the Psychological Mechanisms Behind the Norm Spillover Effect



Note. The re-specified model of the psychological mechanisms behind the norm spillover effect. The pathways in red are added beyond the original model. All the coefficients are standardized. The values in brackets are standard errors.

Discussion

Summary of Results

In this chapter, I describe an online experiment that examines the psychological processes that give rise to the norm spillover effect. I show that individuals will adopt a symbolic norm when they perceive a positive correlation between a symbolic choice and a practical norm compliance behavior on the population level. The stronger the correlation they perceive, the more likely they will adopt the symbolic norm. The experiment also identifies two pathways that mediate the relationship between perceived correlation and behavioral intention. One is that the perceived correlation increases the perceived direct function of the symbolic norm (H1) and thus increases its adoption (H2). The other is that performing the symbolic choice signals good traits in the performer (H3) so that people are more likely to select symbolic norm followers as partners (H4). Such enforcement from others motivates people to adopt the symbolic norm themselves (H5). *All* the hypotheses are supported. These results show that there exists a concurrent psychological process that can make the norm spillover effect happen.

Some findings in this experiment are worth further discussion. First, Supplementary Section **4.2.1 Factor Analysis Across the 20 Traits** show that when participants perceive the correlation between a symbolic and a practical behavior, they perceive the symbolic norm followers to possess good traits in all the four dimensions (i.e., warmth, competence, conformity, and commitment) and do not differentiate between these dimensions. Such a result may indicate a halo effect (Nisbett & Wilson, 1977) in which people give a global positive evaluation of symbolic norm followers. Future research can examine the halo effect in norm spillover more carefully.

Second, in the pathway of signaling, I propose two steps. The first is that the perceived signaling efficacy makes people select symbolic norm followers as their own partners. The second is that the perceived norm enforcement from others makes people adopt the norm themselves. There is a gap between one's own partner selection and their perceived partner selection of others. Results show that these two variables are highly correlated ($r(533) = 0.66$, $p < 0.001$ when the practical behavior is unknown). This may indicate that people are projecting their own partner selection strategy to others. But it is also possible that these two variables do not have a causal relationship. In fact, a participant's own partner selection also predicts their behavioral intention ($r(538) = 0.45$, $p < 0.001$ when the practical behavior is known; $r(541) = 0.52$, $p < 0.001$ when the practical behavior is unknown). It even predicts the behavioral intention more accurately than the perceived partner selection of others ($\Delta r = 0.22$, $z = 5.07$, $p < 0.001$ when the practical behavior is known; $\Delta r = 0.14$, $z = 4.60$, $p < 0.001$ when the practical behavior is unknown). In this sense, it becomes unclear what is the relationship between one's own and perceived others' partner selection and which variable mediates the norm enforcement pathway. In fact, the re-specified SEM model shows that this pathway may be mediated by one's own partner selection, not the perceived norm enforcement (see Section ***Structural Equation Model***). This gap requires further examination in the future.

Moreover, in this experiment, instead of measuring actual behavior, the DV is behavioral intention. There can be discrepancies between behavioral intention and actual behavior (Ajzen, 1985; Madden et al., 1992). Additionally, behavioral intention of isolated individuals can be different from their behavior in an interactive environment. It is possible that in a more interactive setting, when people can get others' feedback, their behavior will be influenced more by perceived norm enforcement. This can be an interesting future direction.

Third, as a part of the exploratory analyses, results show that when a symbolic choice co-occurs more often with a practically beneficial behavior, people are more likely to perceive the symbolic choice as the descriptive norm. They even perceive the symbolic choice to be more prevalent. This indicates that the norm spillover effect may have changed people's fundamental frequency representation. The illusory descriptive norm can be another mechanism of the norm spillover effect that awaits future research.

Finally, the exploratory model re-specification shows that some other pathways may also mediate of the norm spillover effect. First, the perceived direct function of a symbolic norm may increase the perceived good traits in its performers. Second, the perceived direct function may increase partner selection based on the symbolic norm. Third, the perceived good traits in symbolic norm followers may increase norm adoption directly. These mechanisms should be re-tested in future research with new data.

Limitations and Future Directions

Experiments aim to make causal inferences. However, it is noteworthy that while some hypotheses in this dissertation are tested causally, some others are not. The experiment shows that the correlation causes the adoption of the symbolic norm, causes the perceived direct function of the symbolic norm, and causes the perceived good traits in symbolic norm followers. However, the experiment is unable to show whether the perceived direct function of and partner selection based on the symbolic norm cause its adoption. The mediation effect is tested only through a structural equation model, not a controlled experiment. One way to test the first pathway causally is to directly manipulate the perceived direct function. For example, if participants see the correlation but are also explicitly told that the symbolic choice does not influence work performance, this may decrease the norm spillover effect. One way to test the second pathway is to show participants that their symbolic choice will not be known by others or will not be used as a criterion for partner selection. Thus, participants will not feel the norm enforcement from others. Such manipulation will help the current experiment make better causal inference.

Moreover, in this experiment, I only examine a one-shot behavioral intention on the individual level. It can be interesting to explore how a symbolic norm dynamically emerges through interactions. I assume that each person may have a private value of conducting a symbolic choice but also a normative pressure to adopt the norm. When the norm is not clear, their choice may be driven more by the private value, or they will try out different choices to test others' reactions. However, once they receive feedback, their choice will be influenced more by the normative pressure. In the worst case scenario, pluralistic ignorance may happen, in which people privately dislike a norm but publicly enforce it (Prentice & Miller, 1996). The descriptive norm in the population may also influence people's perceived direct function of the norm. These processes can be tested in an interactive experiment where a group of people make repeated decisions together.

Chapter 4: General Discussion

There are many social norms that people adhere to on a daily basis but are unaware of their origins. Most societies have rules about dress codes, hair styles, ornaments, social etiquettes, and food taboos that are functionally opaque but actively practiced by their members (Chandler & Schwarz, 2009; Köster et al., 2022; Navarrete & Fessler, 2003). Many of these norms are harmless, and have become a part of our cultural heritage, showcasing the vibrant richness of our culture. Some others, however, can be costly or even hazardous, with their practical functions remaining unclear. It is important to understand how these socially important but functionally opaque norms emerge in a society. In this dissertation, I argue that one possible mechanism is the norm spillover effect: When a behavior is statistically correlated with a practically beneficial behavior on the population level, the behavior will be adopted and enforced, and thus emerge as a symbolic norm.

I test the norm spillover effect through two levels of analysis. On the population level, agent-based evolutionary game theoretic models are used to test whether the co-occurrence of a practical and a neutral behavior creates an evolutionary force that drives the emergence of the symbolic norm. On the individual level, an experiment is used to examine whether humans have such a psychology that allows the norm spillover effect to happen.

Two Levels of Analyses

The two levels of analyses show consistent results. Both analyses show that the correlation makes people adopt the symbolic norm. The stronger the correlation, the stronger the symbolic norm. Both analyses also show that people are more likely to enforce the symbolic norm when they perceive the correlation.

The evolutionary game theoretic model shows that when the correlation between a symbolic and a practical behavior exists, it is a theoretically more adaptive behavior if an individual can adopt the symbolic norm. However, to achieve this adaptive behavior, humans need to possess a concurrent psychological process that leads to corresponding decisions. The consistency between the theoretical model and empirical data shows that humans do possess such a psychology that makes the norm spillover effect possible.

From the perspective of evolutionary psychology, an individual's thoughts and feelings serve one goal—to facilitate the behaviors that are adaptive to the environment (Dawkins, 1976; W. D. Hamilton, 1964). In this sense, if an evolutionary model predicts the adaptiveness of a behavior, people are likely to process the thoughts and feelings that facilitate this behavior. Put it differently, if humans have the propensity for a certain behavior, it is also likely that such behavior increases individuals' adaptiveness under certain circumstances. One of the goals of this dissertation is to find the link between the behavior and its adaptiveness. The dissertation thus shows that although adopting a symbolic norm seems to be a redundancy, it is in fact an adaptive behavior for surviving in a group.

However, consistency between the model and the empirical data does not mean the isomorphism of constructs across these two levels (Kozlowski & Klein, 2000). In the model, the selection happens passively. Individuals have limited cognitive ability to reason whether they should adopt the symbolic norm. Instead, any single individual only takes a fixed strategy. It is the environment that determines which strategies can survive by selecting agents (Hofbauer & Sigmund, 2003). Given that an individual only takes a single strategy, the adoption of a symbolic norm is reflected on the population level by the proportion of agents taking that strategy. In other words, the adoption of a behavior means that the strategy becomes more prevalent than the alternative options in the population.

In contrast, in the empirical experiment, the adoption of a symbolic norm is the result of an individual's reasoning. I assume that when perceiving the correlation between a neutral and a practical behavior, individuals engage in two reasoning processes. First is to figure out what the correlation means and second is to determine which symbolic choice they should adopt. Accordingly, perceiving direct functions in the symbolic norm is the result of an individual's making sense of the correlation (Deutsch & Gerard, 1955; Legare & Souza, 2012; Lin et al., 2022; Matute et al., 2015). On the other hand, evaluating the extent to which others will enforce the norm is one of the processes to determine which symbolic choice to make. It is noteworthy that the two processes in this dissertation are not necessarily the exclusive processes. For example, if individuals see reasonable alternative explanations for the correlation—for example, if people's symbolic choices are decided by an external force—the perceived direction function will reduce. On the other hand, if there are other forces that make an individual has to or does not have to adopt the symbolic norm, such as external punishment or impunity from punishment, these forces may play a more important role in norm adoption. To gain a more comprehensive view of the psychology behind norm spillover, some qualitative methods, such as interviews, may be used as a future direction.

All in all, the model and the experiment describe two different processes. The convergence between them only indicates that the norm spillover effect should happen and can happen, but the model should not be considered as a virtual counterpart of the empirical experiment.

Nevertheless, the model can generate many testable predictions for empirical research. The robustness tests show that the symbolic norm should be stronger when the symbolic strategy is more visible, when norm enforcement is more effective, when ecological threat is higher, and when network mobility is low. These hypotheses can be tested either through a controlled laboratory experiment or by examining relevant archival data. For example, the model predicts that when threat is higher, everyone's resource is reduced. Based on the principle of diminishing marginal utility (Diener & Biswas-Diener, 2002; Foster, 2004), when the baseline resource is low, every gain and cost from social interactions becomes more influential. In this case, finding a cooperative partner becomes even more important. As the symbolic strategy can be used as a signal for cooperative behavior, more individuals should adopt and enforce the symbolic norm under higher threat (see Supplementary Section ***2.3 Ecological Threat***). For another example, the model also predicted that as network mobility increases, the symbolic norm becomes weaker because it is harder to maintain norms under high social network mobility (De et al., 2015) (see Supplementary Section ***2.8 Network Mobility***). Future research can examine whether these predictions are supported by empirical data.

Connections With and Differences From Other Theories

The findings in the dissertation show both consistency with and distinctiveness from previous literature.

First, consistent with theories on costly signaling, the empirical experiment shows that people use others' symbolic choices as a signal (Murray & Moore, 2009; Salahshour, 2019; Sosis, 2009). In fact, one's symbolic choice is not only used as a signal for a specific practical behavior, but also for other good traits in general, from warmth, conformity, commitment, to competence. Nevertheless, the difference between the norm spillover effect and costly signaling is that the symbolic choice does not require a cost. Admittedly, a signal without a cost is not reliable. Thus, in the norm spillover model, the symbolic norm eventually fades after practical norm followers and violators both adopt the symbolic choice.

Second, consistent with the literature on ethnic marker (McElreath et al., 2003) and tag-based cooperation (Riolo et al., 2001), the model shows that when the correlation is weak, agents prefer to interact with others with the same symbolic choices as themselves. Under this circumstance, tag-based cooperation drives the results. However, as the correlation becomes stronger, even symbolic norm violators prefer to interact with symbolic norm followers. Under this circumstance, the norm spillover effect overrides tag-based cooperation and generates opposite results from the theories on ethnic markers. Moreover, the norm spillover model is also consistent with tag-based cooperation on another aspect. The norm spillover model also shows that, although the symbolic norm itself does not generate direct benefit, allowing agents to enforce a symbolic norm benefits cooperation on the population level. Such consistency indicates that one's symbolic choice can still serve as an ethnic marker and facilitate tag-based cooperation. It just has some other functions beyond a simple ethnic marker.

Third, consistent with the model of silly rules, the norm spillover effect also shows that enforcing the symbolic norm benefits cooperation in the population. The difference is that the norm spillover effect does not need a centralized classification scheme that labels behaviors as right vs. wrong. In the norm spillover effect, individuals spontaneously decide which norm to enforce and the symbolic norm can emerge through a bottom-up process.

Finally, while taking a different perspective from cultural transmission theories, the norm spillover effect provokes interesting discussions about these theories. Research on cultural transmission argues that people have the propensity to imitate more successful individuals (Boyd et al., 2011). Experimental simulations also showed that "copy-successful-individuals" is more adaptive than individual learning (Mesoudi, 2008; Mesoudi & O'Brien, 2008). However, followers of practical norms are not necessarily successful individuals. If following a practical social norm means that one has to restrain their own resource for the common good, under certain circumstances, norm followers can be less successful than violators. In this case, if people copy more successful individuals, they should copy the opposite symbolic norm—the symbolic choice correlated with defection.

In the agent-based simulations, I have intentionally created some ecological environments in which the practical norm fails to evolve. Results show that under these circumstances, the symbolic norm may still emerge. Such a result shows that the norm spillover effect has its own valence. The norm that emerges is the symbolic choice that correlates with the practical behavior that benefits the group. If a symbolic choice is correlated with a harmful behavior, even if that harming behavior becomes the norm, it will not spill over to the symbolic domain. From a cultural transmission perspective, symbolic and practical norms will not always be adopted as a package. People may selectively adopt a symbolic norm, depending on its valence.

Generalizability and Limitations

Although this dissertation originally aims to understand symbolic norms, the norm spillover effect can be generalized to other behaviors. For example, if the most altruistic members in a community happen to wear masks more often, it should also promote mask-wearing in the community, though wearing masks itself is also a beneficial norm. When the symbolic norm is related to a controllable behavior, the norm spillover effect explains why such behavior is adopted. When the symbolic norm is related to an uncontrollable characteristic, the norm spillover effect can explain why such a norm is enforced, such as why people ostracize some marginalized groups.

Notably, a symbolic norm can be enforced both institutionally through top-down processes or socially through bottom-up processes. Oftentimes, the top-down and bottom-up processes are interdependent. For example, as Target employees rallied for wearing jeans at work, the company changed their classic uniform, and allows employees to wear blue jeans at work (*What Is The Target's Dress Code? (Updated)*, 2022). On the other hand, when a behavior is enforced institutionally, people also tend to perceive it as important and enforce it socially (Galbiati et al., 2021). This dissertation considers only the bottom-up norm emergence. In this case, whether to adopt a norm is solely based on individual decisions. In real life, this represents the situations where there is no formal regulation regarding a behavior. However, when top-down processes also influence people's choice, the mechanisms might be different. The interaction between bottom-up and top-down enforcement can be an interesting future direction.

Future Directions

A few future directions have been discussed throughout the paper. In this section, I highlight some of them.

First, in the model, I described two pathways that mediate the norm adoption. One is through perceived direct function and the other is through the pressure from norm enforcement. The first pathway can also be understood as norm internalization, in which people adopt a behavior based on the congruence between the behavior and their own value (Bagozzi & Lee, 2002; Gavrilets & Richerson, 2017). Internalization arises from the information communicated from others (Deutsch & Gerard, 1955). In the case of norm spillover, internalization happens when people infer causal relationship between the symbolic choice and a positive outcome when they see the statistical correlation on the population level. In this case, following the norm is believed to give them direct benefit. The second pathway is through social pressure. Conforming to the norm in this case is instrumental for gaining another indirect benefit, such as avoiding punishment (Molho et al., 2020), gaining approval (Masson & Fritsche, 2019), or maintaining positive self-assessment by identifying with valued groups (Cialdini & Goldstein, 2004).

Although both pathways exist at the same time, future research can investigate the conditions under which one pathway becomes more predominant than the other. Cross-cultural differences may exist in the internalization of symbolic norm (Bagozzi & Lee, 2002; Bontempo et al., 1990; Lin et al., 2022; Oh, 2013). Moreover, the relevant strength of the two pathways may change across the time course of norm emergence. Even within a same individual, different cues may elicit different saliency of the pathways. Norm adoption and enforcement may also be mediated by different pathways. Future research can examine the distinction and interaction between norm internalization vs. conformity more carefully.

Second, the empirical experiment only considers and shows the situation when following the symbolic norm signals a good trait. However, sometimes violating the symbolic norm can signal some other traits simultaneously. For example, norm violation may be related to higher creativity (Petrou et al., 2020) and signal higher power (Stamkou & van Kleef, 2014). This falls into the situation that I described in Section **Future Directions** in Chapter 2, where one symbolic behavior is correlated to multiple practical behaviors in different directions. In this case, I hypothesize that whether the symbolic norm emerges depends on the importance of the practical behaviors and the strength of the correlations. For example, if following the symbolic norm strongly signals a critical practical behavior, individuals are unlikely to risk violating the symbolic norm (Demetriou, 2015). However, if the signaled practical behavior is relatively minor, individuals may violate the symbolic norm to signal their creativity instead. Future research can examine the situations when violating the symbolic norm signals good traits of the violator.

Third, the symbolic norm does not need to be costly for the norm spillover effect to happen. However, it would be interesting to study the emergence of costly and even hazardous symbolic norms, such as extreme rituals (Sterelny, 2014). More importantly, future research can examine how harmful symbolic norms fade way and how a harmful norm can be replaced by a less harmful one that serves the same function in maintaining group coherence. An agent-based model can be used to provide theoretical suggestions on the interventions, and an experiment can then be designed to test these interventions empirically.

Fourth, people perceive the world through their own subjective lenses. Their prior knowledge, fears, and expectations all influence what they see. Many factors make people perceive a different correlation between events from the reality. Words that have a strong associative connection are reported as correlated in their occurrence when not actually correlated (Chapman, 1967). When presented with a minority population and a majority population in which some rare/frequent features occur at the same rate in both groups, people overestimate the occurrence of the rare feature in the minority group (Costello & Watts, 2019; D. L. Hamilton & Gifford, 1976; Mullen & Johnson, 1990). When presented with sentences in which the members of different occupational groups were described by trait adjectives, the perceived correlation between traits and occupations was more consistent with existing stereotypic beliefs than was the actual correlation (D. L. Hamilton & Rose, 1980). When the stimuli are fear-relevant (e.g., snakes, spiders, weapons), fearful individuals are more likely to overestimate the relationship between fear-relevant stimuli and aversive consequences (Wiemer & Pauli, 2016). Contrarily, a skeptical mind-set about the undesirable potential correlation leads to more accurate report of no correlation (Munro & Stansbury, 2009). Illusory correlation has been used to explain the maintenance of stereotypic judgments (D. L. Hamilton & Gifford, 1976; D. L. Hamilton & Rose, 1980; Mullen & Johnson, 1990), superstitious beliefs (Tobacyk, 1991), erroneous cultural beliefs (e.g., vaccines cause autism) (A. L. Watts et al., 2015), and anxiety disorders (Wiemer & Pauli, 2016). Future research can investigate how inaccurate perception of the correlation influences the norm spillover effect. More interestingly, future research can also examine how the norm spillover effect influences illusory correlation and perceived prevalence of events. This dissertation shows that the symbolic choice correlated with the practically beneficial behavior is perceived to be more prevalent. It suggests that the norm spillover effect is interacting with perceived correlation and jointly shape people's behavior.

Moreover, people may inaccurately perceive others' intention of norm enforcement. Students tend to overestimated peer normative support for aggression and underestimated peer normative support for nonviolent problem-solving strategies (Henry et al., 2013). Parents tend to overestimate other parents' approval of alcohol use by their children (LaBrie et al., 2011). On the other hand, people also tend to overestimate others' judgement on themselves (Savitsky et al., 2001). Either way, people's perception of others' norm enforcement may be inaccurate, and such misperception shapes their behavior and even their own norm enforcement behavior. Consequently, these behaviors create a new perceived reality in others' minds. Such interaction between perception and behavior can be an exciting future direction.

Finally, it would be interesting to conduct a repeated-decision interactive study to examine the dynamics in the adoption, enforcement, maintenance, and change of symbolic norms. The feedback from each other will influence individuals' perceptions and behaviors across the phases of the norm spillover process. When individuals adjust their behaviors, the correlation between the symbolic and practical behaviors also changes and influences the perceived correlation. At the same time, it would be interesting to examine how people internalize and de-internalize the norm as the correlation fades away. According to the prediction of the agent-based model, we may observe a symbolic norm that first emerges and then disappears. Such an empirical study will be an important complement to the agent-based model and reveal the interactive dynamics in the norm spillover effect.

Bibliography

Ajzen, I. (1985). *From intentions to actions: A theory of planned behavior*. Springer.

Allen, B., Sample, C., Jencks, R., Withers, J., Steinhagen, P., Brizuela, L., Kolodny, J., Parke, D., Lippner, G., & Dementieva, Y. A. (2020). Transient amplifiers of selection and reducers of fixation for death-Birth updating on graphs. *PLOS Computational Biology*, *16*(1), e1007529. https://doi.org/10.1371/journal.pcbi.1007529

Allen, B., Traulsen, A., Tarnita, C. E., & Nowak, M. A. (2012). How mutation affects evolutionary games on graphs. *Journal of Theoretical Biology*, *299*, 97–105. https://doi.org/10.1016/j.jtbi.2011.03.034

Axelrod, R. (1986). An Evolutionary Approach to Norms. *The American Political Science Review*, *80*(4), 1095–1111. JSTOR. https://doi.org/10.2307/1960858

Bagozzi, R. P., & Lee, K.-H. (2002). Multiple Routes for Social Influence: The Role of Compliance, Internalization, and Social Identity. *Social Psychology Quarterly*, *65*(3), 226–247. https://doi.org/10.2307/3090121

Bainbridge, W. A., Isola, P., & Oliva, A. (2013). The intrinsic memorability of face photographs. *Journal of Experimental Psychology: General*, *142*(4), 1323–1334. https://doi.org/10.1037/a0033872

Bem, D. J. (1967). Self-perception: An alternative interpretation of cognitive dissonance phenomena. *Psychological Review*, *74*(3), 183–200. https://doi.org/10.1037/h0024835

Bem, D. J. (1972). Self-perception theory. In *Advances in experimental social psychology* (Vol. 6, pp. 1–62). Elsevier.

Bontempo, R., Lobel, S., & Triandis, H. (1990). Compliance and value internalization in Brazil and the US: Effects of allocentrism and anonymity. *Journal of Cross-Cultural Psychology*, *21*(2), 200–213.

Boyd, R., Richerson, P. J., & Henrich, J. (2011). The cultural niche: Why social learning is essential for human adaptation. *Proceedings of the National Academy of Sciences*, *108*(supplement_2), 10918–10925. https://doi.org/10.1073/pnas.1100290108

Bureau, U. C. (n.d.-a). *2020 Census Illuminates Racial and Ethnic Composition of the Country*. Census.Gov. Retrieved March 13, 2023, from https://www.census.gov/library/stories/2021/08/improved-race-ethnicity-measures-reveal-united-states-population-much-more-multiracial.html

Bureau, U. C. (n.d.-b). *Frequently Occurring Surnames from the 2010 Census*. Census.Gov. Retrieved April 20, 2023, from https://www.census.gov/topics/population/genealogy/data/2010_surnames.html

Cavalli-Sforza, L. L., & Feldman, M. W. (1981). *Cultural Transmission and Evolution: A Quantitative Approach*. Princeton University Press.

Chandler, J., & Schwarz, N. (2009). How extending your middle finger affects your perception of others: Learned movements influence concept accessibility. *Journal of Experimental Social Psychology*, *45*(1), 123–128. https://doi.org/10.1016/j.jesp.2008.06.012

Chapman, L. J. (1967). Illusory correlation in observational report. *Journal of Verbal Learning and Verbal Behavior*, *6*(1), 151–155.

Cialdini, R. B., & Goldstein, N. J. (2004). Social Influence: Compliance and Conformity. *Annual Review of Psychology*, *55*(1), 591–621. https://doi.org/10.1146/annurev.psych.55.090902.142015

Cialdini, R. B., Reno, R. R., & Kallgren, C. A. (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*(6), 1015.

Cialdini, R. B., & Trost, M. R. (1998). *Social influence: Social norms, conformity and compliance.*

Costello, F., & Watts, P. (2019). The rationality of illusory correlation. *Psychological Review*, *126*(3), 437.

Cumberbatch, S. (2021). When Your Identity Is Inherently "Unprofessional": Navigating Rules of Professional Appearance Rooted in Cisheteronormative Whiteness as Black Women and Gender Non-Conforming Professionals. *Journal of Civil Rights and Economic Development*, *34*(2), 81–124.

David, E. J. R., & Derthick, A. O. (2023). *What Is Internalized Oppression, and So What?* Springer Publishing Company. https://connect.springerpub.com/content/book/978-0-8261-9926-3/part/part01/chapter/ch01

Dawkins, R. (1976). *The Selfish Gene*. Oxford University Press. https://doi.org/10.4324/9781912281251

De, S., Gelfand, M. J., Nau, D., & Roos, P. (2015). The inevitability of ethnocentrism revisited: Ethnocentrism diminishes as mobility increases. *Scientific Reports*, *5*, 17963. https://doi.org/10.1038/srep17963

De, S., Nau, D. S., Pan, X., & Gelfand, M. J. (2018). Tipping points for norm change in human cultures. *International Conference on Social Computing, Behavioral-Cultural Modeling and Prediction and Behavior Representation in Modeling and Simulation*, 61–69.

Derex, M., Bonnefon, J.-F., Boyd, R., & Mesoudi, A. (2019). Causal understanding is not necessary for the improvement of culturally evolving technology. *Nature Human Behaviour*, *3*(5), 446–452.

Deutsch, M., & Gerard, H. B. (1955). A study of normative and informational social influences upon individual judgment. *The Journal of Abnormal and Social Psychology*, *51*(3), 629.

Diener, E., & Biswas-Diener, R. (2002). Will money increase subjective well-being? *Social Indicators Research*, *57*(2), 119–169. https://doi.org/10.1023/A:1014411319119

Eriksson, K., Strimling, P., Gelfand, M., Wu, J., Abernathy, J., Akotia, C. S., Aldashev, A., Andersson, P. A., Andrighetto, G., Anum, A., Arikan, G., Aycan, Z., Bagherian, F., Barrera, D., Basnight-Brown, D., Batkeyev, B., Belaus, A., Berezina, E., Björnstjerna, M., … Van Lange, P. A. M. (2021). Perceptions of the appropriate response to norm violation in 57 societies. *Nature Communications*, *12*(1), Article 1. https://doi.org/10.1038/s41467-021-21602-9

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*(4), 1149–1160. https://doi.org/10.3758/BRM.41.4.1149

Fiske, S. T., Cuddy, A. J. C., & Glick, P. (2007). Universal dimensions of social cognition: Warmth and competence. *Trends in Cognitive Sciences*, *11*(2), 77–83. https://doi.org/10.1016/j.tics.2006.11.005

Foster, K. R. (2004). Diminishing returns in social evolution: The not-so-tragic commons. *Journal of Evolutionary Biology*, *17*(5), 1058–1072. https://doi.org/10.1111/j.1420-9101.2004.00747.x

Furnham, A., & Boo, H. C. (2011). A literature review of the anchoring effect. *The Journal of Socio-Economics*, *40*(1), 35–42. https://doi.org/10.1016/j.socec.2010.10.008

Galbiati, R., Henry, E., Jacquemet, N., & Lobeck, M. (2021). How laws affect the perception of norms: Empirical evidence from the lockdown. *PLOS ONE*, *16*(9), e0256624. https://doi.org/10.1371/journal.pone.0256624

García, J., & Traulsen, A. (2012). The Structure of Mutations and the Evolution of Cooperation. *PLOS ONE*, *7*(4), e35287. https://doi.org/10.1371/journal.pone.0035287

García, J., & Traulsen, A. (2019). Evolution of coordinated punishment to enforce cooperation from an unbiased strategy space. *Journal of The Royal Society Interface*, *16*(156), 20190127. https://doi.org/10.1098/rsif.2019.0127

Gavrilets, S., & Richerson, P. J. (2017). Collective action and the evolution of social norm internalization. *Proceedings of the National Academy of Sciences*, *114*(23), 6068–6073. https://doi.org/10.1073/pnas.1703857114

Gelfand, M. J., Harrington, J. R., & Jackson, J. C. (2017). The Strength of Social Norms Across Human Groups. *Perspectives on Psychological Science*, *12*(5), 800–809. https://doi.org/10.1177/1745691617708631

Gelfand, M. J., Raver, J. L., Nishii, L., Leslie, L. M., Lun, J., Lim, B. C., Duan, L., Almaliach, A., Ang, S., Arnadottir, J., Aycan, Z., Boehnke, K., Boski, P., Cabecinhas, R., Chan, D., Chhokar, J., D'Amato, A., Ferrer, M., Fischlmayr, I. C., … Yamaguchi, S. (2011). Differences Between Tight and Loose Cultures: A 33-Nation Study. *Science*, *332*(6033), 1100–1104. https://doi.org/10.1126/science.1197754

Gintis, H., Smith, E. A., & Bowles, S. (2001). Costly Signaling and Cooperation. *Journal of Theoretical Biology*, *213*(1), 103–119. https://doi.org/10.1006/jtbi.2001.2406

Goel, S., Mason, W., & Watts, D. J. (2010). Real and perceived attitude agreement in social networks. *Journal of Personality and Social Psychology*, *99*(4), 611–621. https://doi.org/10.1037/a0020697

Goldsmith, R. E., Clark, R. A., & Lafferty, B. A. (2005). Tendency to Conform: A New Measure and its Relationship to Psychological Reactance. *Psychological Reports*, *96*(3), 591–594. https://doi.org/10.2466/pr0.96.3.591-594

Gosling, S. D., Ko, S. J., Mannarelli, T., & Morris, M. E. (2002). A room with a cue: Personality judgments based on offices and bedrooms. *Journal of Personality and Social Psychology*, *82*(3), 379.

Gross, J., & Vostroknutov, A. (2022). Why do people follow social norms? *Current Opinion in Psychology*, *44*, 1–6. https://doi.org/10.1016/j.copsyc.2021.08.016

Hagberg, A. A., Schult, D. A., & Swart, P. J. (2008). *Exploring Network Structure, Dynamics, and Function using NetworkX*. 5.

Hamilton, D. L., & Gifford, R. K. (1976). Illusory correlation in interpersonal perception: A cognitive basis of stereotypic judgments. *Journal of Experimental Social Psychology*, *12*(4), 392–407. https://doi.org/10.1016/S0022-1031(76)80006-6

Hamilton, D. L., & Rose, T. L. (1980). Illusory correlation and the maintenance of stereotypic beliefs. *Journal of Personality and Social Psychology*, *39*(5), 832–845. https://doi.org/10.1037/0022-3514.39.5.832

Hamilton, W. D. (1964). The genetical evolution of social behaviour. II. *Journal of Theoretical Biology*, *7*(1), 17–52. https://doi.org/10.1016/0022-5193(64)90039-6

Hammond, R. A., & Axelrod, R. (2006). The evolution of ethnocentrism. *Journal of Conflict Resolution*, *50*(6), 926–936. https://doi.org/10.1177/0022002706293470

Haslam, S. A., McGarty, C., & Brown, P. M. (1996). The Search for Differentiated Meaning is a Precursor to Illusory Correlation. *Personality and Social Psychology Bulletin*, *22*(6), 611–619. https://doi.org/10.1177/0146167296226006

Henry, D. B., Dymnicki, A. B., Schoeny, M. E., Meyer, A. L., Martin, N. C., & Project, M. V. P. (2013). Middle school students overestimate normative support for aggression and underestimate normative support for nonviolent problem-solving strategies. *Journal of Applied Social Psychology*, *43*(2), 433–445. https://doi.org/10.1111/j.1559-1816.2013.01027.x

Hindersin, L., & Traulsen, A. (2015). Most Undirected Random Graphs Are Amplifiers of Selection for Birth-Death Dynamics, but Suppressors of Selection for Death-Birth Dynamics. *PLOS Computational Biology*, *11*(11), e1004437. https://doi.org/10.1371/journal.pcbi.1004437

Hobson, N. M., Schroeder, J., Risen, J. L., Xygalatas, D., & Inzlicht, M. (2018). The Psychology of Rituals: An Integrative Review and Process-Based Framework. *Personality and Social Psychology Review*, *22*(3), 260–284. https://doi.org/10.1177/1088868317734944

Hoehl, S., Keupp, S., Schleihauf, H., McGuigan, N., Buttelmann, D., & Whiten, A. (2019). ‘Over-imitation’: A review and appraisal of a decade of research. *Developmental Review*, *51*, 90–108. https://doi.org/10.1016/j.dr.2018.12.002

Hofbauer, J., & Sigmund, K. (2003). Evolutionary game dynamics. *Bulletin of the American Mathematical Society*, *40*(4), 479–519.

Jackson, D. L. (2003). Revisiting Sample Size and Number of Parameter Estimates: Some Support for the N:q Hypothesis. *Structural Equation Modeling: A Multidisciplinary Journal*, *10*(1), 128–141. https://doi.org/10.1207/S15328007SEM1001_6

Jackson, J., & Gelfand, M. (2016). On the nature of tightness-looseness in cultural ecosystems. *Religion, Brain, and Behavior*, *7*. https://doi.org/10.1080/2153599X.2015.1132252

Kameda, T., Takezawa, M., & Hastie, R. (2003). The Logic of Social Sharing: An Evolutionary Game Analysis of Adaptive Norm Development. *Personality and Social Psychology Review*, *7*(1), 2–19. https://doi.org/10.1207/S15327957PSPR0701_1

Kenward, B. (2012). Over-imitating preschoolers believe unnecessary actions are normative and enforce their performance by a third party. *Journal of Experimental Child Psychology*, *112*(2), 195–207. https://doi.org/10.1016/j.jecp.2012.02.006

Kline, R. B. (2015). *Principles and practice of structural equation modeling*. Guilford publications.

Köster, R., Hadfield-Menell, D., Everett, R., Weidinger, L., Hadfield, G. K., & Leibo, J. Z. (2022). Spurious normativity enhances learning of compliance and enforcement behavior in artificial agents. *Proceedings of the National Academy of Sciences*, *119*(3), e2106028118. https://doi.org/10.1073/pnas.2106028118

Köster, R., Hadfield-Menell, D., Hadfield, G. K., & Leibo, J. Z. (2020). Silly rules improve the capacity of agents to learn stable enforcement and compliance behaviors. *arXiv:2001.09318 [Cs]*. http://arxiv.org/abs/2001.09318

Kozlowski, S. W., & Klein, K. J. (2000). *A multilevel approach to theory and research in organizations: Contextual, temporal, and emergent processes.*

LaBrie, J. W., Hummer, J. F., Lac, A., Ehret, P. J., & Kenney, S. R. (2011). Parents Know Best, But Are They Accurate? Parental Normative Misperceptions and Their Relationship to Students' Alcohol-Related Outcomes. *Journal of Studies on Alcohol and Drugs*, *72*(4), 521–529.

Legare, C. H., & Souza, A. L. (2012). Evaluating ritual efficacy: Evidence from the supernatural. *Cognition*, *124*(1), 1–15. https://doi.org/10.1016/j.cognition.2012.03.004

Legare, C. H., Watson-Jones, R. E., & Souza, A. L. (2016). A cognitive psychological account of reasoning about ritual efficacy. *Advances in Religion, Cognitive Science, and Experimental Philosophy*, 85–102.

Lin, Y., Zhang, Y. C., & Oyserman, D. (2022). Seeing meaning even when none may exist: Collectivism increases belief in empty claims. *Journal of Personality and Social Psychology*, *122*(3), 351–366. https://doi.org/10.1037/pspa0000280

Liu, X., He, M., Kang, Y., & Pan, Q. (2017). Fixation of strategies with the Moran and Fermi processes in evolutionary games. *Physica A: Statistical Mechanics and Its Applications*, *484*, 336–344. https://doi.org/10.1016/j.physa.2017.04.154

Macy, M. W., & Willer, R. (2002). From factors to actors: Computational sociology and agent-based modeling. *Annual Review of Sociology*, 143–166.

Madden, T. J., Ellen, P. S., & Ajzen, I. (1992). A Comparison of the Theory of Planned Behavior and the Theory of Reasoned Action. *Personality and Social Psychology Bulletin*, *18*(1), 3–9. https://doi.org/10.1177/0146167292181001

Majority of Japan hotels ban tattooed tourists from public baths. (n.d.). Retrieved February 26, 2023, from https://www.telegraph.co.uk/news/worldnews/asia/japan/11947313/Majority-of-Japan-hotels-ban-tattooed-tourists-from-public-baths.html

Masson, T., & Fritsche, I. (2019). Loyal peripherals? The interactive effects of identification and peripheral group membership on deviance from non-beneficial ingroup norms. *European Journal of Social Psychology*, *49*(1), 76–92. https://doi.org/10.1002/ejsp.2501

Matute, H., Blanco, F., Yarritu, I., Díaz-Lago, M., Vadillo, M. A., & Barberia, I. (2015). Illusions of causality: How they bias our everyday thinking and how they could be reduced. *Frontiers in Psychology*, *6*. https://www.frontiersin.org/articles/10.3389/fpsyg.2015.00888

Matute, H., Yarritu, I., & Vadillo, M. A. (2011). Illusions of causality at the heart of pseudoscience. *British Journal of Psychology*, *102*(3), 392–405. https://doi.org/10.1348/000712610X532210

McElreath, R., Boyd, R., & Richerson, P. J. (2003). Shared Norms and the Evolution of Ethnic Markers. *Current Anthropology*, *44*(1), 122–130. https://doi.org/10.1086/345689

Mesoudi, A. (2008). An experimental simulation of the “copy-successful-individuals” cultural learning strategy: Adaptive landscapes, producer–scrounger dynamics, and informational access costs. *Evolution and Human Behavior*, *29*(5), 350–363. https://doi.org/10.1016/j.evolhumbehav.2008.04.005

Mesoudi, A. (2011). An experimental comparison of human social learning strategies: Payoff-biased social learning is adaptive but underused. *Evolution and Human Behavior*, *32*(5), 334–342. https://doi.org/10.1016/j.evolhumbehav.2010.12.001

Mesoudi, A., & O’Brien, M. J. (2008). The Cultural Transmission of Great Basin Projectile-Point Technology I: An Experimental Simulation. *American Antiquity*, *73*(1), 3–28.

Milgram, S. (1967). The small world problem. *Psychology Today*, *2*(1), 60–67.

Molho, C., Tybur, J. M., Van Lange, P. A. M., & Balliet, D. (2020). Direct and indirect punishment of norm violations in daily life. *Nature Communications*, *11*(1), Article 1. https://doi.org/10.1038/s41467-020-17286-2

Mowday, R. T., Steers, R. M., & Porter, L. W. (1979). The measurement of organizational commitment. *Journal of Vocational Behavior*, *14*(2), 224–247.

Mullen, B., & Johnson, C. (1990). Distinctiveness-based illusory correlations and stereotyping: A meta-analytic integration*. *British Journal of Social Psychology*, *29*(1), 11–28. https://doi.org/10.1111/j.2044-8309.1990.tb00883.x

Munro, G. D., & Stansbury, J. A. (2009). The Dark Side of Self-Affirmation: Confirmation Bias and Illusory Correlation in Response to Threatening Information. *Personality and Social Psychology Bulletin*, *35*(9), 1143–1153. https://doi.org/10.1177/0146167209337163

Murray, M. J., & Moore, L. (2009). Costly Signaling and the Origin of Religion. *Journal of Cognition & Culture*, *9*(3/4), 225–245. https://doi.org/10.1163/156770909X12489459066264

Navarrete, C. D., & Fessler, D. (2003). Meat Is Good to Taboo: Dietary Proscriptions as a Product of the Interaction of Psychological Mechanisms and Social Processes. *Journal of Cognition and Culture*, *3*(1), 1–40. https://doi.org/10.1163/156853703321598563

Nisbett, R. E., & Wilson, T. D. (1977). The halo effect: Evidence for unconscious alteration of judgments. *Journal of Personality and Social Psychology*, *35*(4), 250.

Nowak, M. A. (2006). Five Rules for the Evolution of Cooperation. *Science*, *314*(5805), 1560–1563. https://doi.org/10.1126/science.1133755

np_123_aos. (2022, June 29). What Is The Target's Dress Code? (Updated). *AisleofShame.Com*. https://www.aisleofshame.com/what-is-the-target-dress-code/

Oh, S. H. (2013). Do collectivists conform more than individualists? Cross-cultural differences in compliance and internalization. *Social Behavior and Personality: An International Journal*, *41*(6), 981–994. https://doi.org/10.2224/sbp.2013.41.6.981

Ohtsuki, H., Hauert, C., Lieberman, E., & Nowak, M. A. (2006). A simple rule for the evolution of cooperation on graphs and social networks. *Nature*, *441*(7092), 502.

Omobola, O. C. (2013). An Overview of Taboo and Superstition among the Yoruba of Southwest of Nigeria. *Mediterranean Journal of Social Sciences*. https://doi.org/10.5901/mjss.2013.v4n2p221

Organizational Citizenship Behavior Checklist OCB-C. (n.d.). *Paul Spector*. Retrieved March 13, 2023, from https://paulspector.com/assessments/pauls-no-cost-assessments/organizational-citizenship-behavior-checklist-ocb-c/

Pan, X., Gelfand, M., & Nau, D. (2021). Integrating evolutionary game theory and cross-cultural psychology to understand cultural dynamics. *American Psychologist*, *76*(6), 1054–1066. https://doi.org/10.1037/amp0000851

Pan, X., Landry, A., & Gelfand, M. (2023). *The Evolution of Intergroup Relations*.

Pepitone, A. (1976). Toward a normative and comparative biocultural social psychology. *Journal of Personality and Social Psychology*, *34*(4), 641–653. https://doi.org/10.1037/0022-3514.34.4.641

Petrou, P., van der Linden, D., & Salcescu, O. C. (2020). When Breaking the Rules Relates to Creativity: The Role of Creative Problem-Solving Demands and Organizational Constraints. *The Journal of Creative Behavior*, *54*(1), 184–195. https://doi.org/10.1002/jocb.354

Popp, A. (2006). The effects of natural disasters on long run growth. *Major Themes in Economics*, *8*(1), 61–82.

Popular Baby Names by Decade. (n.d.). Retrieved March 14, 2023, from https://www.ssa.gov/oact/babynames/decades/index.html

Prentice, D. A., & Miller, D. T. (1996). Pluralistic Ignorance and the Perpetuation of Social Norms by Unwitting Actors. In *Advances in Experimental Social Psychology* (Vol. 28, pp. 161–209). Elsevier. https://doi.org/10.1016/S0065-2601(08)60238-5

Pryor, C., Perfors, A., & Howe, P. D. L. (2018). Even arbitrary norms influence moral decision-making. *Nature Human Behaviour*, *3*(1), 57–62. https://doi.org/10.1038/s41562-018-0489-y

Rakoczy, H., Warneken, F., & Tomasello, M. (2008). The sources of normativity: Young children's awareness of the normative structure of games. *In Press, Developmental Psychology*, 875–881.

Ramos-Villagrasa, P. J., Barrada, J. R., Fernández-del-Río, E., & Koopmans, L. (2019). Assessing Job Performance Using Brief Self-report Scales: The Case of the Individual Work Performance Questionnaire. *Revista de Psicología Del Trabajo y de Las Organizaciones*, *35*(3), 195–205. https://doi.org/10.5093/jwop2019a21

Rand, D. G., & Nowak, M. A. (2011). The evolution of antisocial punishment in optional public goods games. *Nature Communications*, *2*(1), Article 1. https://doi.org/10.1038/ncomms1442

Riolo, R. L., Cohen, M. D., & Axelrod, R. (2001). Evolution of cooperation without reciprocity. *Nature*, *414*(6862), 441–443. https://doi.org/10.1038/35106555

Robertson, K., & Tustin, K. (2018). Students Who Limit Their Drinking, as Recommended by National Guidelines, Are Stigmatized, Ostracized, or the Subject of Peer Pressure: Limiting Consumption Is All But Prohibited in a Culture of Intoxication. *Substance Abuse: Research and Treatment*, *12*, 1178221818792414. https://doi.org/10.1177/1178221818792414

Roca, C. P., Cuesta, J. A., & Sánchez, A. (2009). Evolutionary game theory: Temporal and spatial effects beyond replicator dynamics. *Physics of Life Reviews*, *6*(4), 208–249. https://doi.org/10.1016/j.plrev.2009.08.001

Roos, P., Gelfand, M., Nau, D., & Lun, J. (2015). Societal threat and cultural variation in the strength of social norms: An evolutionary basis. *Organizational Behavior and Human Decision Processes*, *129*, 14–23. https://doi.org/10.1016/j.obhdp.2015.01.003

RStudio Team. (2020). *RStudio: Integrated Development Environment for R*. RStudio, PBC. http://www.rstudio.com/

Salahshour, M. (2019). Evolution of costly signaling and partial cooperation. *Scientific Reports*, *9*(1), 8792. https://doi.org/10.1038/s41598-019-45272-2

Savitsky, K., Epley, N., & Gilovich, T. (2001). Do others judge us as harshly as we think? Overestimating the impact of our failures, shortcomings, and mishaps. *Journal of Personality and Social Psychology*, *81*, 44–56. https://doi.org/10.1037/0022-3514.81.1.44

Schmidt, M. F. H., Butler, L. P., Heinz, J., & Tomasello, M. (2016). Young Children See a Single Action and Infer a Social Norm: Promiscuous Normativity in 3-Year-Olds. *Psychological Science*, *27*(10), 1360–1370. https://doi.org/10.1177/0956797616661182

Schwartz, S. H., & Howard, J. A. (1984). Internalized Values as Motivators of Altruism. In E. Staub, D. Bar-Tal, J. Karylowski, & J. Reykowski (Eds.), *Development and Maintenance of Prosocial Behavior: International Perspectives on Positive Morality* (pp. 229–255). Springer US. https://doi.org/10.1007/978-1-4613-2645-8_14

Sigmund, K., & Nowak, M. A. (1999). Evolutionary game theory. *Current Biology*, *9*(14), R503–R505.

Sivakumar, M. V. K. (2005). Impacts of natural disasters in agriculture, rangeland and forestry: An overview. In M. V. K. Sivakumar, R. P. Motha, & H. P. Das (Eds.), *Natural Disasters and Extreme Events in Agriculture: Impacts and Mitigation* (pp. 1–22). Springer. https://doi.org/10.1007/3-540-28307-2_1

Sosis, R. (2009). The adaptationist-byproduct debate on the evolution of religion: Five misunderstandings of the adaptationist program. *Journal of Cognition and Culture*, *9*(3–4), 315–332.

Stamkou, E., & van Kleef, G. (2014). Do we give power to the right people? When and how norm violators rise to the top. In *Power, Politics, and Paranoia: Why People are Suspicious About their Leaders* (pp. 33–52). https://doi.org/10.1017/CBO9781139565417.005

Sterelny, K. (2014). A Paleolithic Reciprocation Crisis: Symbols, Signals, and Norms. *Biological Theory*, *9*(1), 65–77. https://doi.org/10.1007/s13752-013-0143-x

Stutterheim, S. E., Bos, A. E. R., Pryor, J. B., Brands, R., Liebregts, M., & Schaalma, H. P. (2011). Psychological and Social Correlates of HIV Status Disclosure: The Significance of Stigma Visibility. *AIDS Education and Prevention*, *23*(4), 382–392. https://doi.org/10.1521/aeap.2011.23.4.382

Summers, J. K., Howe, M., McElroy, J. C., Ronald Buckley, M., Pahng, P., & Cortes-Mejia, S. (2018). A typology of stigma within organizations: Access and treatment effects. *Journal of Organizational Behavior*, *39*(7), 853–868. https://doi.org/10.1002/job.2279

Tattoos: From Social Taboo to Social Acceptance | Free Essay Example. (2022, January 20). StudyCorgi.Com. https://studycorgi.com/tattoos-from-social-taboo-to-social-acceptance/

Teamwork Perceptions Questionnaire (T-TPQ). (n.d.). Retrieved March 13, 2023, from https://www.ahrq.gov/teamstepps/instructor/reference/teampercept.html

Tetlock, P. E. (1985). Accountability: A Social Check on the Fundamental Attribution Error. *Social Psychology Quarterly*, *48*(3), 227–236. https://doi.org/10.2307/3033683

The world's least tattoo-friendly countries. (2016, April 25). Andrea Catton Laser Clinic. https://www.andreacatton.co.uk/2016/04/worlds-least-tattoo-friendly-countries/

Tkadlec, J., Pavlogiannis, A., Chatterjee, K., & Nowak, M. A. (2020). Limits on amplifiers of natural selection under death-Birth updating. *PLOS Computational Biology*, *16*(1), e1007494. https://doi.org/10.1371/journal.pcbi.1007494

Tobacyk, J. J. (1991). Superstition and Beliefs about the Prediction of Future Events. *Psychological Reports*, *68*(2), 511–512. https://doi.org/10.2466/pr0.1991.68.2.511

Traulsen, A., Pacheco, J. M., & Nowak, M. A. (2007). Pairwise comparison and selection temperature in evolutionary game dynamics. *Journal of Theoretical Biology*, *246*(3), 522–529. https://doi.org/10.1016/j.jtbi.2007.01.002

Wang, E. S.-T., & Chu, Y.-H. (2021). How social norms affect consumer intention to purchase certified functional foods: The mediating role of perceived effectiveness and attitude. *Foods*, *10*(6), 1151.

Watts, A. L., Smith, S. F., & Lilienfeld, S. O. (2015). Illusory Correlation. In R. L. Cautin & S. O. Lilienfeld (Eds.), *The Encyclopedia of Clinical Psychology* (pp. 1–8). John Wiley & Sons, Inc. https://doi.org/10.1002/9781118625392.wbecp573

Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *Nature*, *393*(6684), Article 6684. https://doi.org/10.1038/30918

Weeden, K., & Cornwell, B. (2020). The Small-World Network of College Classes: Implications for Epidemic Spread on a University Campus. *Sociological Science*, *7*, 222–241. https://doi.org/10.15195/v7.a9

Wiemer, J., & Pauli, P. (2016). Fear-relevant illusory correlations in different fears and anxiety disorders: A review of the literature. *Journal of Anxiety Disorders*, *42*, 113–128. https://doi.org/10.1016/j.janxdis.2016.07.003

Young, H. P. (2015). The Evolution of Social Norms. *Annual Review of Economics*, *7*(1), 359–387. https://doi.org/10.1146/annurev-economics-080614-115322

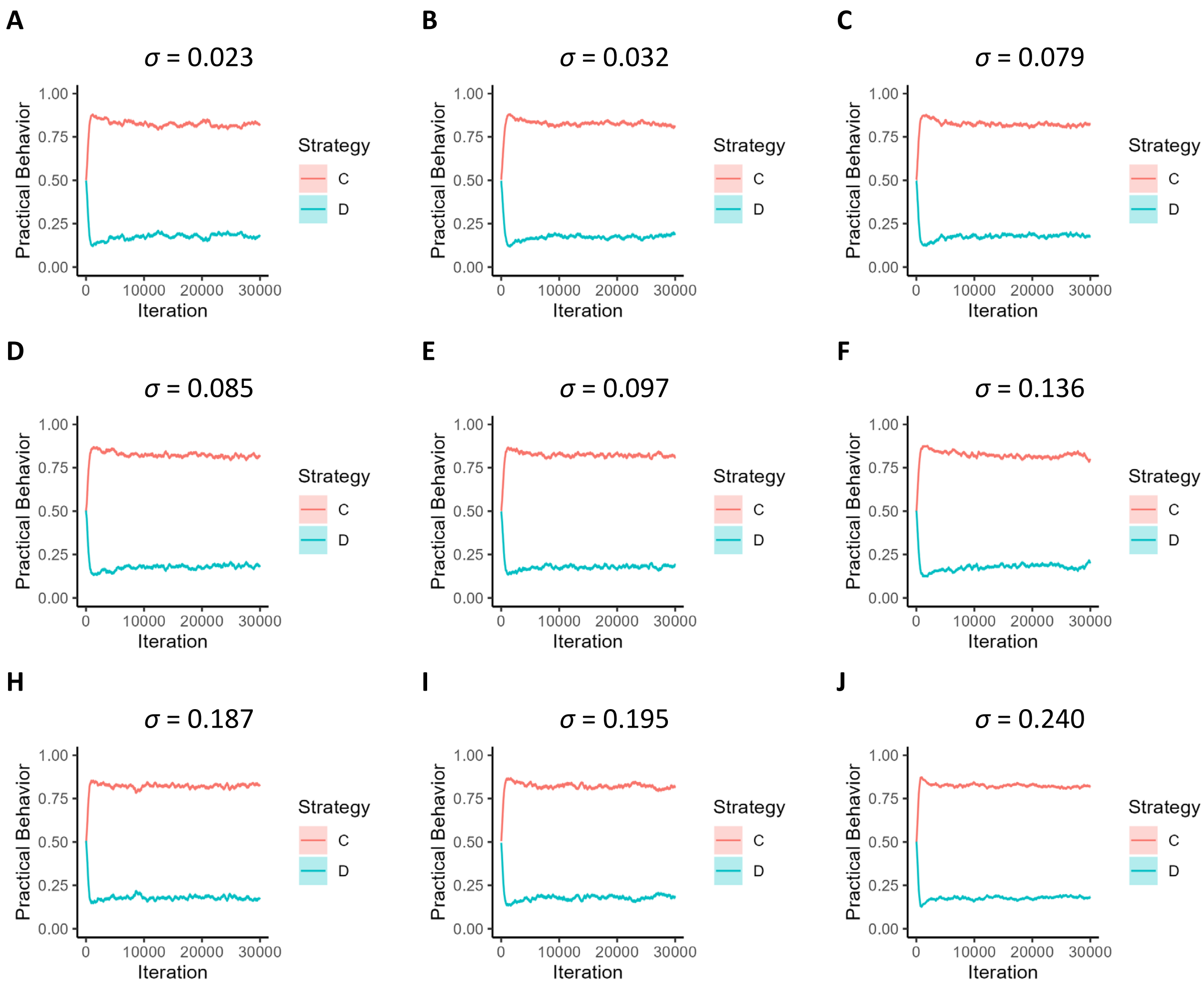
Supplementary Materials

1 Supplementary Results for the Main Simulation

1.1 The Evolution of Cooperation

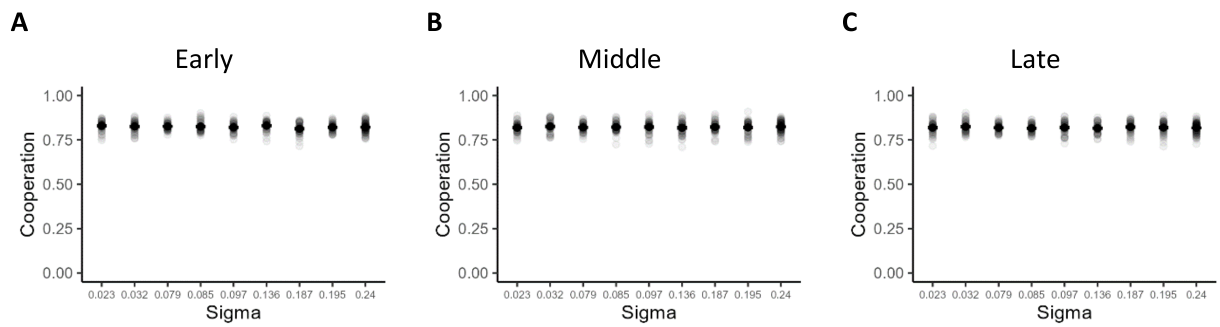
Figure 49 shows the evolutionary trajectory of cooperation in the default model. **Figure 50** shows the cooperation rates in the early, middle, and late phases of the simulation under different levels of σ.

Figure 49 The Evolution of Cooperation



Note. The evolution of cooperation under different levels of *σ*. C represents cooperators and D represents defectors. Plot J is averaged across 100 simulation runs under *σ* = 0.240, because this *σ* value was randomly chosen twice by an R program.

Figure 50 Cooperation Rate Under Different Levels of *σ*



1.2 The Impact of Symbolic Enforcement on Cooperation

1.2.1 The Impact of Symbolic Enforcement When Symbolic Strategies Are Transparent

Three 2 (with vs. without symbolic enforcement) * 6 ($\sigma \in [0, 0.05, 0.1, 0.15, 0.2, 0.25]$) ANOVA were done on the early, middle, and late phases of the simulation to examine the impact of symbolic enforcement on cooperation under different *σ* levels. In these conditions, agents' cooperation strategies are non-transparent and their symbolic strategies are transparent.

For the early phase, the interaction between enforcement and *σ* is not significant ($F(5, 588) = 1.996$, $p = 0.078$). The main effect of enforcement is significant ($F(1, 588) = 453.590$, $p < 0.001$). The cooperation rate is significantly higher with symbolic enforcement ($M = 0.829$, $SE = 0.002$) than without symbolic enforcement ($M = 0.779$, $SE = 0.002$). The main effect of *σ* is not significant ($F(5, 588) = 1.132$, $p = 0.342$).

For the middle phase, the interaction between enforcement and *σ* is not significant ($F(5, 588) = 0.341$, $p = 0.888$). The main effect of enforcement is significant ($F(1, 588) = 158.294$, $p < 0.001$). The cooperation rate is significantly higher with symbolic enforcement ($M = 0.820$, $SE = 0.002$) than without symbolic enforcement ($M = 0.786$, $SE = 0.002$). The main effect of σ is not significant ($F(5, 588) = 1.823$, $p = 0.316$).

For the late phase, the interaction between enforcement and σ is not significant ($F(5, 588) = 0.938$, $p = 0.456$). The main effect of enforcement is significant ($F(1, 588) = 119.027$, $p < 0.001$). The cooperation rate is significantly higher with symbolic enforcement ($M = 0.821$, $SE = 0.002$) than without symbolic enforcement ($M = 0.790$, $SE = 0.002$). The main effect of σ is not significant ($F(5, 588) = 1.175$, $p = 0.320$).

1.2.2 The Impact of Symbolic Enforcement When Symbolic Strategies Are Non-transparent

For the symbolic enforcement to promote cooperation, agents' symbolic strategies must be easily visible. This section compares the conditions with vs. without symbolic enforcement when the symbolic strategies are *non-transparent*. Three 2 (with vs. without symbolic enforcement) * 6 ($\sigma \in [0, 0.05, 0.1, 0.15, 0.2, 0.25]$) ANOVA were done on the early, middle, and late phases of the simulation.

For the early phase, the interaction between enforcement and σ is not significant ($F(5, 588) = 0.956$, $p = 0.444$). The main effect of enforcement is significant but small ($F(1, 588) = 6.846$, $p = 0.009$). The cooperation rate is slightly higher with symbolic enforcement ($M = 0.786$, $SE = 0.002$) than without symbolic enforcement ($M = 0.779$, $SE = 0.002$). The main effect of σ is not significant ($F(5, 588) = 0.928$, $p = 0.462$).

For the middle phase, the interaction between enforcement and σ is not significant ($F(5, 588) = 0.864$, $p = 0.505$). The main effect of enforcement is not significant ($F(1, 588) = 0.941$, $p = 0.333$). The main effect of σ is not significant ($F(5, 588) = 0.431$, $p = 0.827$).

For the late phase, the interaction between enforcement and σ is not significant ($F(5, 588) = 0.892$, $p = 0.486$). The main effect of enforcement is not significant ($F(1, 588) = 0.082$, $p = 0.774$). The main effect of σ is not significant ($F(5, 588) = 1.867$, $p = 0.098$).

1.2.3 Another Model Option to Remove Symbolic Enforcement

I also tried another model option to examine the impact of symbolic enforcement on cooperation. In this set of simulations, agents are prohibited from ostracizing others based on symbolic strategies.

Specifically, only four ostracism strategies are implemented this time: 1) DQs who ostracize defectors, 2) CQs who ostracize cooperators, 3) NQs who never ostracize, and 4) AQs who ostracize everyone.

To match the proportions of the strategies in the original strategy set, I set the initial proportions of DQ, CQ, NQ, and AQ as 3/10, 3/10, 3/10, and 1/10, respectively. The rationale is that 1) the new proportion of DQ equals to the sum of the proportions of DQ, DXQ, and DYQ in the original model, 2) the new proportion of CQ equals to the sum of the proportions of CQ, CXQ, and CYQ in the original model, 3) the new proportion of NQ equals to the sum of the proportions of NQ, XQ, and YQ in the original model, and 4) the new proportion of AQ equals to the proportion of AQ in the original model. Similarly, when mutation happens, the probability that an agent's strategy mutates to DQ, CQ, NQ, and AQ equals to 3/10, 3/10, 3/10, and 1/10, respectively.

Figure 51 compares the cooperation rate with vs. without symbolic enforcement in the early, middle, and late phases of the simulation. In all these three phases, the symbolic enforcement boosts cooperation.

Three 2 (with vs. without symbolic enforcement) * 6 ($\sigma \in [0, 0.05, 0.1, 0.15, 0.2, 0.25]$) ANOVA were done on the early, middle, and late phases of the simulation to examine the impact of symbolic enforcement on cooperation under different σ levels. In these conditions, agents' cooperation strategies are non-transparent and their symbolic strategies are transparent.

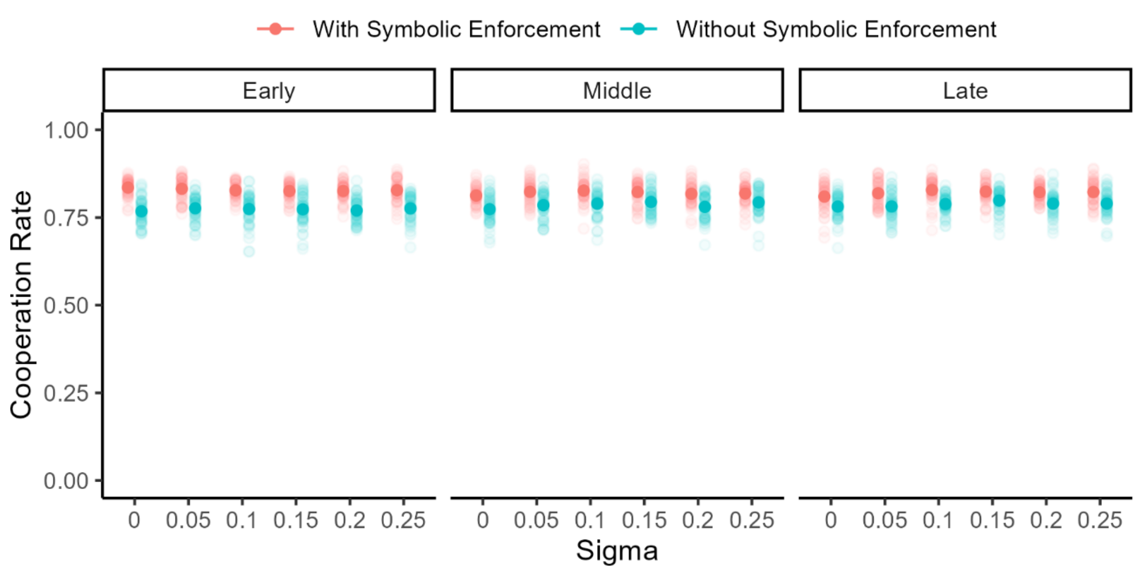
For the early phase, the interaction between enforcement and σ is not significant ($F(5, 588) = 0.848$, $p = 0.516$). The main effect of enforcement is significant ($F(1, 588) = 448.883$, $p < 0.001$). The cooperation rate is significantly higher with symbolic enforcement ($M = 0.829$, $SE = 0.002$) than without symbolic enforcement ($M = 0.773$, $SE = 0.002$). The main effect of σ is not significant ($F(5, 588) = 0.437$, $p = 0.823$).

For the middle phase, the interaction between enforcement and σ is not significant ($F(5, 588) = 0.810$, $p = 0.543$). The main effect of enforcement is significant ($F(1, 588) = 163.775$, $p < 0.001$). The cooperation rate is significantly higher with symbolic enforcement ($M = 0.820$, $SE = 0.002$) than without symbolic enforcement (M = 0.786, SE = 0.002). The main effect of σ is significant ($F(5, 588) = 3.314$, p = 0.006), but does not have a consistent monotonic effect.

For the late phase, the interaction between enforcement and σ is not significant ($F(5, 588)$ = 0.741, p = 0.593). The main effect of enforcement is significant ($F(1, 588) = 156.4948$, $p < 0.001$). The cooperation rate is significantly higher with symbolic enforcement (M = 0.821, SE = 0.002) than without symbolic enforcement (M = 0.788, SE = 0.002). The main effect of σ is significant ($F(5, 588) = 3.203$, p = 0.007), but does not have a consistent monotonic effect.

These results replicate the results in the main text and show that although it seems irrational to enforce a symbolic norm, being able to ostracize others based on their symbolic strategy helps agents avoid defectors and build mutual cooperation.

Figure 51 The Impact of Symbolic Norm Enforcement on Cooperation Using Another Model Option



2 Supplementary Methods and Results for Robustness Tests

2.1 Partial Strategy Transparency

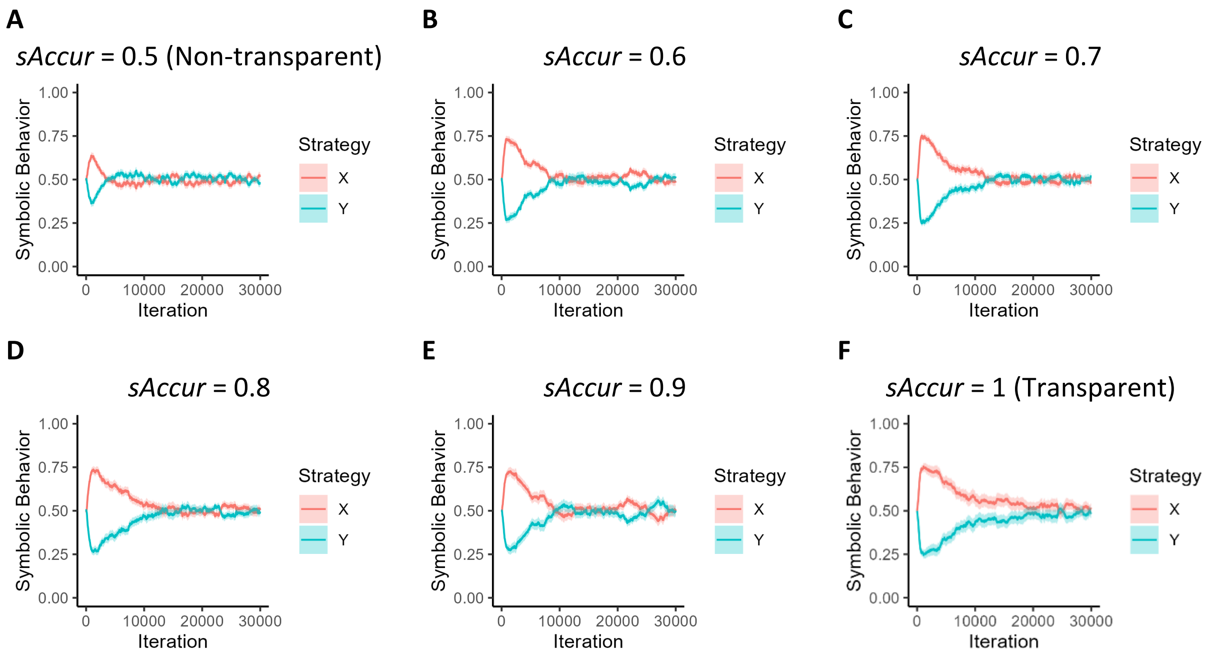
In the default model, when a strategy is transparent, everyone knows each other's strategy with 100% accuracy. When a strategy is non-transparent, individuals' initial impressions of that strategy are completely random, with 50% accuracy. In this robustness test, I made the strategy transparency as a continuous variable between 0.5 and 1. Specifically, two sets of simulations were conducted.

2.1.1 Non-transparent Cooperation Strategy and Partial Transparent Symbolic Strategy

In the first set of simulations, as in the default model, agents' cooperation strategies are non-transparent. An agent's initial impression of another agent's cooperation strategy is completely random. However, agents' symbolic strategies are partial transparent. I manipulated the initial accuracy of the symbolic strategy. Before *Agent A* interacts with *Agent B* for the first time, *Agent A* randomly guesses the symbolic strategy of *Agent B* with an accuracy of *sAccur* ∈ [0.6, 0.7, 0.8, 0.9]. Agents know each other's symbolic strategy after they interact once in the cooperation game. This manipulation aims to test whether the symbolic strategy needs to be completely transparent for norm spillover to happen or it can happen as long as the symbolic strategy is not completely opaque.

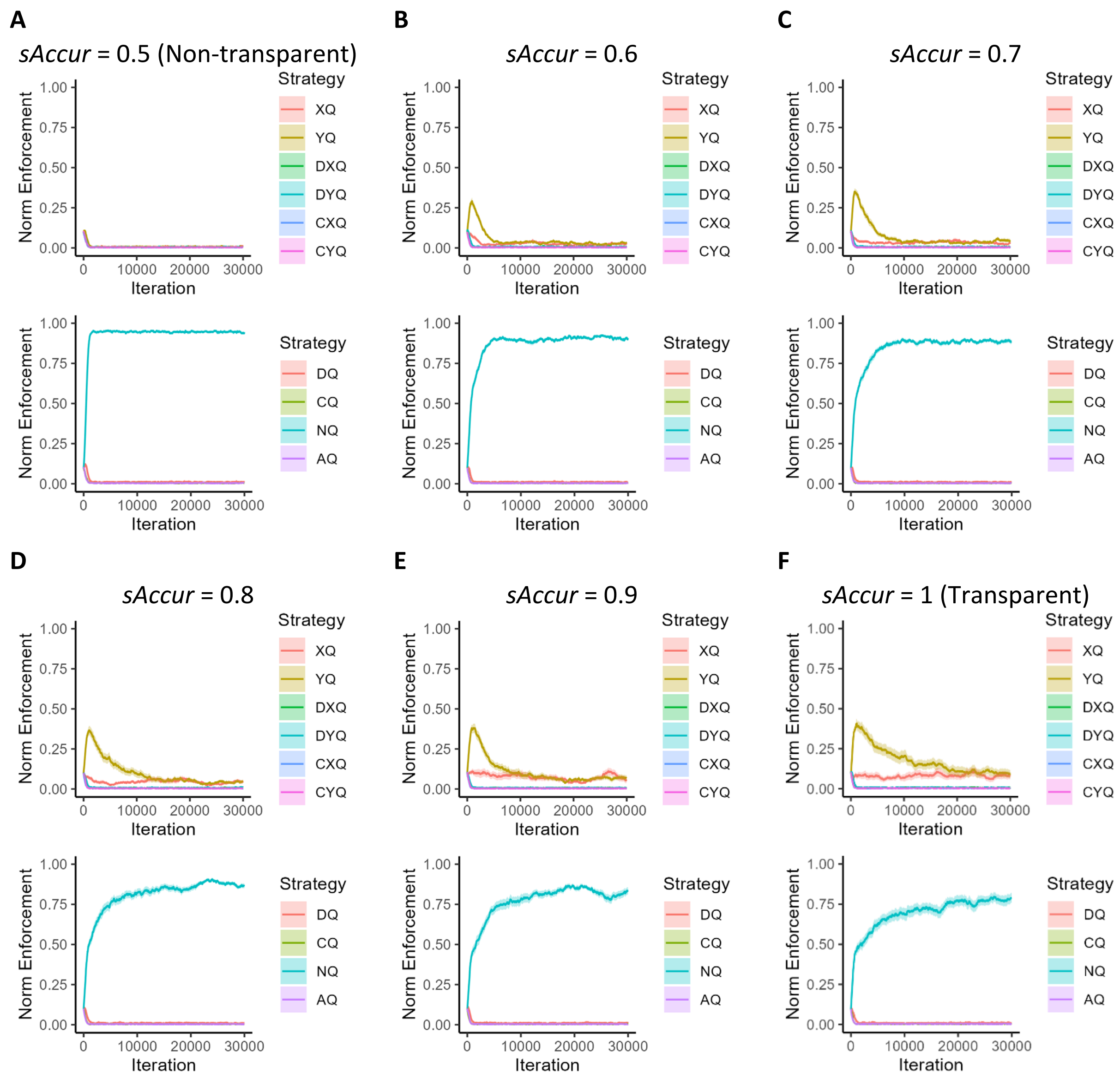
50 simulation runs were run under each *sAccur* ∈ [0.6, 0.7, 0.8, 0.9] under $\sigma = 0.15$. **Figure 52** shows that as agents' symbolic strategies become increasingly transparent, the norm spillover effect becomes stronger. **Figure 53** shows that when symbolic strategies are completely non-transparent (i.e., *sAccur* = 0.5), agents do not ostracize others based on their symbolic strategies. However, as long as agents' initial impressions of others' symbolic strategies are more accurate than chance (i.e., *sAccur* > 0.5), even when the accuracy is low (e.g., as low as 0.6), agents will ostracize Y-performers. This result shows that symbolic strategies do *not* need to be completely transparent for the norm spillover effect to happen. As long as agents have some (even imperfect) knowledge of each other's symbolic strategy before they interact directly, agents will ostracize symbolic norm violators.

Figure 52 The Evolution of Symbolic Norm When Symbolic Strategies Are Partial Transparent



Note. The trajectories of the evolution of symbolic strategies. For this and next figure, different *sAccur* levels represent the accuracy of agents' initial impressions of each other's symbolic strategies. When *sAccur* = 0.5, agents' initial impressions of each other's symbolic strategies are completely random, meaning that the symbolic strategies are non-transparent. When *sAccur = 1*, agents' initial impressions of each other's symbolic strategies are completely accurate, meaning that the symbolic strategies are transparent.

Figure 53 The Evolution of Norm Enforcement When Symbolic Strategies Are Partial Transparent



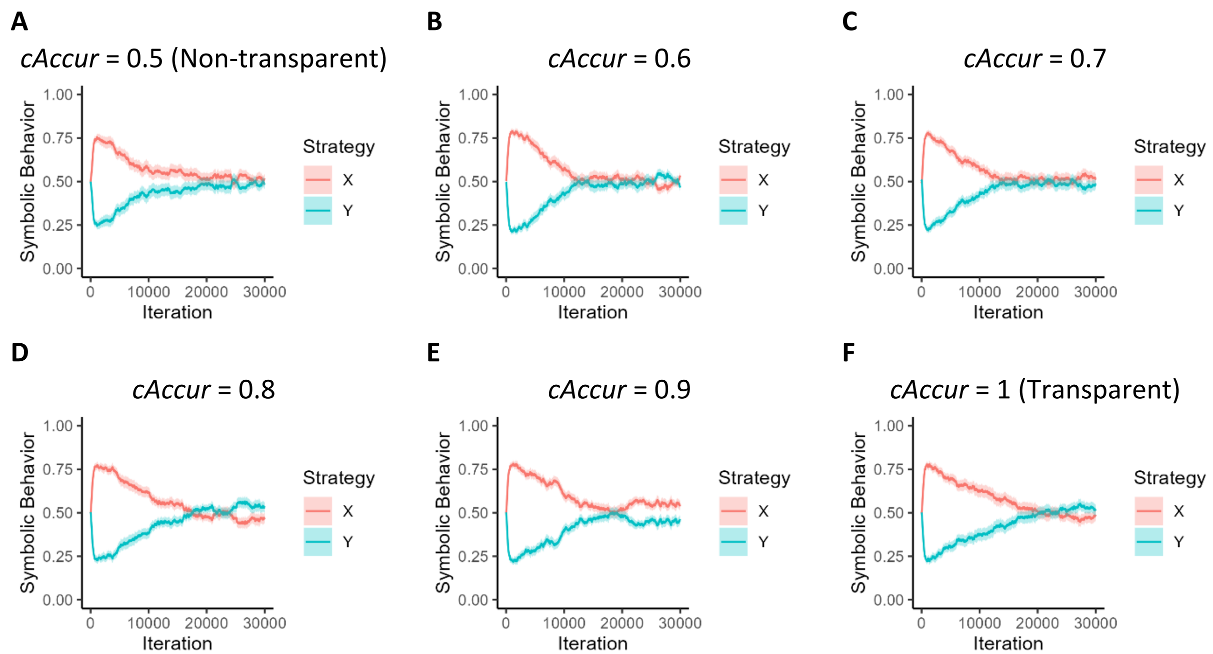
2.1.2 Transparent Symbolic Strategy and Partial Transparent Cooperation Strategy

In the second set of simulations, as in the default model, agents' symbolic strategies are completely transparent (i.e., initial impression with 100% accuracy). However, agents' cooperation strategies are partial transparent. I manipulated the initial accuracy of the cooperation strategy. Before *Agent A* interacts with *Agent B* for the first time, *Agent A* randomly guesses the cooperation strategy of *Agent B* with an accuracy of *cAccur* ∈ [0.6, 0.7, 0.8, 0.9]. Agents know each other's cooperation strategy after they interact once in the cooperation game. This manipulation aims to examine how the transparency of the practical behavior influences the norm spillover effect.

Figure 54 shows that the symbolic norm *X* evolved regardless of the transparency of the practical behavior. **Figure 55** shows that the prevalence of X-performers did not differ under different levels of transparency of the cooperation strategy (i.e., *cAccur*). Three one-way ANOVA tests were done for the early, middle, and late phases of the simulations to compare the prevalence of X-performers across different levels of *cAccur*. The main effects of *cAccur* were not significant (early phase: $F(5, 294) = 0.565$, $p = 0.727$; middle phase: $F(5, 294) = 0.353$, $p = 0.880$; late phase: $F(5, 294) = 1.444$, $p = 0.208$).

Importantly, **Figure 56** shows that the composition of ostracism strategies varies as a function of the transparency of the cooperation strategy. As agents' cooperation strategies become more transparent, more agents become DQs and DYQs and ostracize defectors. This is because when agents have more knowledge about who are cooperators vs. defectors, it is more effective to ostracize defectors directly. Nevertheless, a substantial proportion agents are still YQs and DYQs who ostracize symbolic norm violators. This result suggests that the norm spillover effect can happen even when people have much knowledge of each other's practical behavior, supporting the robustness of our model.

Figure 54 The Evolution of Symbolic Norm When Cooperation Strategies Are Partial Transparent



Note. The trajectories of the evolution of symbolic strategies. For this and **Figure 56**, different *cAccur* levels represent the accuracy of agents' initial impressions of each other's cooperation strategies. When *cAccur* = 0.5, agents' initial impressions of each other's cooperation strategies are completely random, meaning that the cooperation strategies are non-transparent. When *cAccur = 1*, agents' initial impressions of each other's cooperation strategies are completely accurate, meaning that the cooperation strategies are transparent.

Figure 55 Prevalence of X-performers Under Different Transparency of the Cooperation Strategy

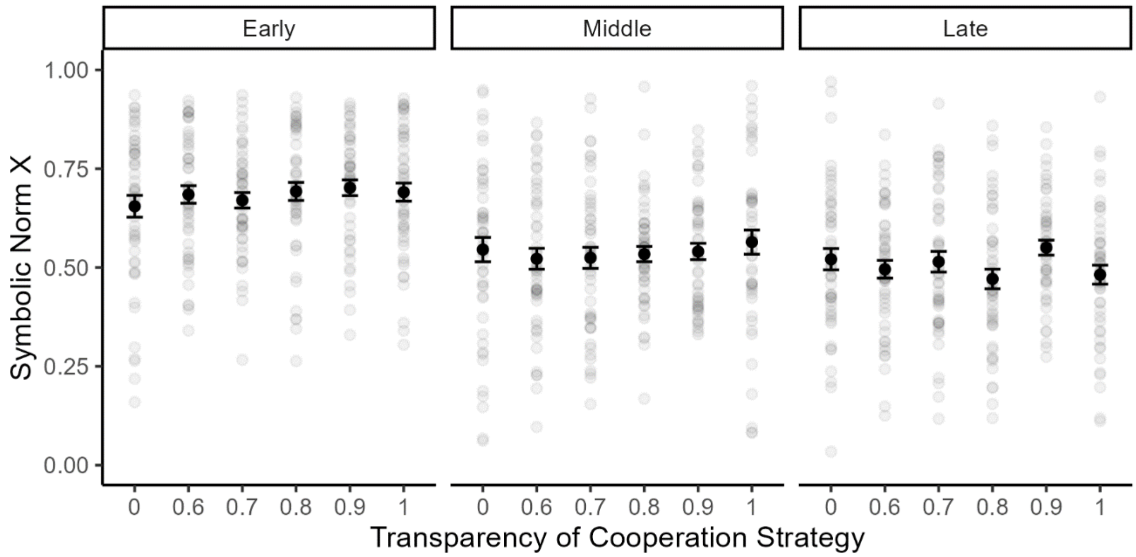
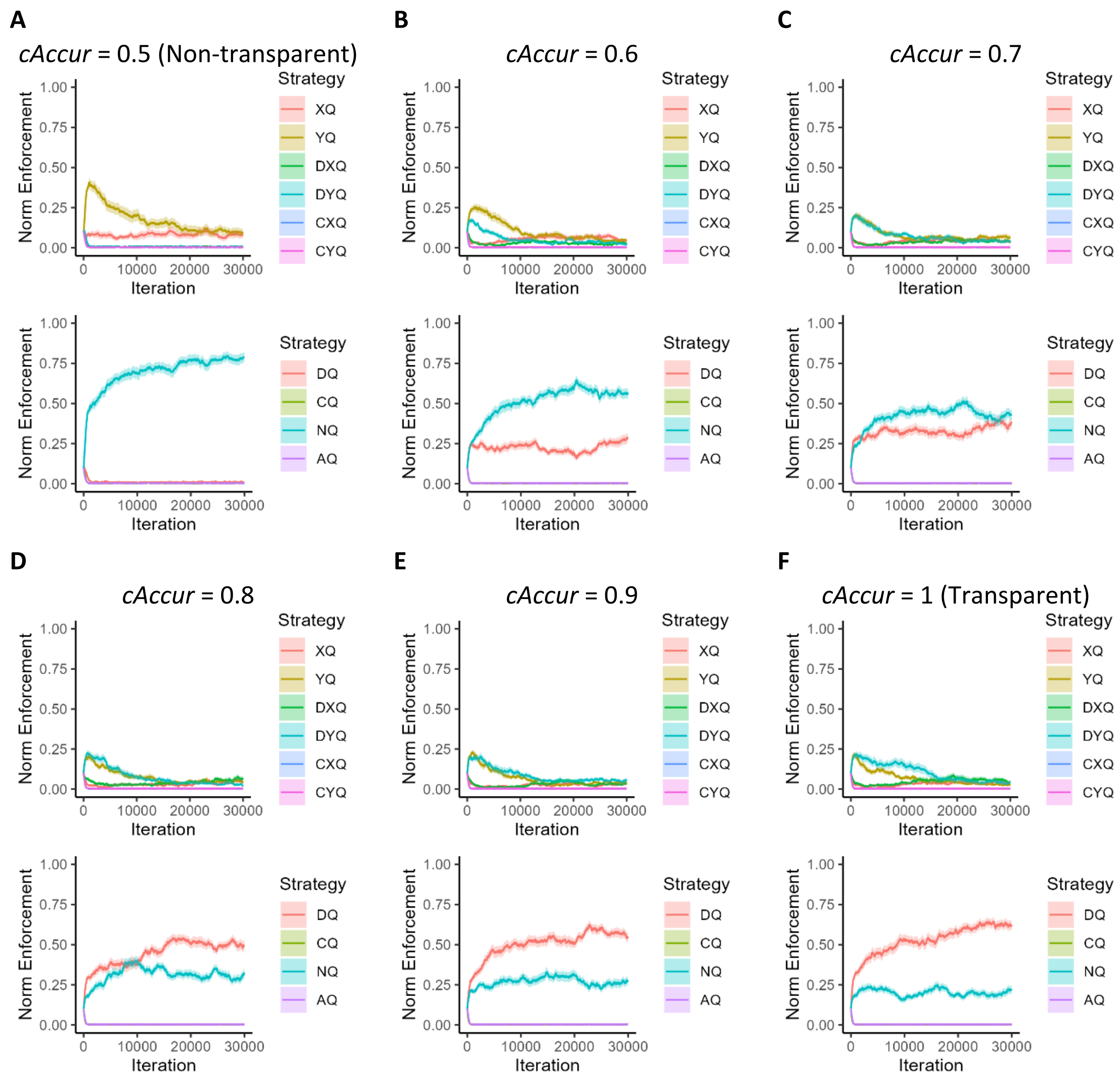


Figure 56 The Evolution of Norm Enforcement When Cooperation Strategies Are Partial Transparent



2.2 Noise in Norm Enforcement

In the default model, if *Agent A* believes that *Agent B* possesses a strategy that *Agent A* ostracizes, *Agent A* will certainly ostracize *Agent B*. This represents a situation with zero noise in norm enforcement. In this section of robustness test, I relaxed this assumption and added a noise *e* to ostracism behaviors. Specifically, when *Agent A* encounters *Agent B*, with a probability of 1 – *e*, *Agent A* decides whether to ostracize *Agent B* based on *Agent A*'s ostracism strategy and *Agent A*'s impression of *Agent B*'s strategy, as in the original model. However, with a probability of *e*, *Agent A* randomly makes decisions. I manipulated the magnitude of *e* and examined its effect. 50 simulation runs were run under each $e \in$ [0.05, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9] under $\sigma = 0.15$.

Figure 57 and **Figure 58** show that as the noise in norm enforcement gets larger, the symbolic norm becomes weaker and there are fewer agents who enforce the symbolic norm. This is a reasonable result because as ostracism becomes more random, the difference between different ostracism strategies becomes smaller. In this case, neither XQ nor YQ can outperform the other. It becomes harder for the symbolic norm to evolve.

Figure 57 The Evolution of Symbolic Norm With Noise in Norm Enforcement

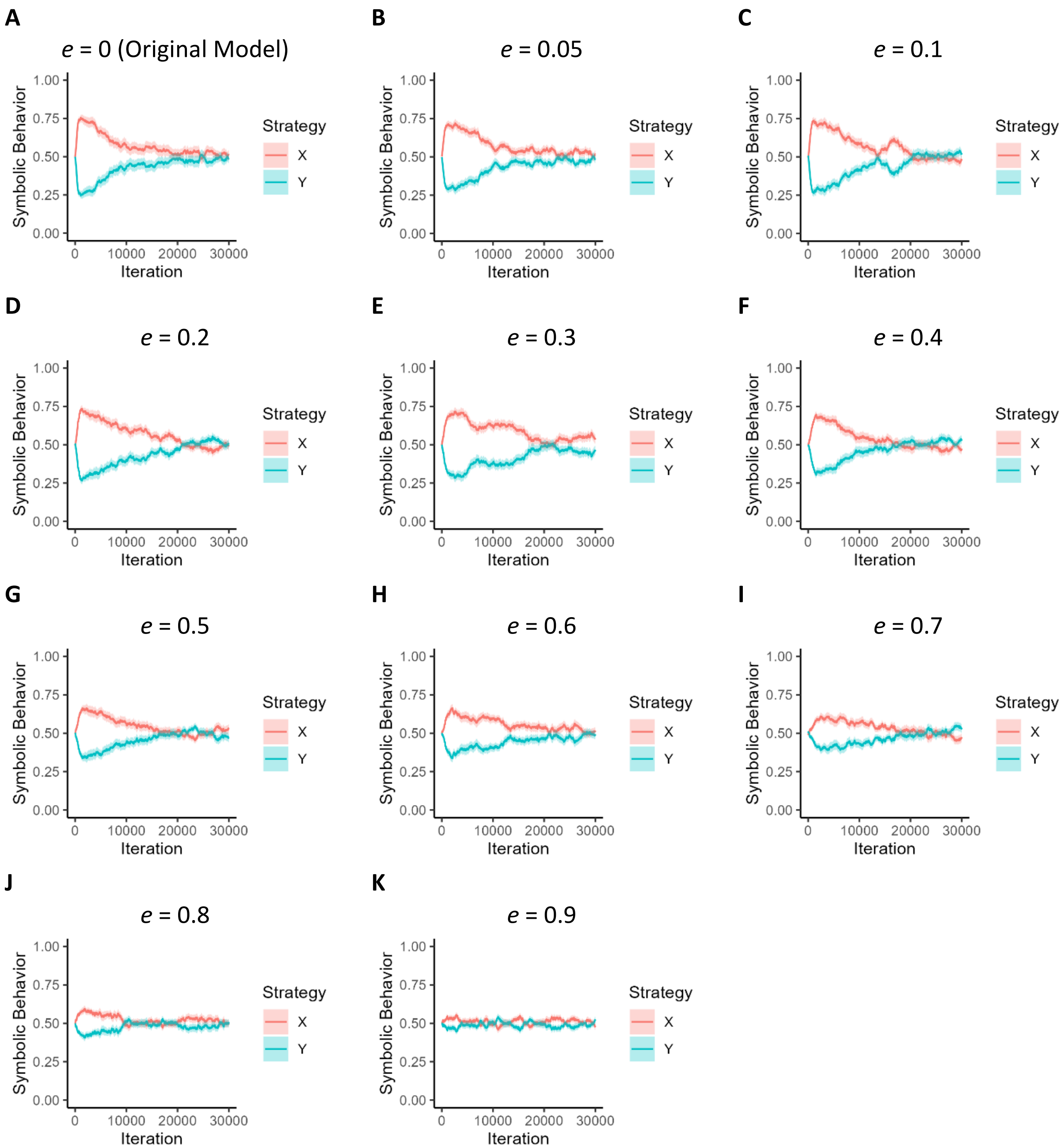
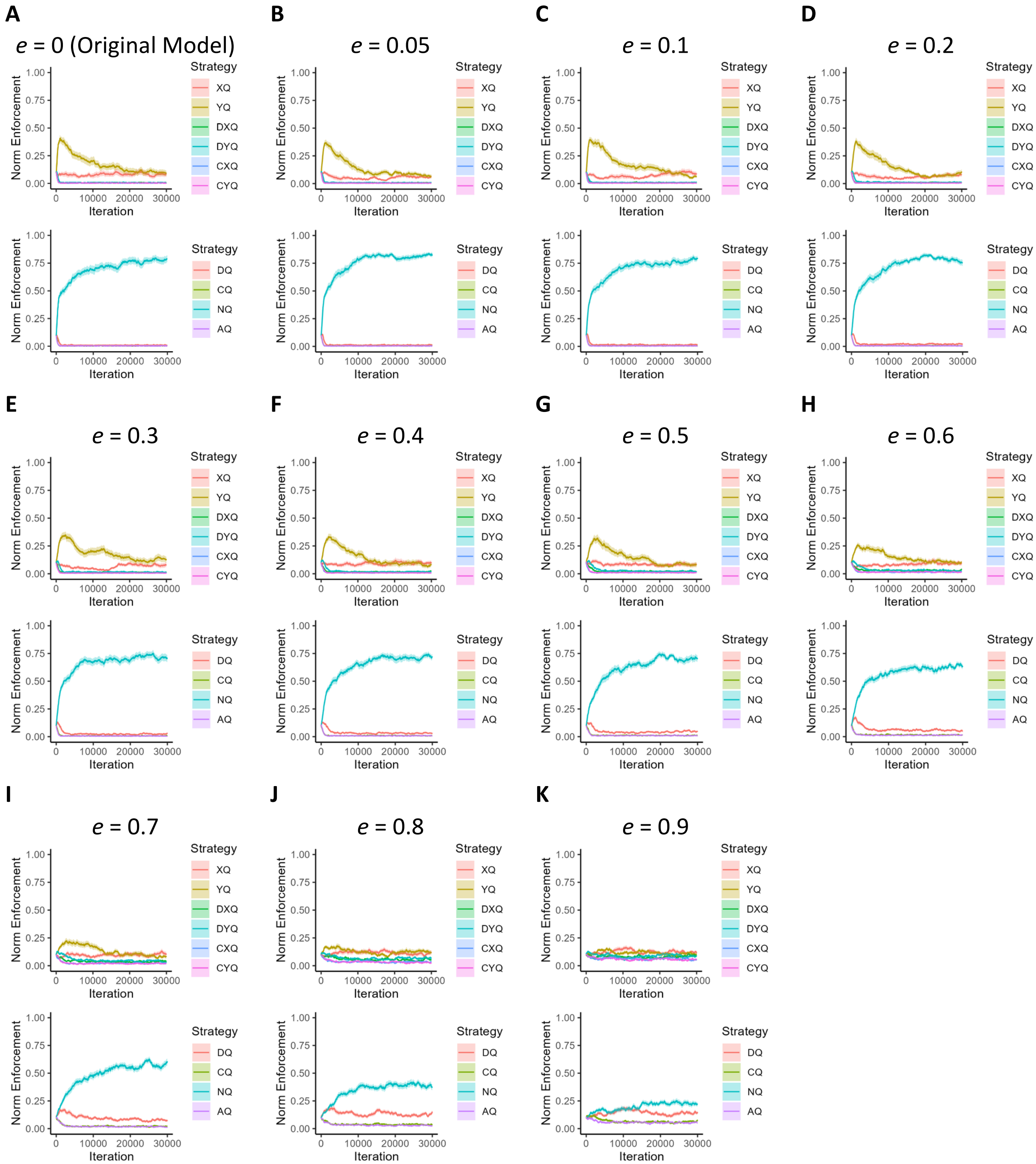


Figure 58 The Evolution of Norm Enforcement With Noise in Norm Enforcement



2.3 Ecological Threat

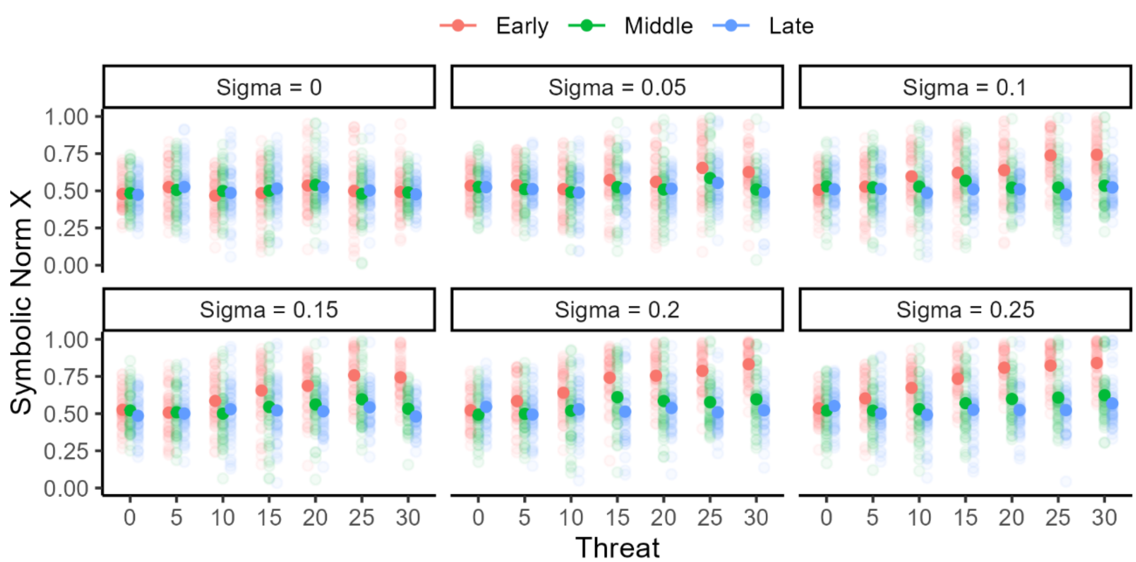
In the default model, the level of threat τ was set as 15, which represents a medium level of threat. In this section, different levels of threat were used to examine the effects of ecological threat on the norm spillover effect.

As mentioned in Section **Evolutionary Phases**, the level of threat is manipulated as a deduction of τ from every agent's payoff (Roos et al., 2015). This setup resembles the scenarios in which threats like drought, hurricanes, tornadoes, or famine reduce the general resources that individuals receive from their environment (Popp, 2006; Sivakumar, 2005). In this section, 50 simulation runs were run under each level of $\tau \in [0, 5, 10, 20, 25, 30]$ under each level of $\sigma \in [0, 0.05, 0.1, 0.15, 0.2, 0.25]$.

Figure 59 shows that, when there is a correlation between the symbolic and the practical behaviors (i.e., $\sigma > 0$), as ecological threat increases, the symbolic norm becomes stronger, especially in the early phase of the simulations. In the middle and late phases of the simulations, the symbolic norm fades away, so the effect of threat becomes less prominent. **Figure 60** shows that as threat increases, there are more agents who enforce the symbolic norm (i.e., YQs) especially in the early phases of the simulations. **Figure 61** to **Figure 72** show the trajectories of the evolution of the symbolic norm and norm enforcement under these conditions.

Why does ecological threat lead to a stronger symbolic norm? When threat is higher, everyone's resource is reduced. Based on the principle of diminishing marginal utility (Diener & Biswas-Diener, 2002; Foster, 2004), when the baseline resource is low, every gain and cost from social interactions becomes more influential. In this case, finding a cooperative partner becomes even more important. As the symbolic strategy can be used as a signal for cooperative behavior, more agents adopt and enforce the symbolic norm under higher threat. This result is consistent with previous literature on threat and cultural tightness (Gelfand et al., 2011). When a society faces higher threat from the environment, the society tends to develop a tighter culture with more and stronger social norms. The cultural tightness does not only regulate functional social norms but can also spillover into less functional domains as people bestow symbolic importance on following normative behaviors in general (J. Jackson & Gelfand, 2016). The results in this section echo this literature and suggest that the stronger social norms in tight cultures may be due to the stronger norm spillover effect under higher ecological threat.

Figure 59 The Evolution of Symbolic Norm Under Different Levels of Threat



Note. The evolution of the symbolic norm X under different levels of threat. Each subplot represents a level of σ (i.e., the initial correlation between the symbolic and practical behaviors). Within each subplot, the X-axis represents different levels of ecological threat τ. The red, green, and blue points represent the average values during the early, middle, and late phases of the simulations, respectively.

Figure 60 The Evolution of Norm Enforcement Under Different Levels of Threat

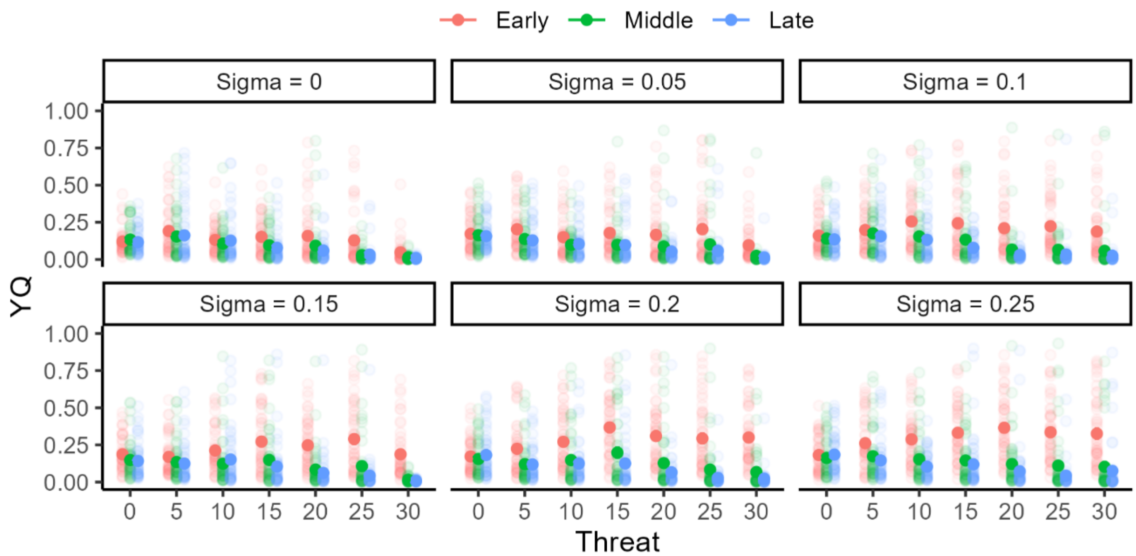


Figure 61 The Evolution of Symbolic Norm When $\tau = 0$

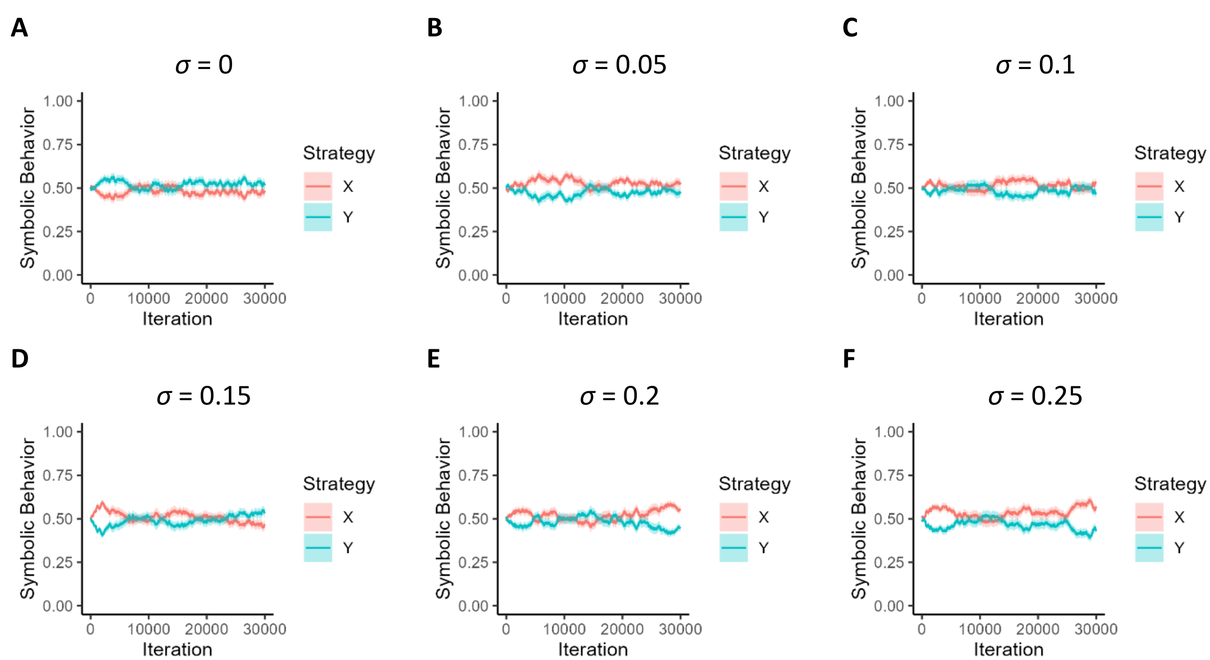


Figure 62 The Evolution of Symbolic Norm When $\tau = 5$

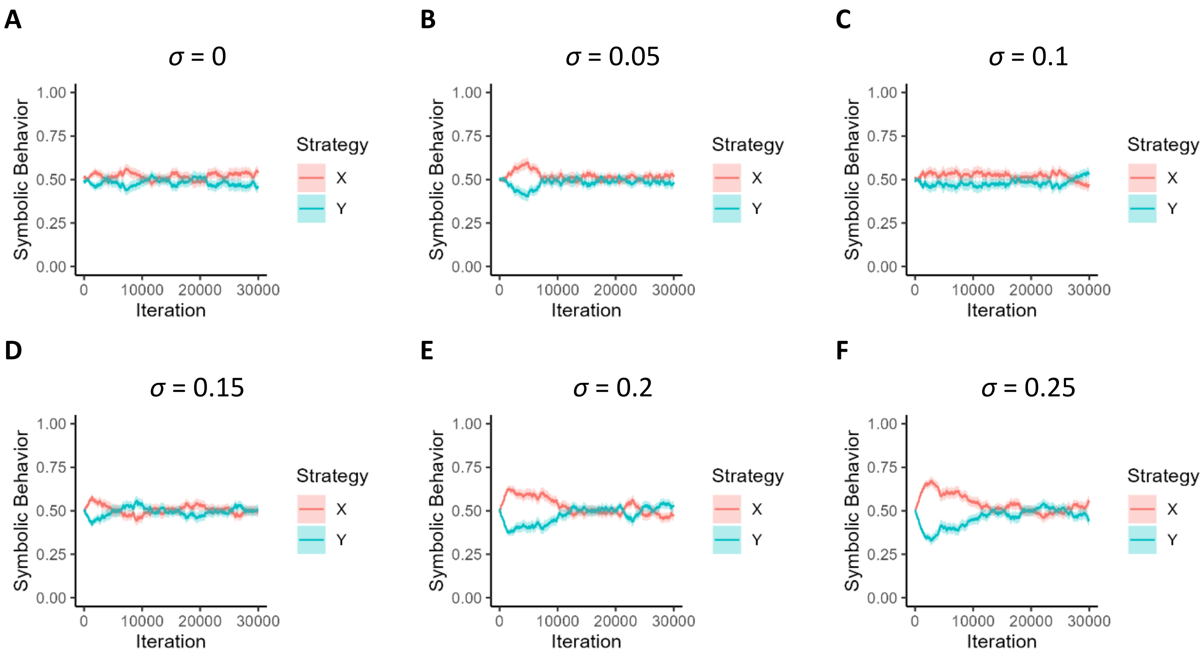


Figure 63 The Evolution of Symbolic Norm When $\tau = 10$

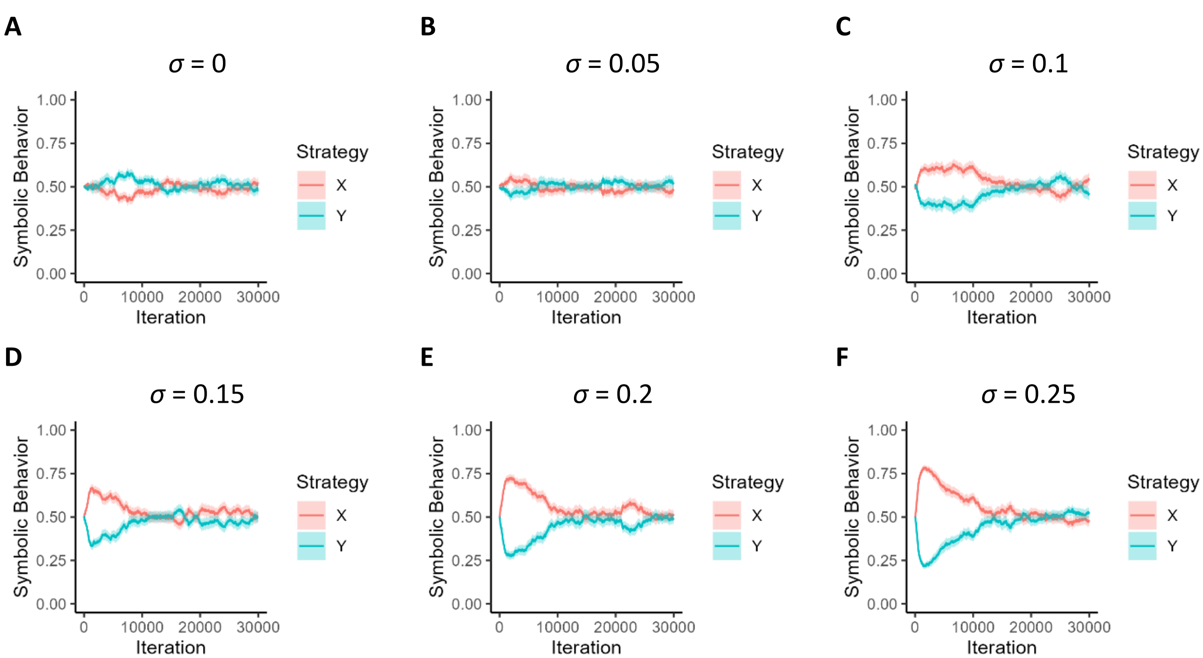


Figure 64 The Evolution of Symbolic Norm When $\tau = 20$

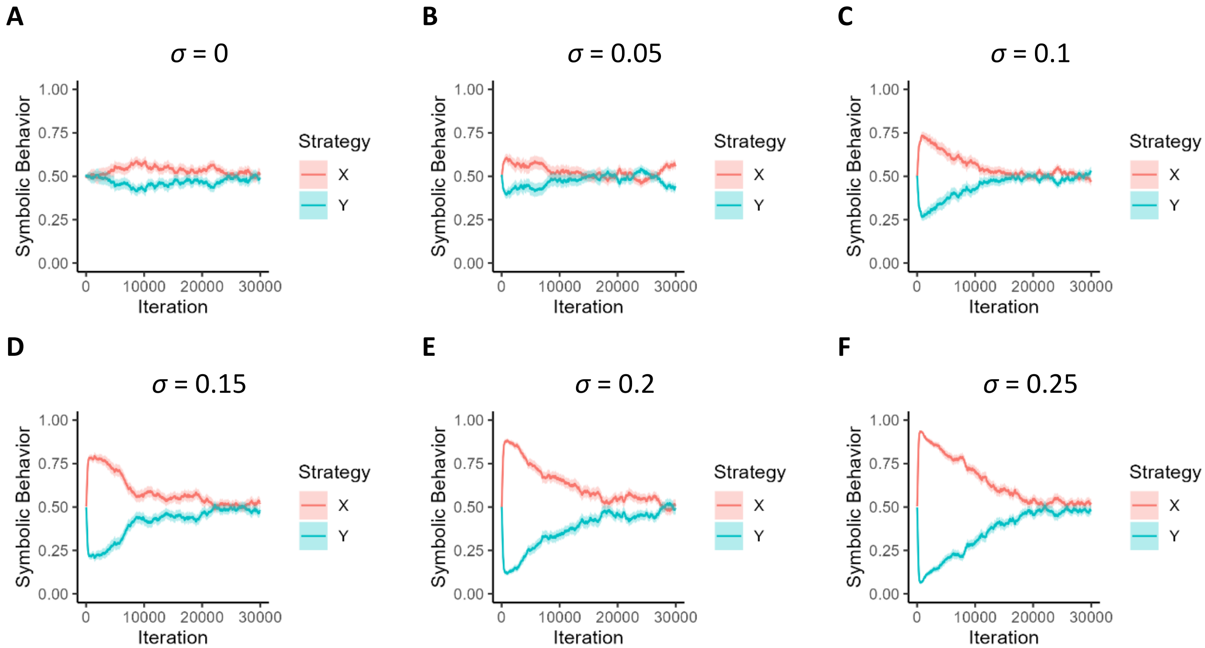


Figure 65 The Evolution of Symbolic Norm When $\tau = 25$

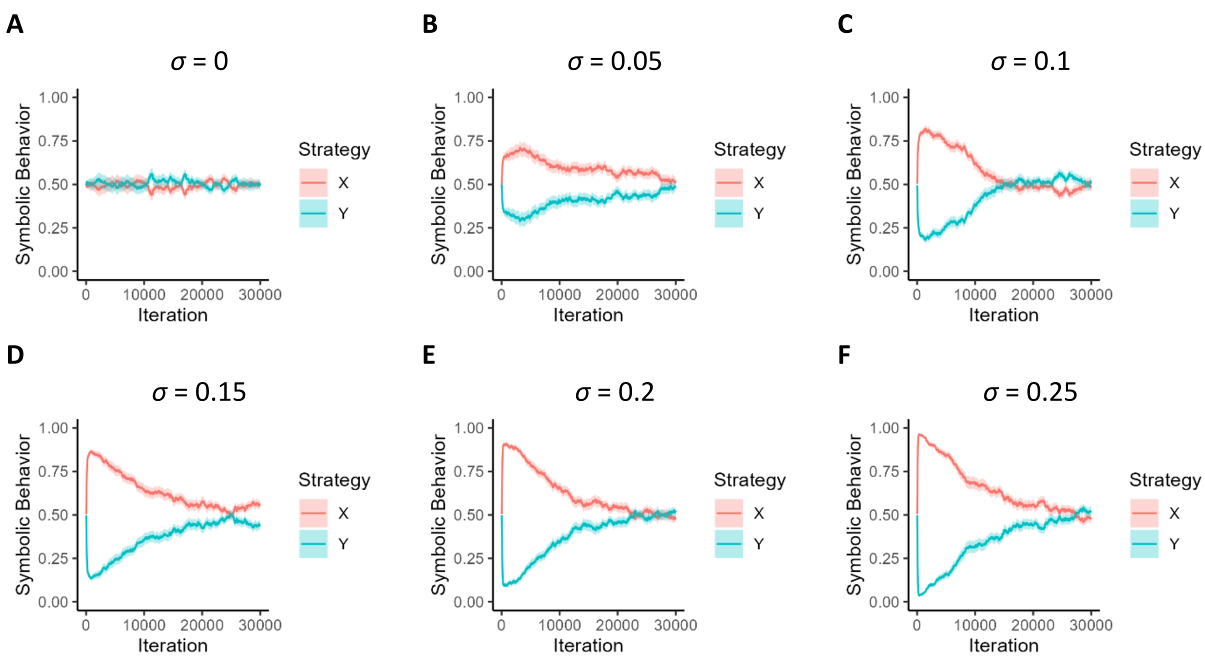


Figure 66 The Evolution of Symbolic Norm When $\tau = 30$

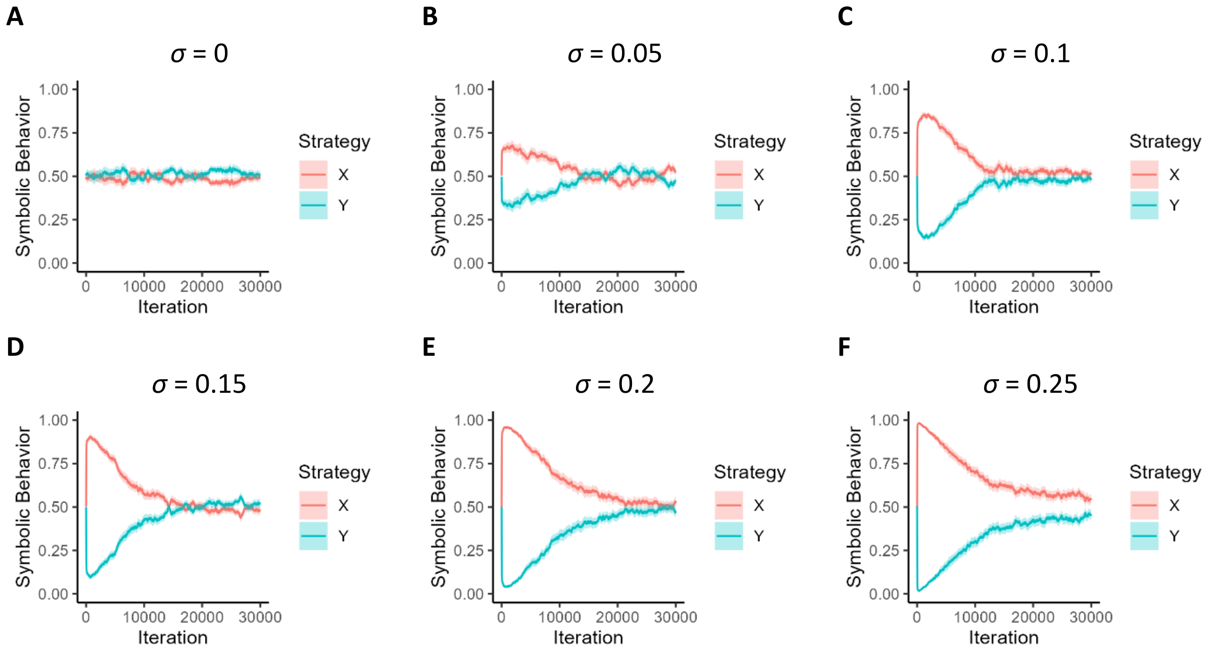


Figure 67 The Evolution of Norm Enforcement When $\tau = 0$

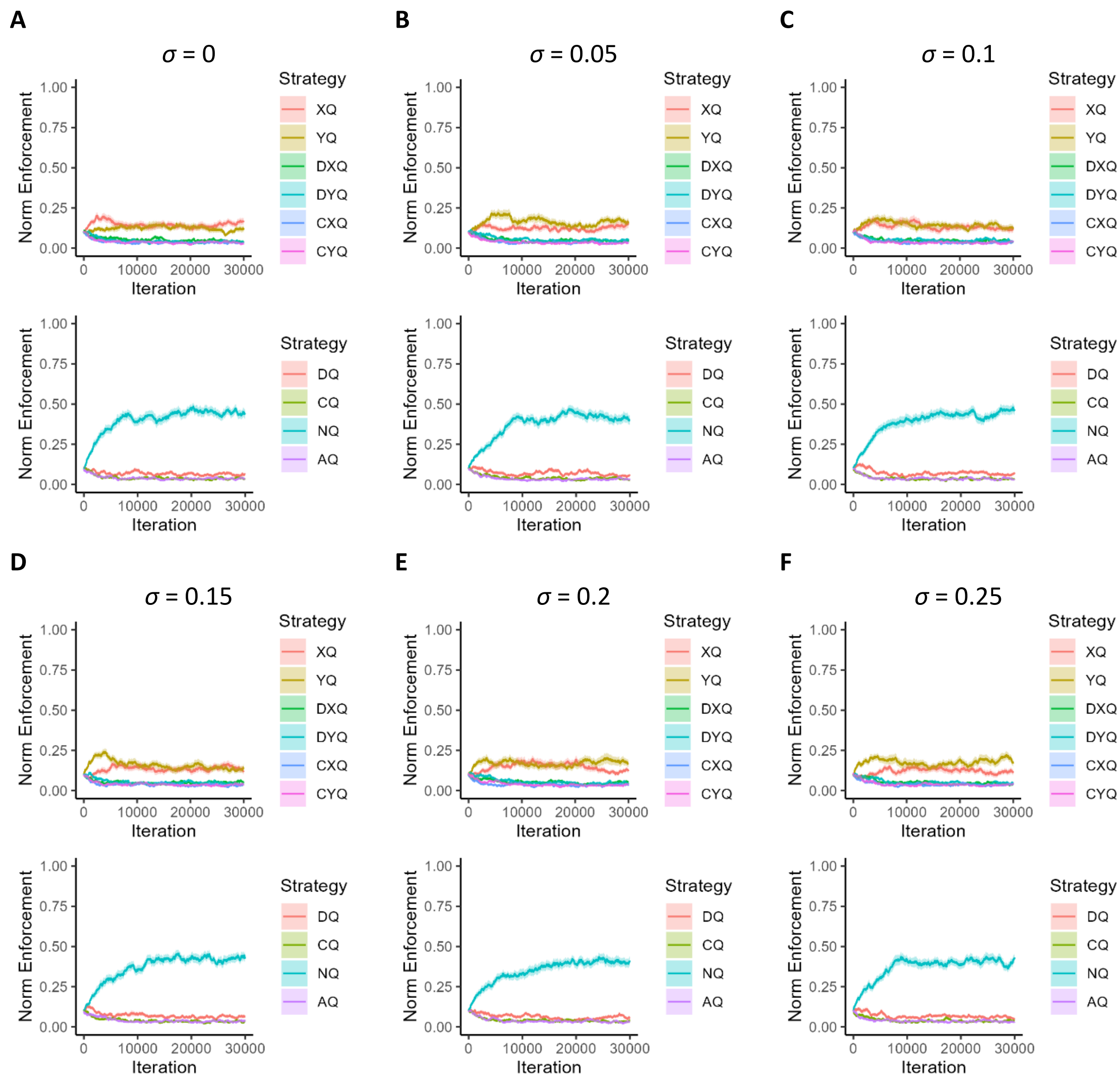


Figure 68 The Evolution of Norm Enforcement When $\tau = 5$

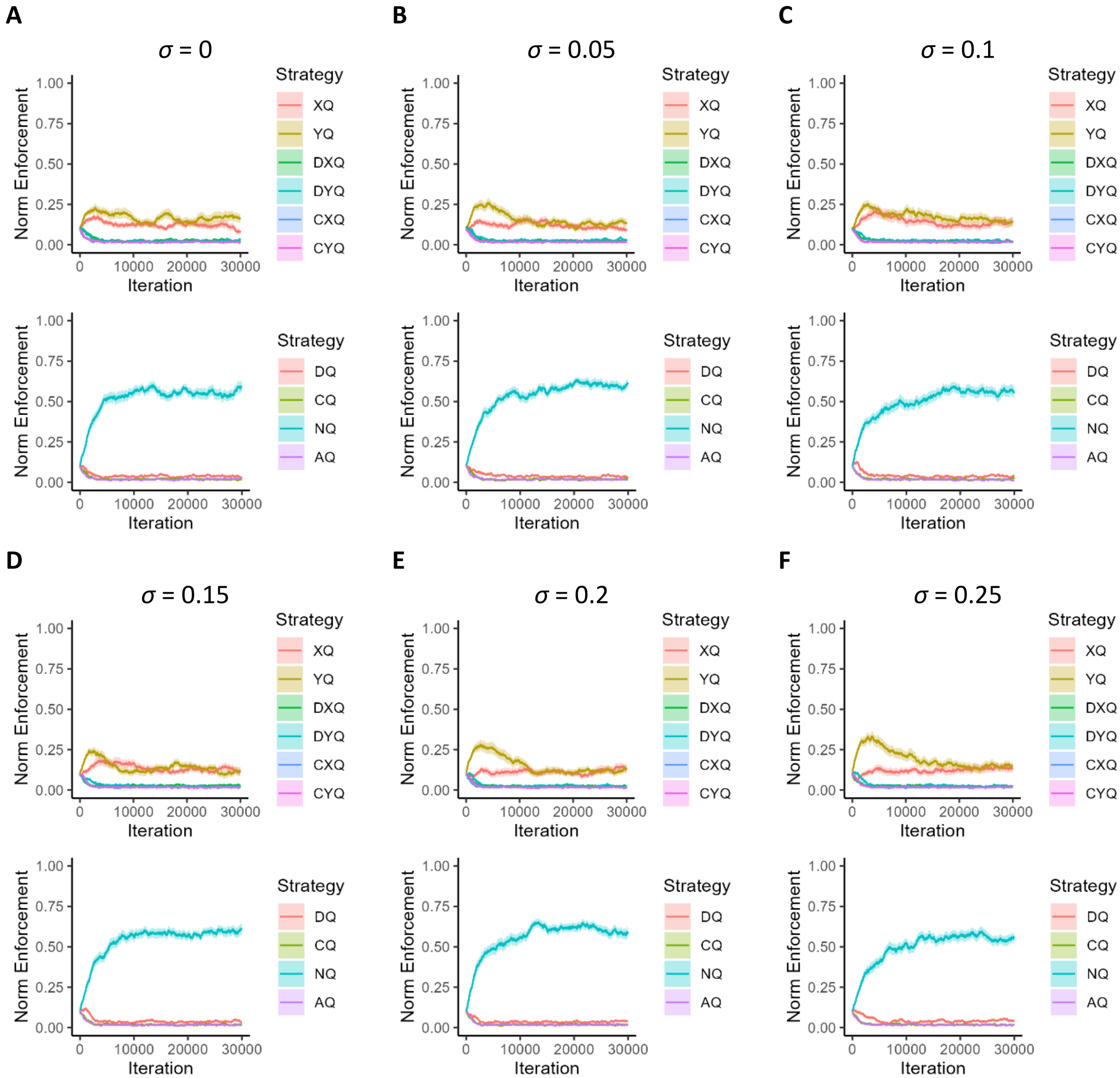


Figure 69 The Evolution of Norm Enforcement When $\tau = 10$

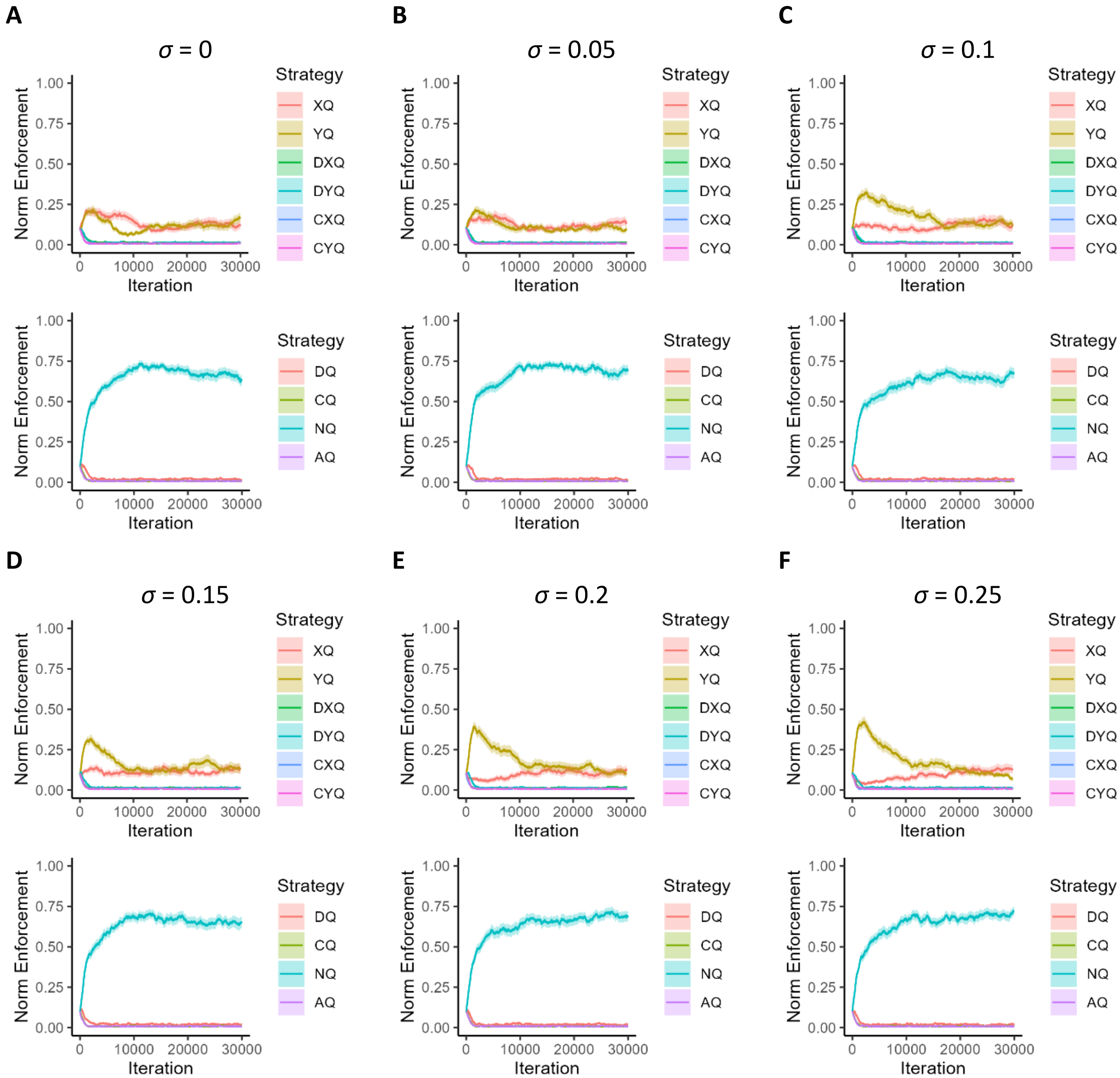


Figure 70 The Evolution of Norm Enforcement When $\tau = 20$

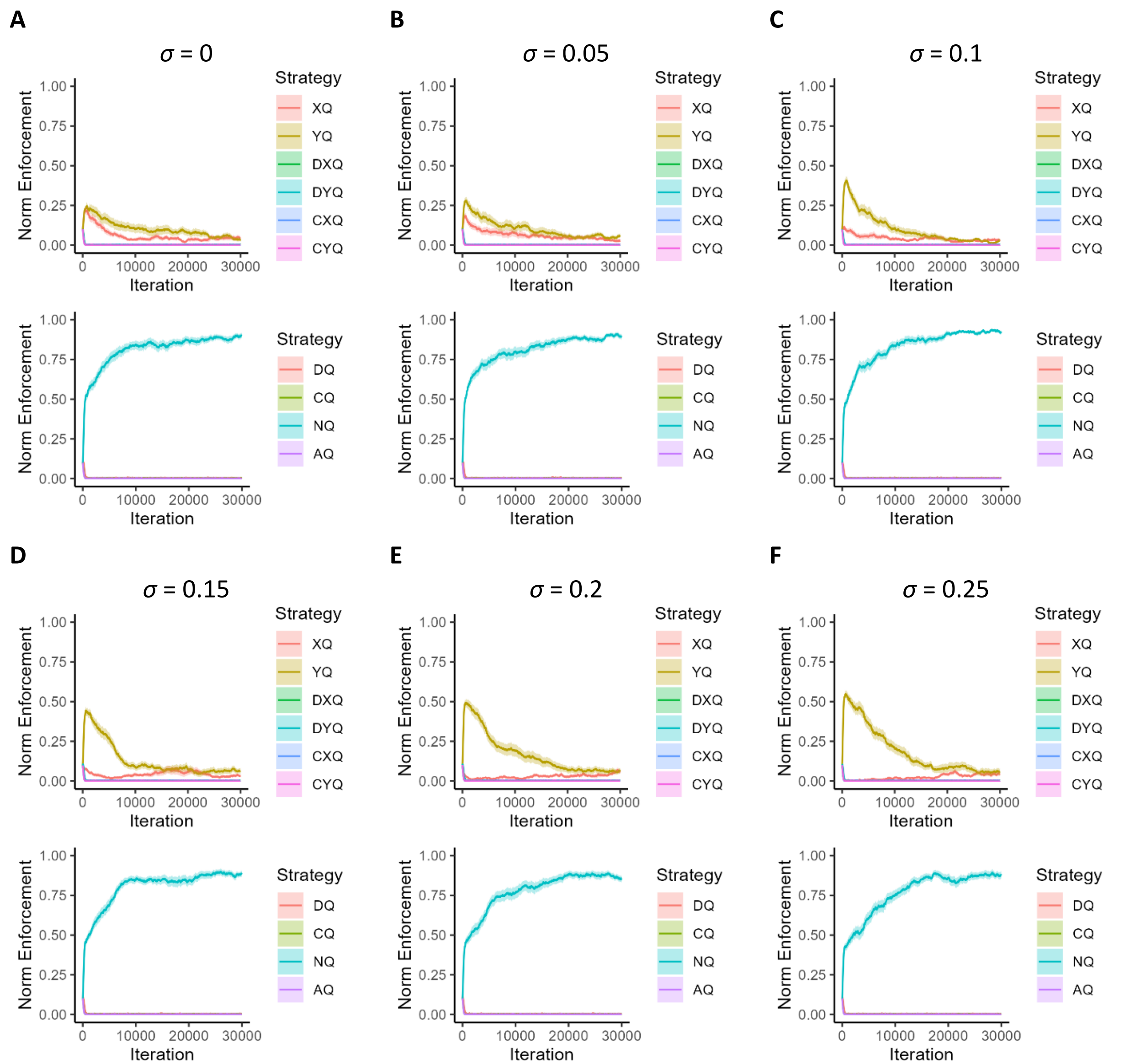


Figure 71 The Evolution of Norm Enforcement When $\tau = 25$

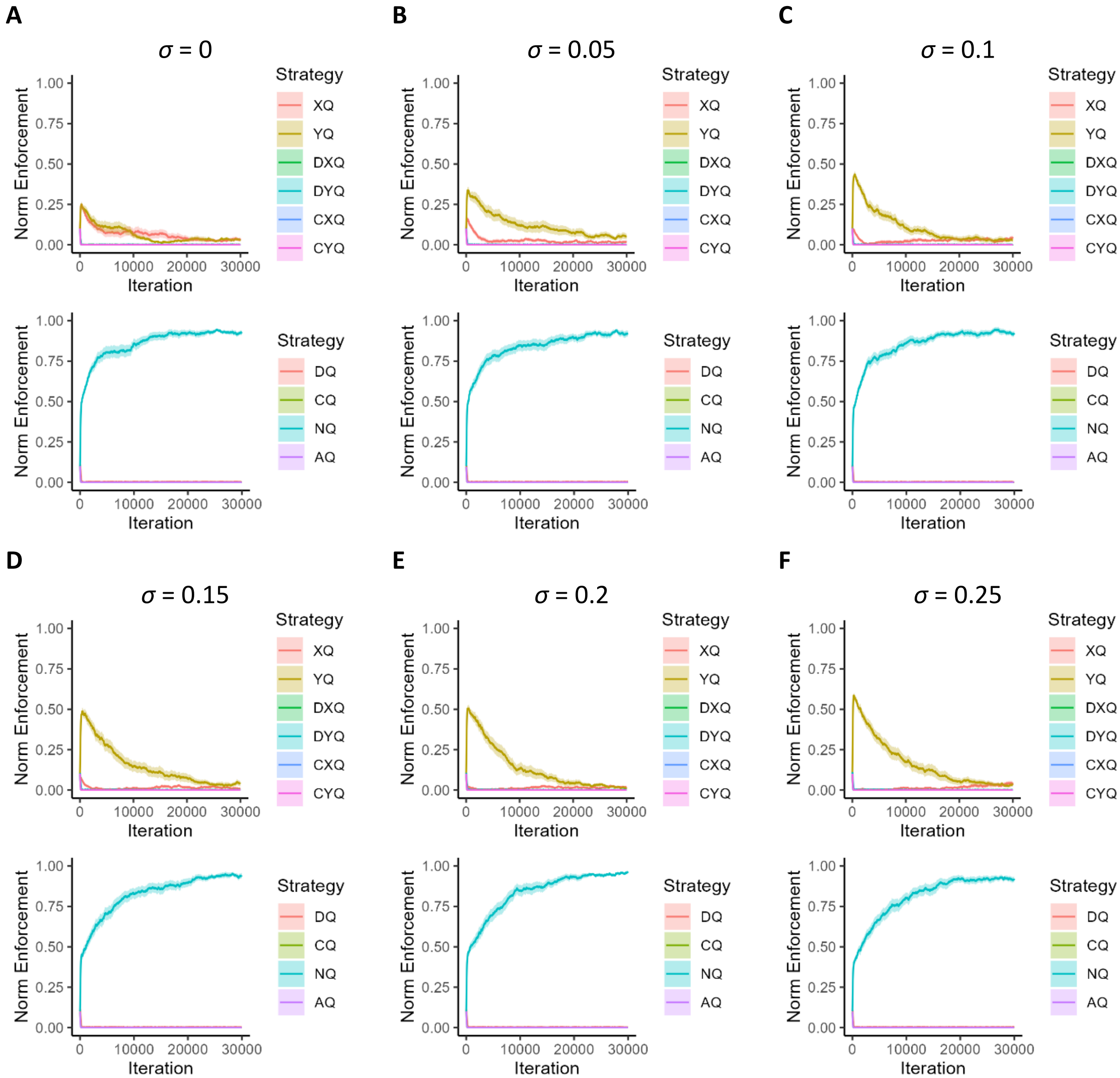
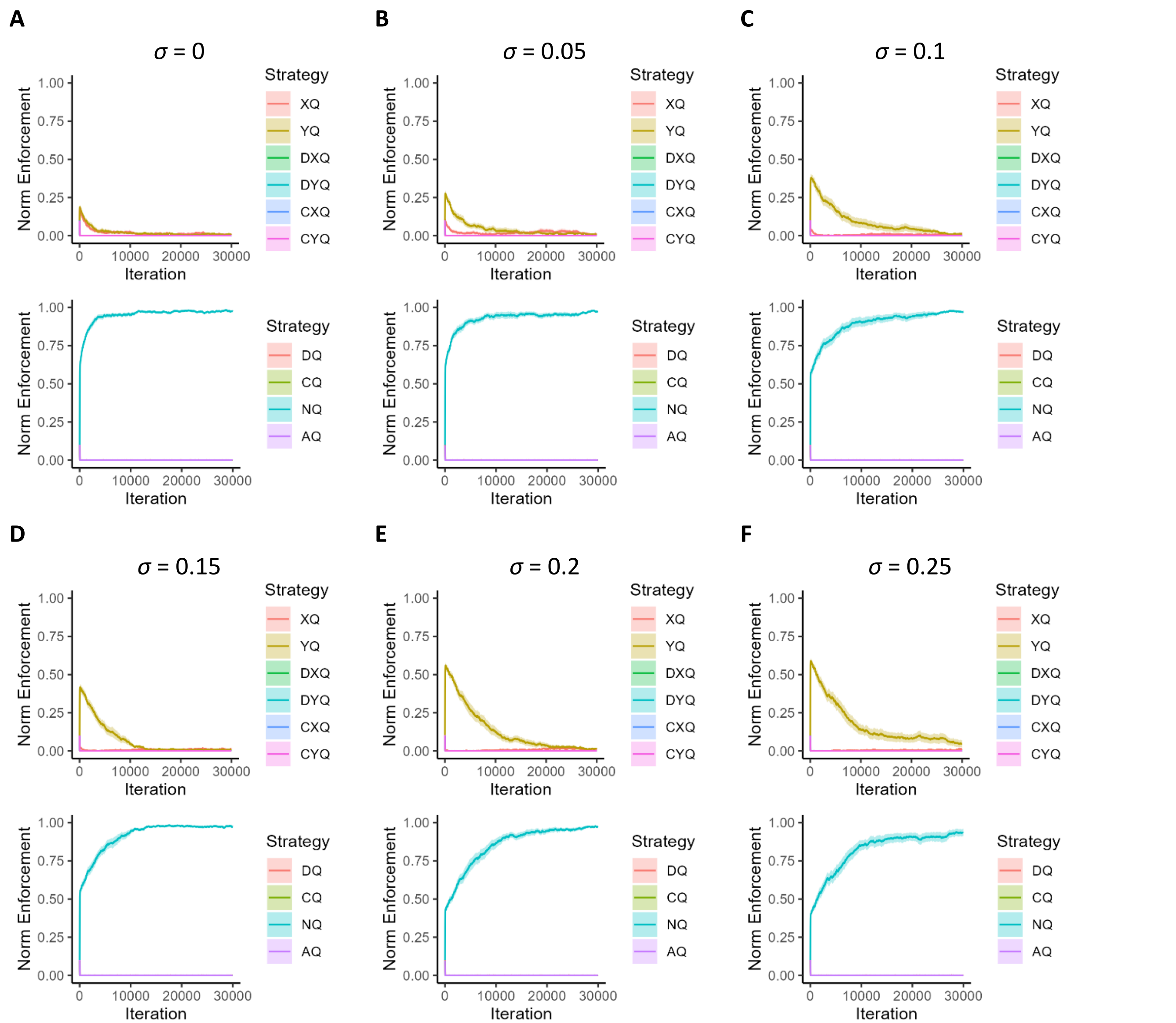


Figure 72 The Evolution of Norm Enforcement When $\tau = 30$



2.4 Death Rate

In this section, I manipulated the death rate δ in the evolutionary dynamic. Death rate is the parameter that controls the speed of evolution. 50 simulation runs were run under each level of $\delta \in$ [0.01, 0.02, 0.05, 0.2, 0.3, 0.4, 0.5] under $\sigma = 0.15$.

Figure 73 and **Figure 74** show that the symbolic norm and norm enforcement evolved in all the conditions when $\delta < 0.5$. When the death rate is too high (i.e., $\delta = 0.5$), few agents could survive, so the strategy composition was nearly random. As the death rate increases, the symbolic norm evolves faster and disappears faster. This is because the parameter death rate controls the speed of evolution and a higher death rate accelerates the change of behaviors.

Figure 73 The Evolution of Symbolic Norm Under Different Death Rates

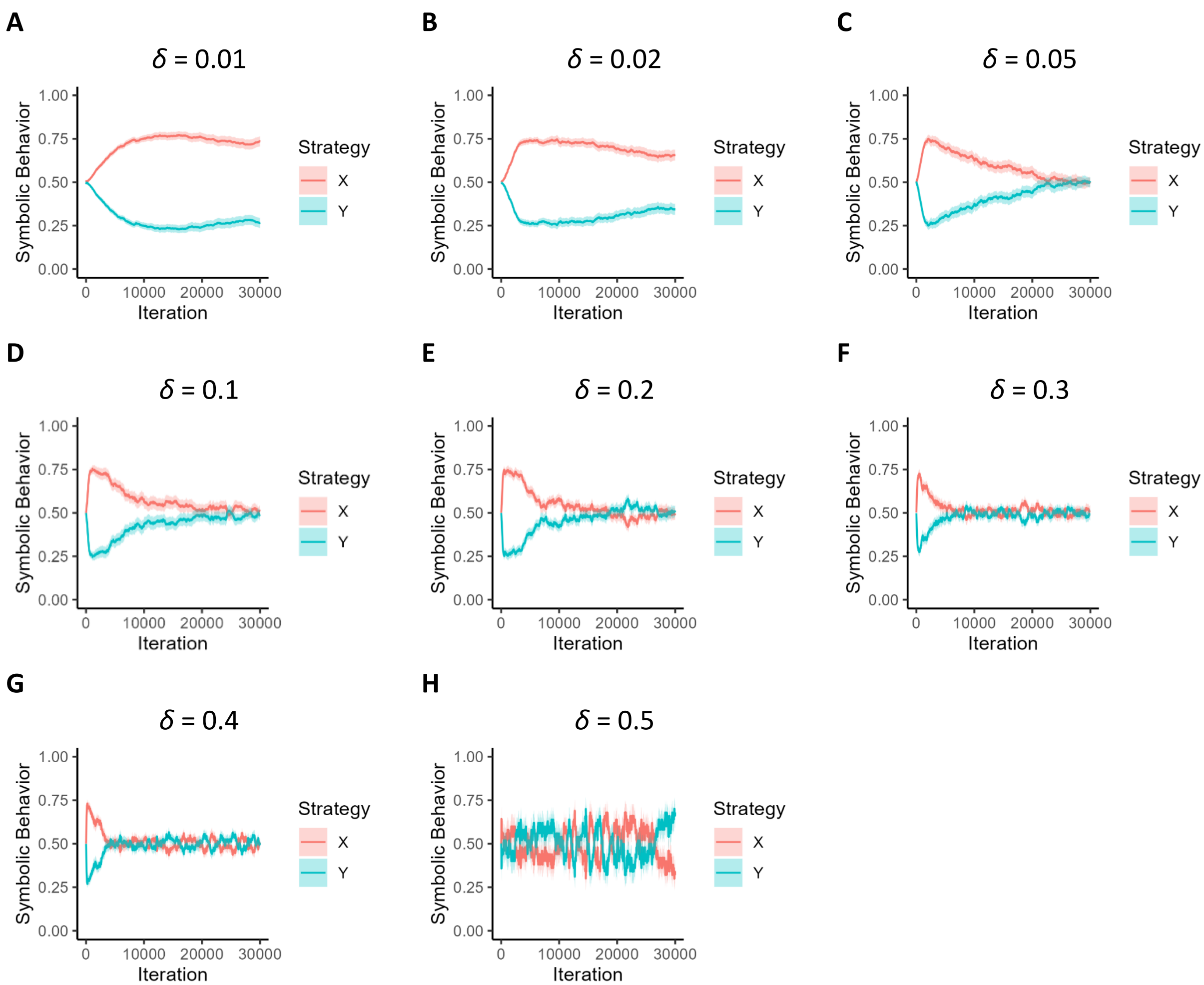
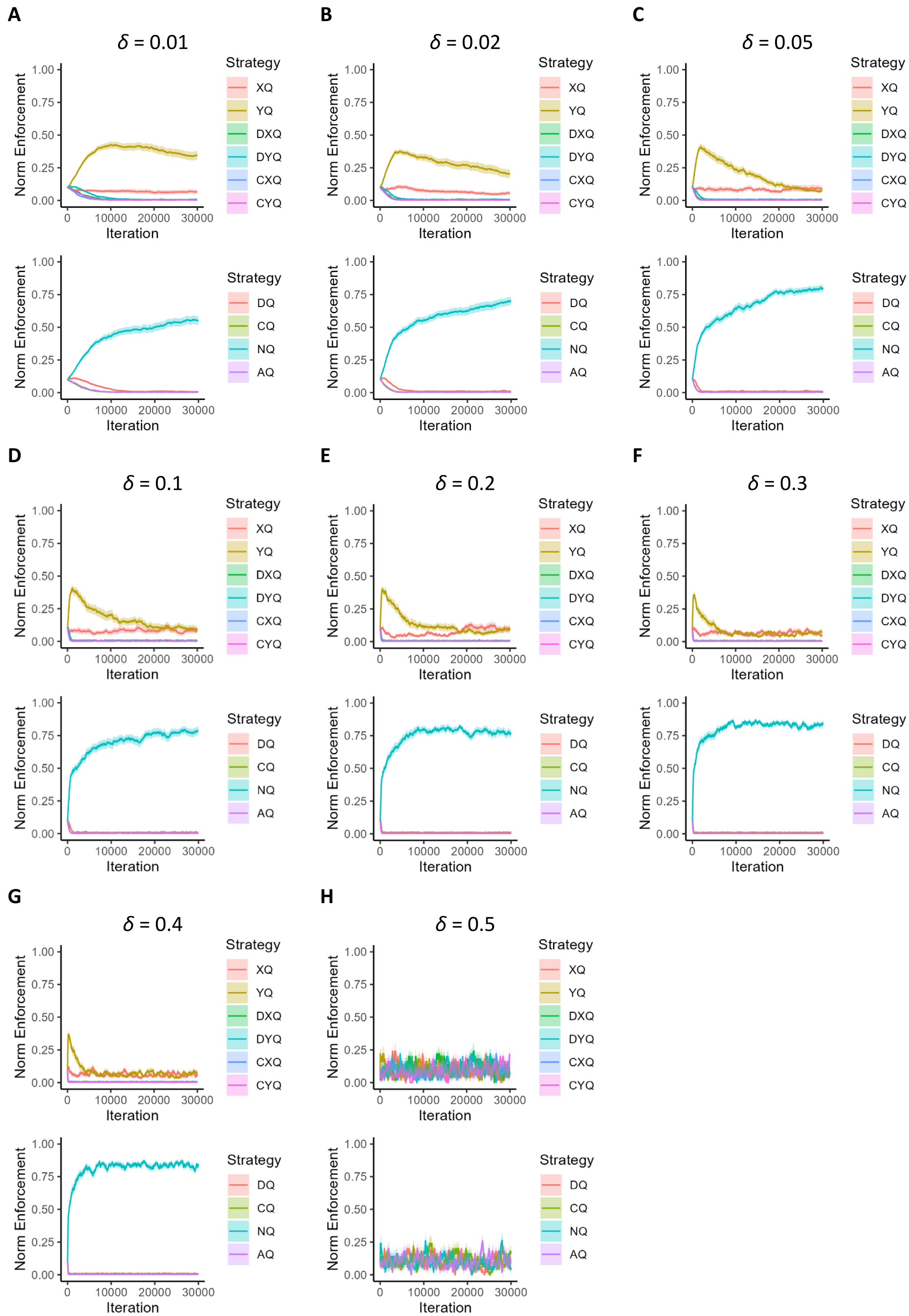


Figure 74 The Evolution of Norm Enforcement Under Different Death Rates



2.5 Mutation

In this section, I tested the robustness of the model under different exploration rates μ and tried another model option of implementing mutation.

2.5.1 Exploration Rate

In the default model, the exploration rate μ was set as 0.005. In this robustness test, 50 additional simulation runs were run under each level of $\mu \in [0.001, 0.002, 0.01, 0.02, 0.05]$ under $\sigma = 0.15$. **Figure 75** and **Figure 76** show that as the exploration rate increases, it becomes harder for the symbolic norm to evolve and maintain. This is reasonable because the exploration rate controls the probability that agents randomly try out new strategies. As random exploration becomes more frequent, behaviors become more random and it is harder to maintain a norm (Allen et al., 2012). Nevertheless, when μ is not too high, the findings in the main text were replicated: The symbolic norm evolved and a substantial proportion of agents evolved to enforce the symbolic norm.

Figure 75 The Evolution of Symbolic Norm Under Different Mutation Rates

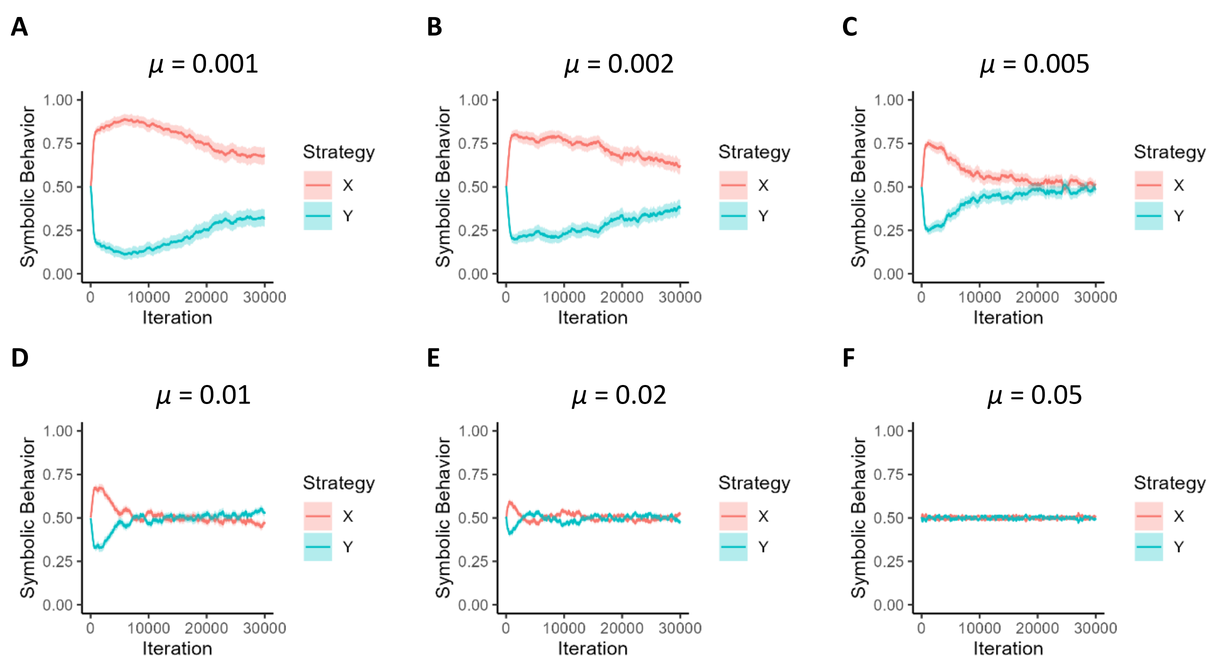
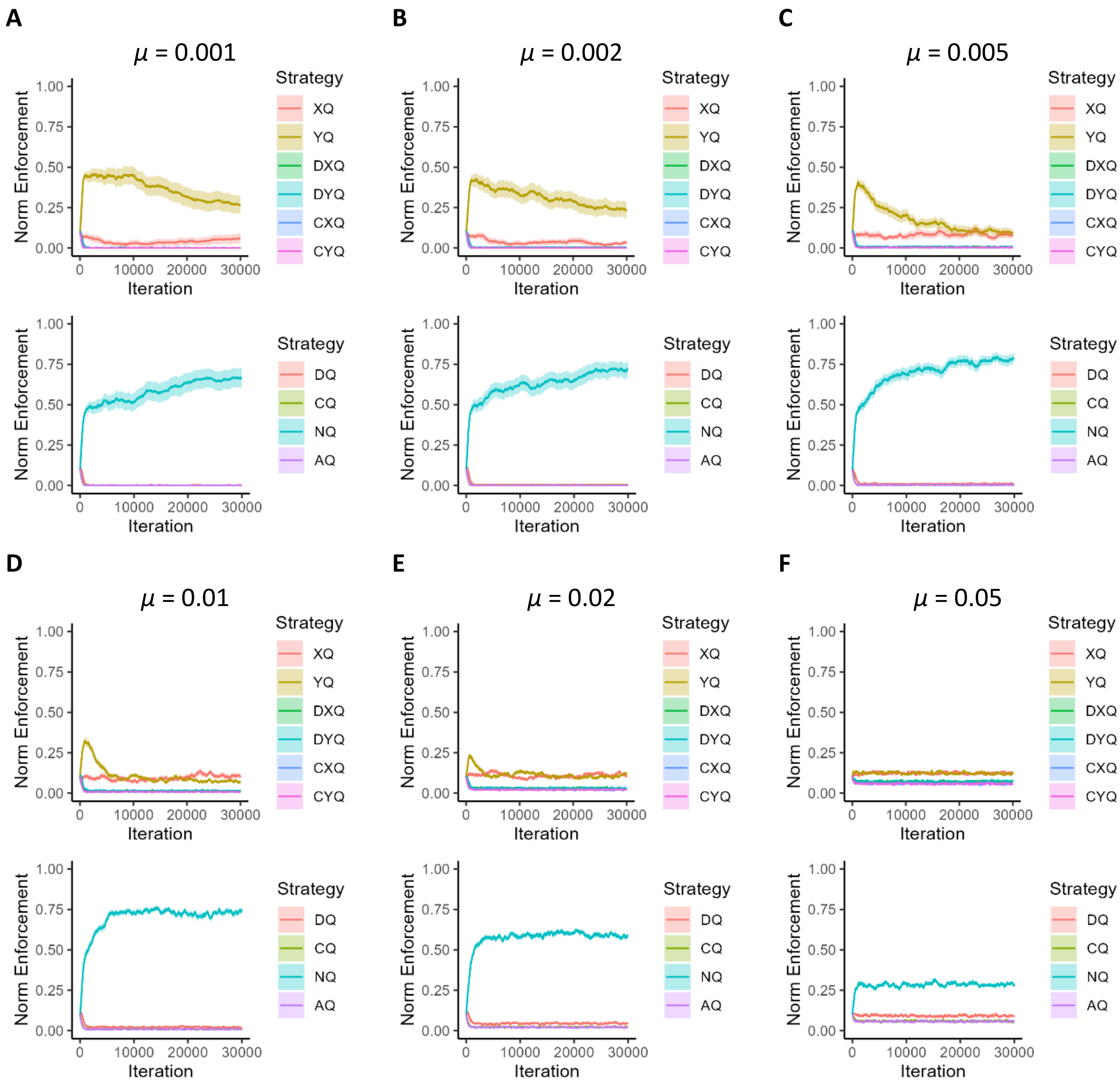


Figure 76 The Evolution of Norm Enforcement Under Different Mutation Rates



2.5.2 When Cooperation, Symbolic, and Ostracism Strategies Mutate Together

In the model, each agent's strategy consists of three parts: a cooperation strategy, a symbolic strategy, and an ostracism strategy. In the default model, these three strategies mutate independently. For each strategy, there is a probability μ that this strategy will be randomly selected from all the possible options. The setup in the default model has the following advantage. Because the three strategies mutate independently, the change of one strategy does not indicate the change of other strategies. In this case, it is less effective to use one's symbolic strategy as a cue to infer whether their cooperation strategy has changed. This is a conservative setup that is disadvantageous for the evolution of the symbolic norm because it decreases the signaling efficacy of the symbolic strategies. I chose this conservative setup because if the norm spillover effect still happens under this setup, I can confidently argue that the norm spillover effect is not built in the mutation process.

In the robustness test in this section, I used a less conservative and more commonly used way to implement random mutation. Instead of letting the three strategies mutate independently, three mutations happen together. In the birth phase, when an agent reproduces, with a probability of μ = 0.015, the new agent randomly selects all the three strategies among all the possible combinations. The new exploration rate μ equals to $1 - (1 - 0.005)^3$ = 0.015 to make the new exploration rate match the exploration rate in the default model. 50 simulation runs were run under each level of $\sigma \in$ [0, 0.05, 0.1, 0.15, 0.2, 0.25] using this new mutation method.

Same as the default model, **Figure 77** shows that the symbolic norm evolved when there is an initial correlation between the symbolic and practical behaviors (i.e., $\sigma > 0$). The symbolic norm is stronger when the initial correlation is stronger. **Figure 78** shows that a substantial proportion of agents enforce the symbolic norm *X* by ostracizing Y-performers (i.e., YQs). **Figure 79** shows that the symbolic norm is not only adopted by cooperators, but also by defectors.

The difference between this section and the results in the default model is that the symbolic norm can be maintained for longer when the three strategies mutate together. This further supports the argument that the default model is a more conservative setup that is adverse for the evolution of symbolic norm.

The reason for the difference between this model and the default model is as follows. When $\sigma > 0$, most cooperators in the population are X-performers. Imagine that there are only cooperators in the population and random mutation is the only force that drives evolution. When the three strategies mutate together, if a Y-performer shows up in the population, this Y-performer is very likely to be an agent who has experienced mutation. Thus, it is very likely the Y-performer's cooperation strategy was also randomly assigned and thus, this Y-performer is probably not a cooperator. In this situation, a better strategy is to ostracize any agent who is not an X-performer to avoid the risk of interacting with a mutated defector. This case is different in the default model. When the three strategies mutate independently, even if an X-performer has mutated to a Y-performer, it does not indicate the mutation of their cooperation strategy. Thus, a Y-performer is still more likely to be maintained as a cooperator. In this case, an agent's symbolic strategy is a less reliable signal for their cooperative behavior. Indeed, independent mutation of the three strategies keeps destroying the correlation between the symbolic and practical behaviors until there is no correlation at all.

Previous literature argued that even when the mutation rate is very small, the structure of mutation may change the evolutionary trajectories substantively (García & Traulsen, 2012). The result in this section provides a good example for this literature. Nevertheless, the main findings of this dissertation remain intact. Regardless of the structure of mutation, the symbolic norm evolved when there is an initial correlation between a symbolic and a practically beneficial behavior and agents will evolve to enforce the symbolic norm by ostracizing its violators.

Figure 77 The Evolution of Symbolic Norm When Cooperation, Symbolic, and Ostracism Strategies Mutate Together

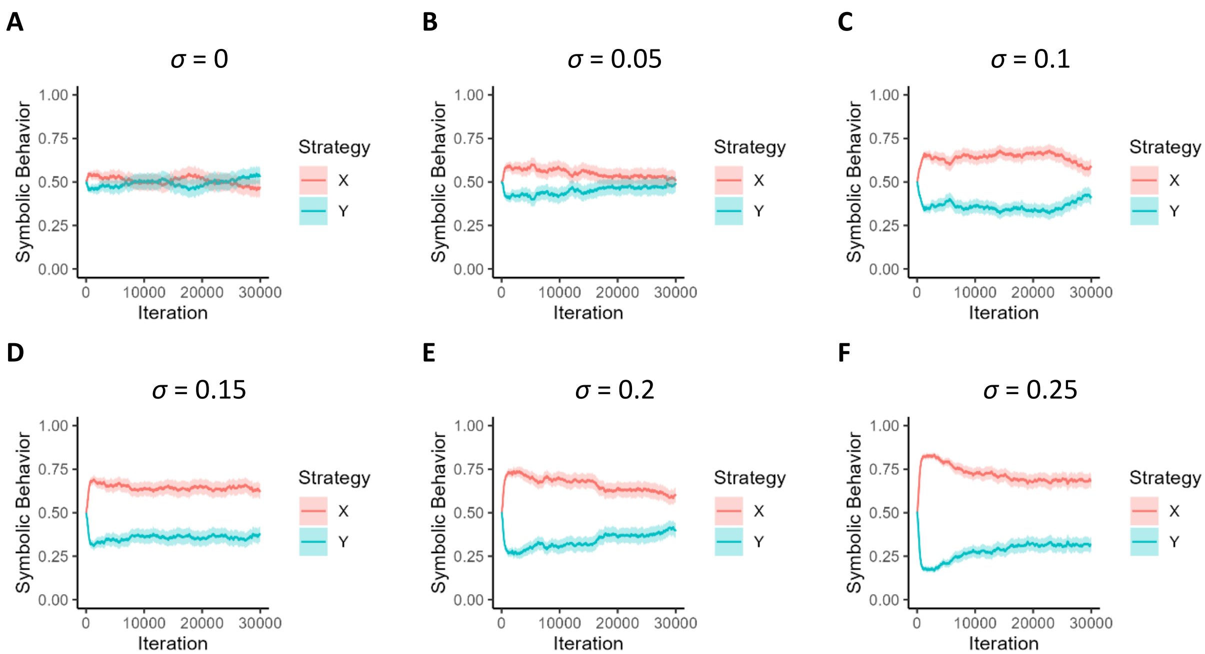


Figure 78 The Evolution of Norm Enforcement When Cooperation, Symbolic, and Ostracism Strategies Mutate Together

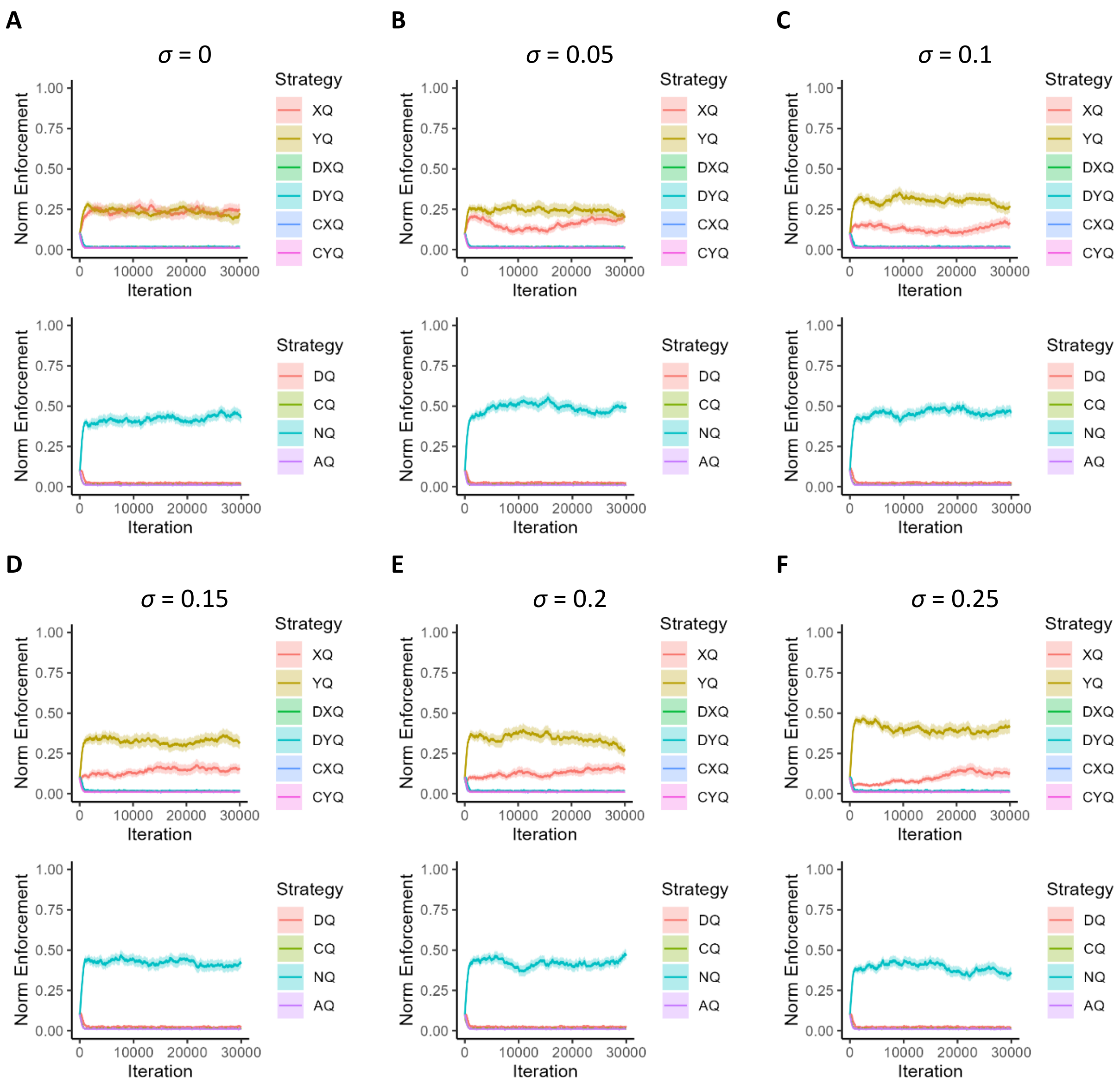
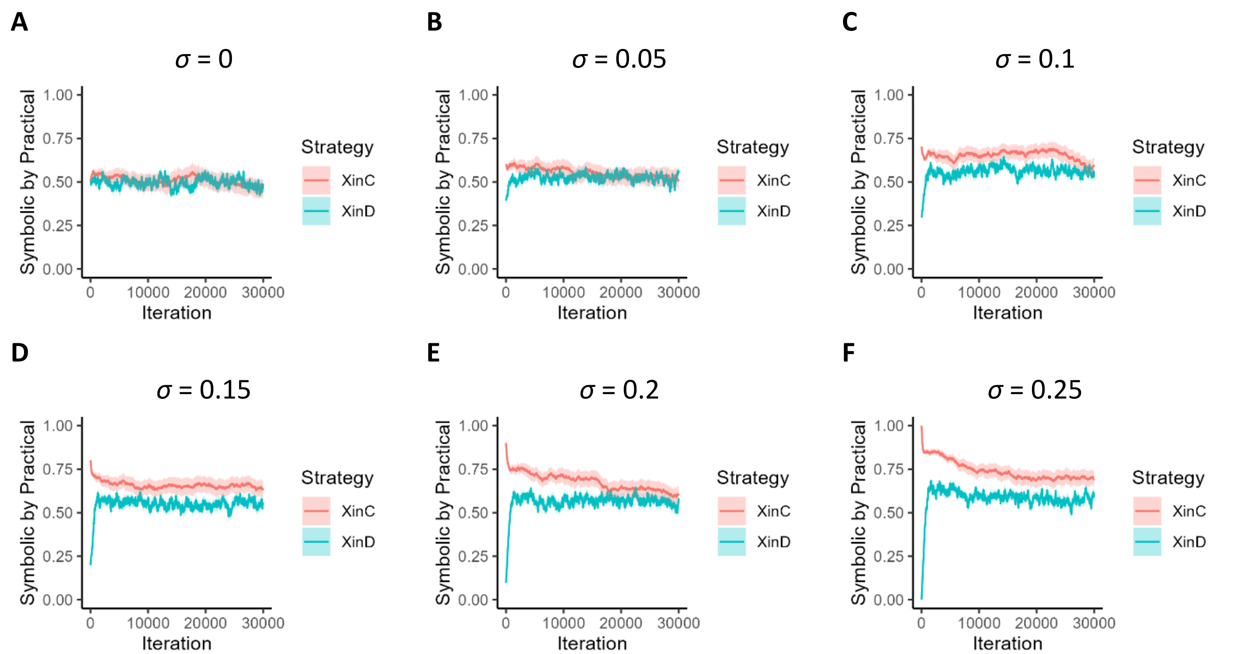


Figure 79 Symbolic Norm Among Cooperators vs. Defectors When Cooperation, Symbolic, and Ostracism Strategies Mutate Together



2.6 Pairwise Comparison as Evolutionary Dynamic

In the default model, the death-birth rule was used as the evolutionary dynamic (Roos et al., 2015). In this section of robustness test, I used pairwise comparison as the evolutionary dynamic and compared it with the death-birth rule.

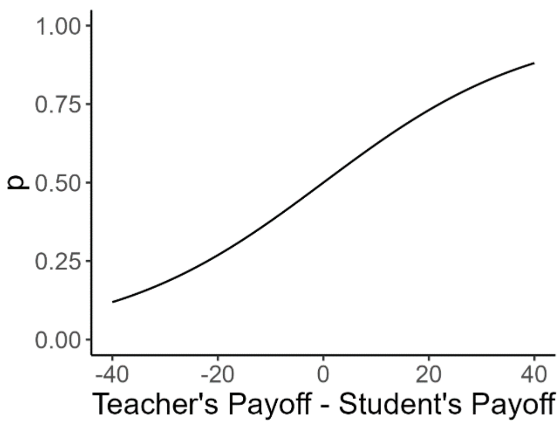
Pairwise comparison updating mimics the social learning process in cultural evolution (Mesoudi, 2011). Empirical research has found that people tend to copy the behavior of others who have higher payoffs. Resembling this process, in each iteration, a proportion of *updF* agents are randomly selected without replacement as "students" who are given a chance to update their strategies. *updF* indicates the frequency of strategy updating or the speed of evolution. *updF* is set at 0.1 to be comparable as the death rate δ in the original model.

To explain how the agents update their strategies, we may consider each agent as a "student" who randomly selects another agent as their "teacher." If agents are connected by the social network, the “teacher” is randomly selected from all the neighbors connected to the “student.” With a probability of p, the “student” will adopt the cooperation, symbolic, and ostracism strategies of the “teacher.”

The magnitude of p is decided by the payoff of the “student,” π_s, and the payoff of the “teacher,” π_t, as shown in Equation (3) and **Figure *80*** (Roca et al., 2009; Traulsen et al., 2007). In general, the higher payoff the “teacher” has compared to the “student,” the more likely the “student” will adopt the “teacher’s” strategies. The parameter *s* = 0.05 represents the strength of selection. An agent’s payoff is calculated as the sum payoff from the current iteration.

$$p = \frac{1}{1+e^{-s\times(\pi_t-\pi_s)}} \quad (3)$$

Figure 80 Probability of Strategy Change in Fermi Rule



Note. The probability of a “student” changing strategy as a function of the payoff difference between the “teacher” and the “student.” Selection strength *s* = 0.05 in this plot.

In addition to the payoff-based strategy updating, for each student, for each of the student’s strategies, there is a small probability $\mu = 0.005$ that this strategy will be randomly selected from all the possible options.

Moreover, two versions of models were implemented in this section. In the first version, agents *keep* their memories even if they change their strategies. This is the normal way of implementing pairwise comparison updating. Strategy updating in this version represents the social learning process within the same individual. In other words, an individual can learn a new strategy but maintain their previous memories. In the second version, however, a variation of the model was used: If an agent is selected as the "student," the agent will be considered as a completely new agent. In this case, this agent will reset all their memories for others and others will also reset their memories for this agent as if they never interacted before. This setup is untypical and may be counterintuitive, but the purpose is to make this version of pairwise comparison comparable to the death-birth rule. In death-birth, after an agent "dies," all memories are erased. Therefore, by comparing the two versions of pairwise comparison (i.e., memory keeping vs. memory reset), I examined whether the different results given by death-birth vs. pairwise comparison are due to the difference in memory.

2.6.1 Pairwise Comparison While Keeping Students' Memories

Same as in the default model, **Figure 81** and **Figure 82** show that the symbolic norm and norm enforcement first evolved and then disappeared when there was an initial correlation between the symbolic and practical behaviors. As the initial correlation increases, the symbolic norm becomes stronger. The difference between this robustness test and the default model is that the symbolic norm disappears faster under pairwise comparison updating. This is probably because the selection strength in pairwise comparison is higher. Nevertheless, the main findings regarding the norm spillover effect remain robust.

Figure 81 The Evolution of Symbolic Norm Using Pairwise Comparison While Keeping Students' Memories

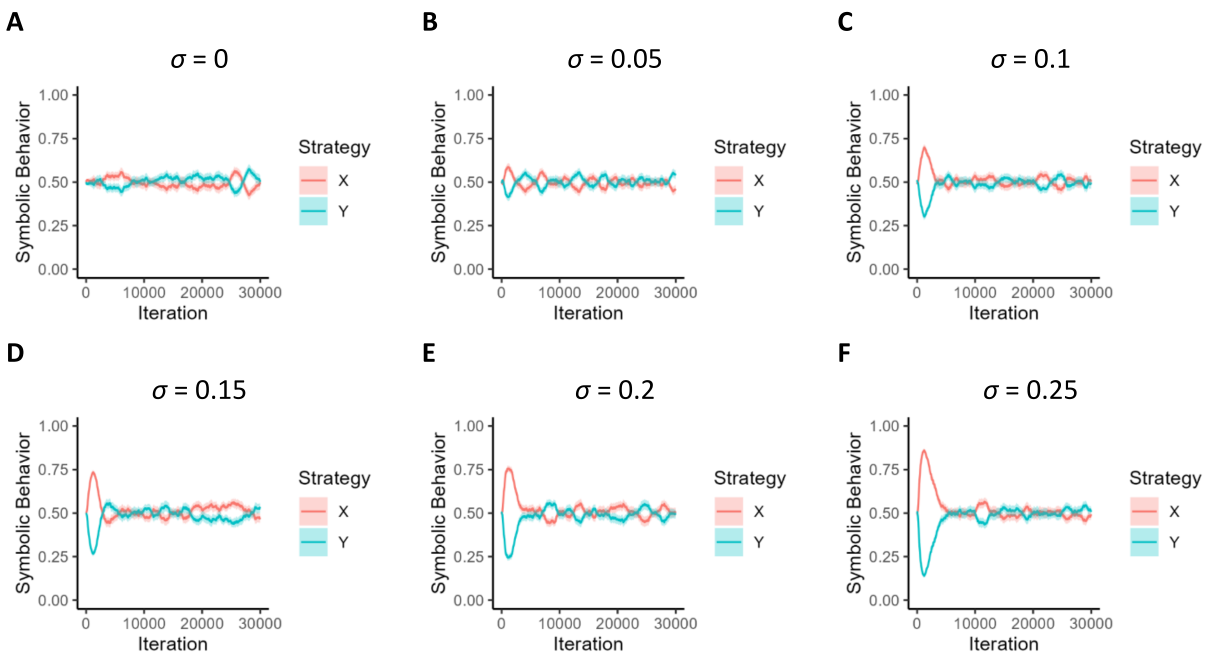
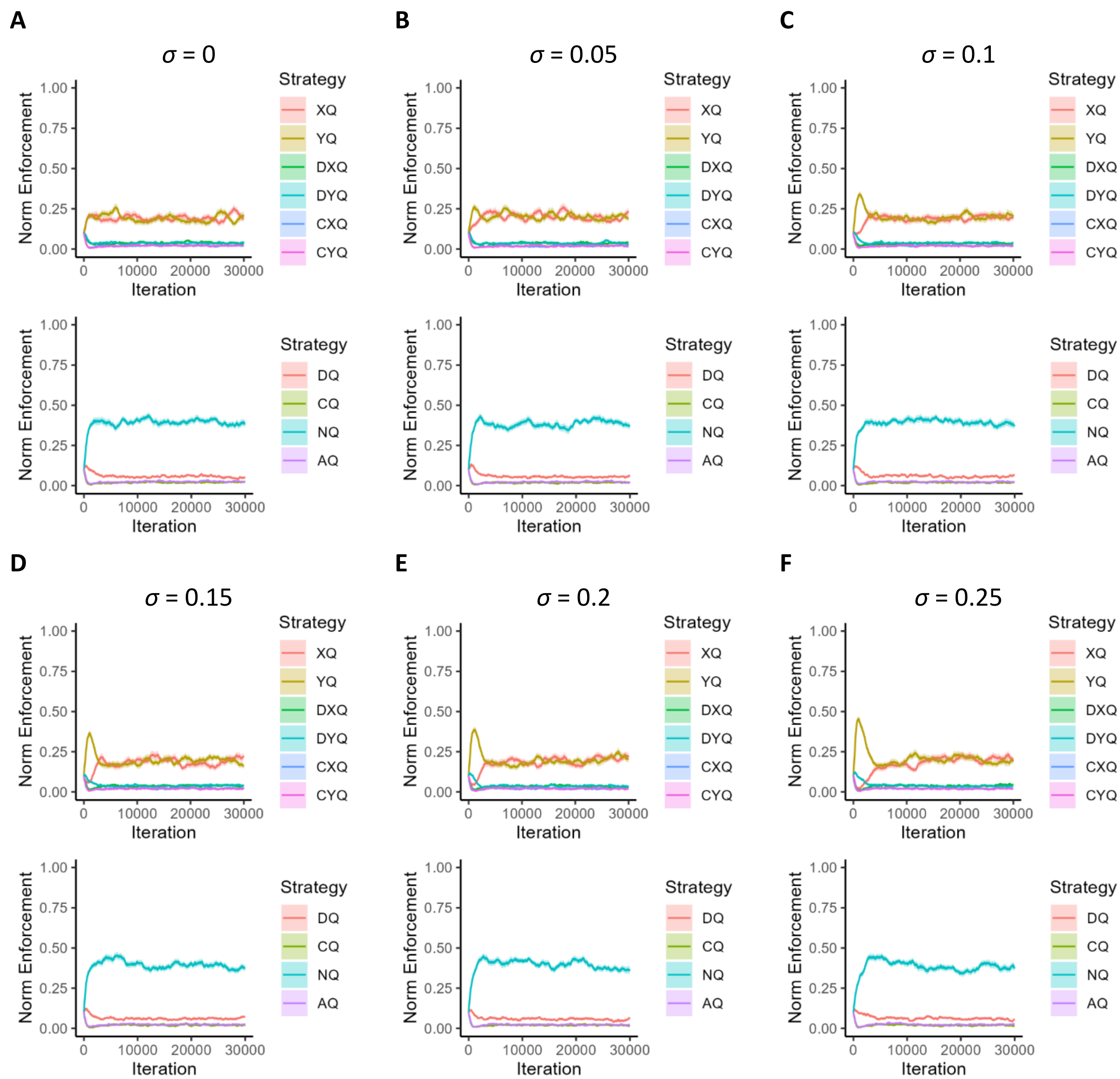


Figure 82 The Evolution of Norm Enforcement Using Pairwise Comparison While Keeping Students' Memories



2.6.2 Pairwise Comparison While Resetting Students' Memories

The pattern of the results remains the same when students are made to reset their memories. **Figure 83** and **Figure 84** show that the symbolic norm and norm enforcement first evolved and then disappeared when there was an initial correlation between the symbolic and practical behaviors. As the initial correlation increases, the symbolic norm becomes stronger. This section supports the robustness of the findings regarding the norm spillover effect.

Figure 83 The Evolution of Symbolic Norm Using Pairwise Comparison While Resetting Students' Memories

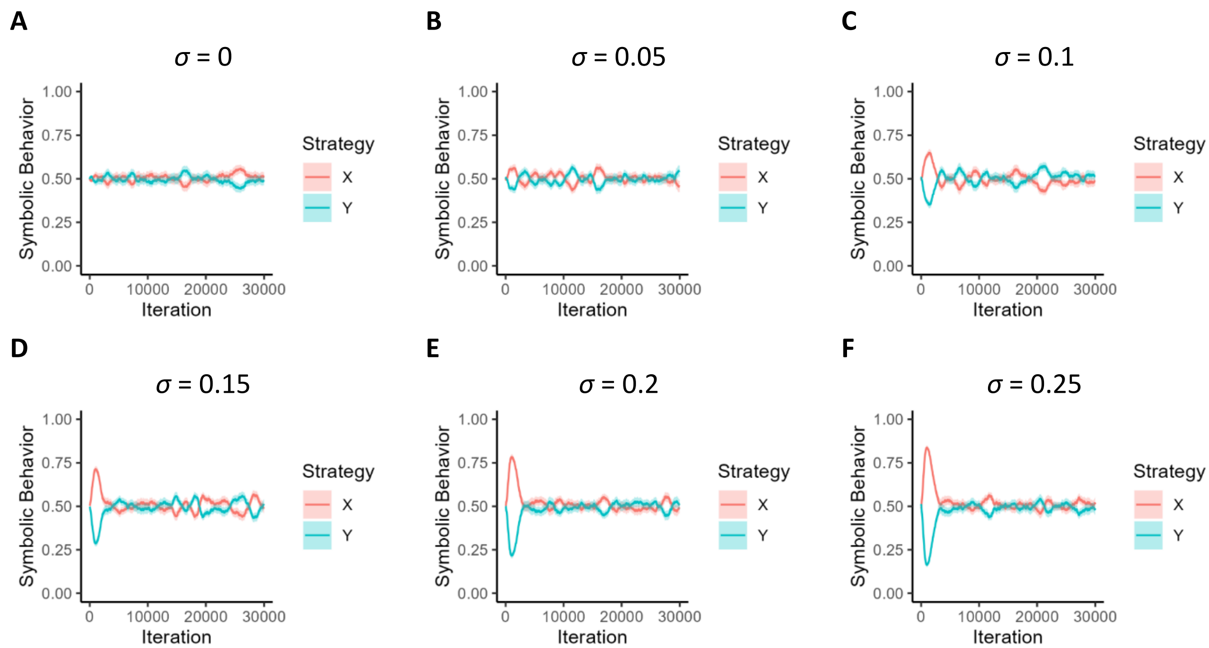
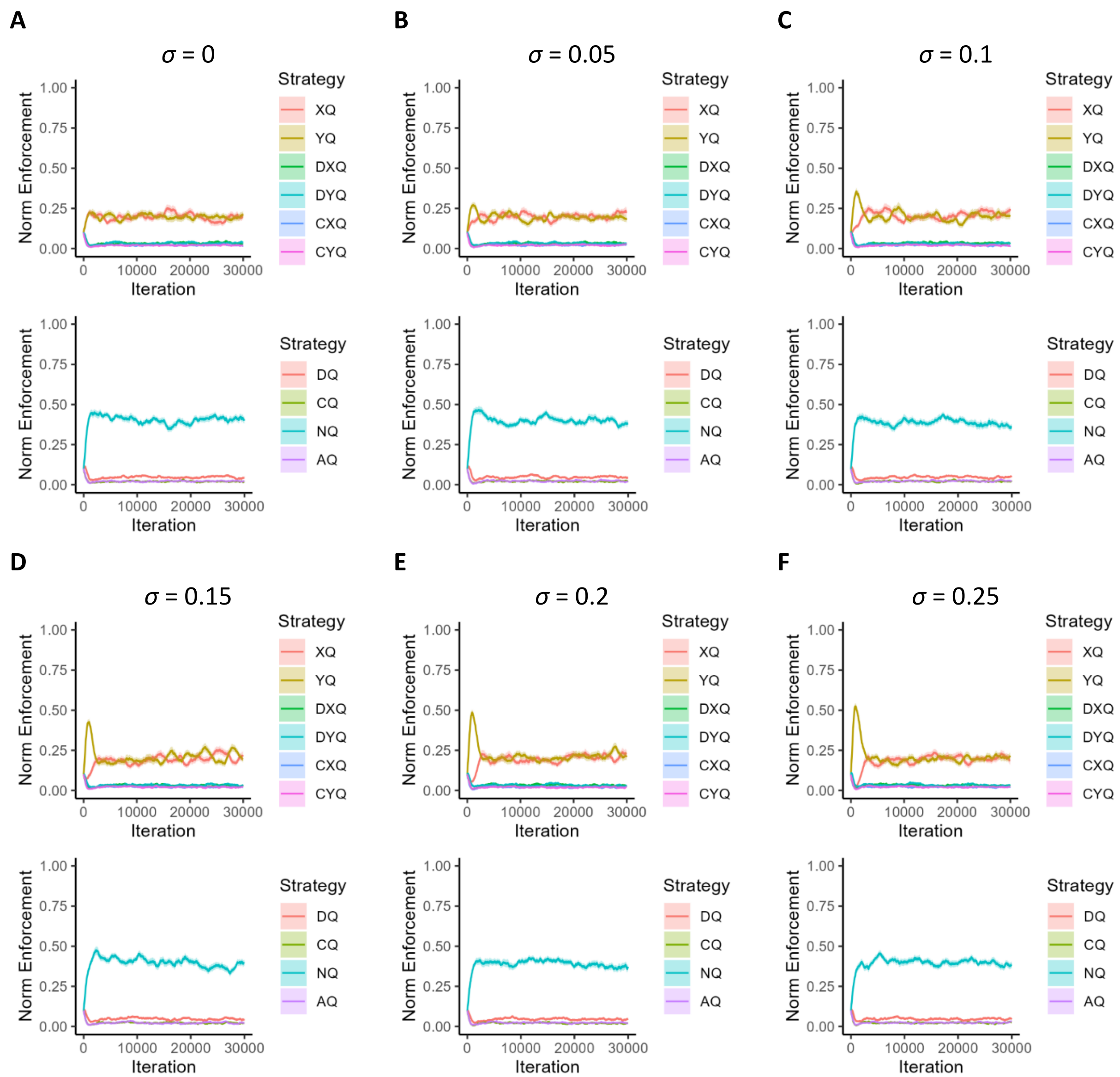


Figure 84 The Evolution of Norm Enforcement Using Pairwise Comparison While Resetting Students' Memories



2.7 Network Structure

In the default model, agents were embedded in a static grid network. In the robustness tests in this section, I used many different network structures to examine the robustness of the results. Specifically, I ran the same simulations on random regular networks with different levels of connectivity, on a small-world network with different levels of connectivity, and also among a well-mixed population.

2.7.1 Random Regular Network

In this section, a series of random regular networks were used to test the robustness of the norm spillover effect. The degree of the random regular network was manipulated. 50 simulation runs were run under each level of *degree* ∈ [4, 5, 10, 15, 20, 25, 30, 35] under $\sigma = 0.15$. When the network degree is higher, if each agent still interacts with all their neighbors in each iteration, agents will have more interactions per iteration. To match the interaction frequency across different network degrees, I randomly selected only a proportion of connected pairs to interact and make the total number of interactions equal to the interactions in the default model.

Figure 85 and **Figure 86** show that the symbolic norm and norm enforcement evolved under all the conditions, supporting the norm spillover effect. As the degree of the network increases, the symbolic norm becomes weaker. One reason may be that as the degree of the network increases, the cooperation rate decreases (see **Figure 87**) (Ohtsuki et al., 2006). As mentioned in Section ***The Impact of Practical Behavior on the Norm Spillover Effect***, the symbolic norm is weaker when fewer agents follow the practical norm. Thus, the symbolic norm is weaker when the network degree is higher. Nevertheless, even when defectors dominant the population (e.g., *degree* > 10), some symbolic norm still evolved. This further suggests that the symbolic choice can emerge as a norm even if the practical behavior correlated to it does not become the norm, a result that is consistent with Section ***The Impact of Practical Behavior on the Norm Spillover Effect***.

Figure 85 The Evolution of Symbolic Norm on Random Regular Networks

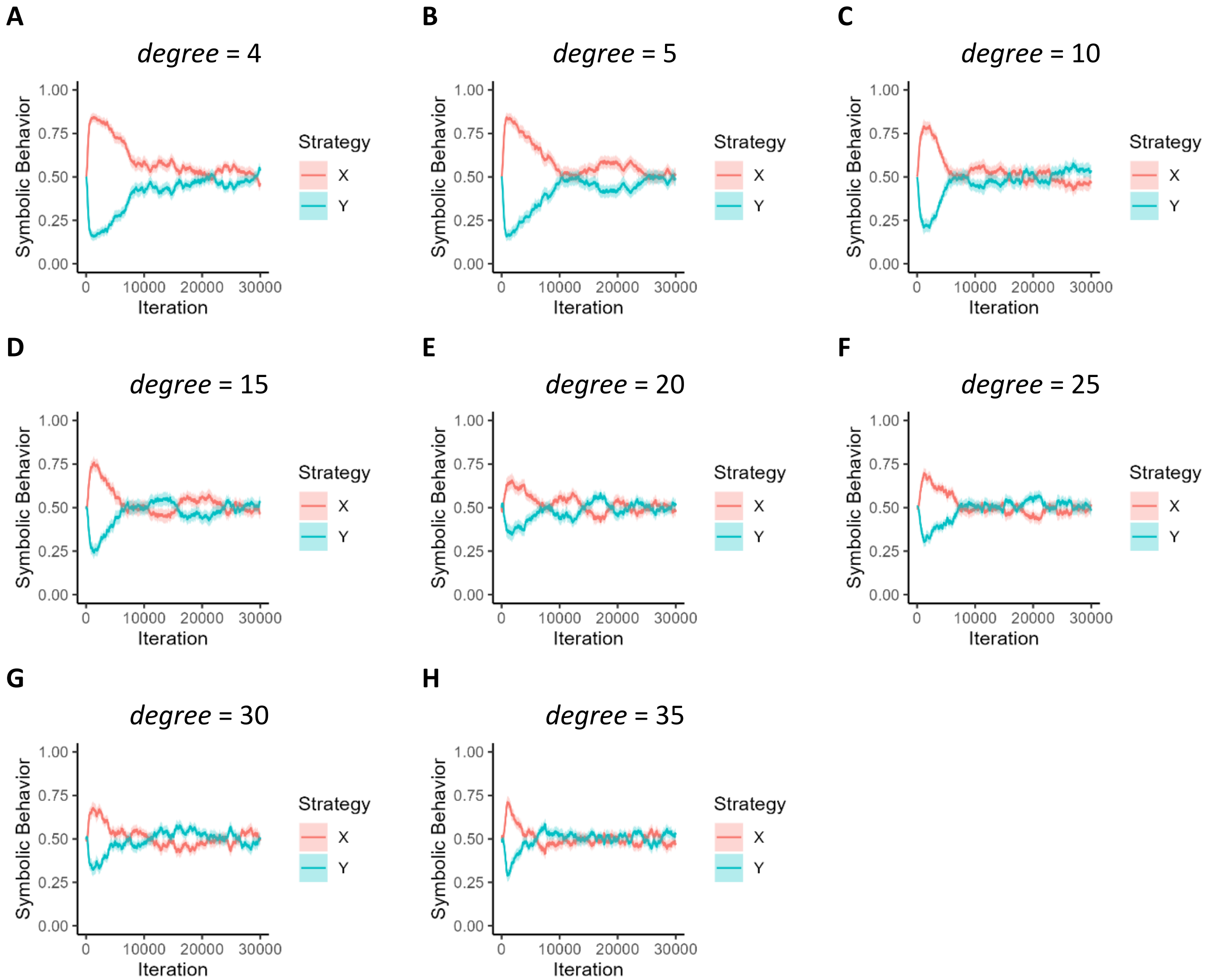


Figure 86 The Evolution of Norm Enforcement on Random Regular Networks

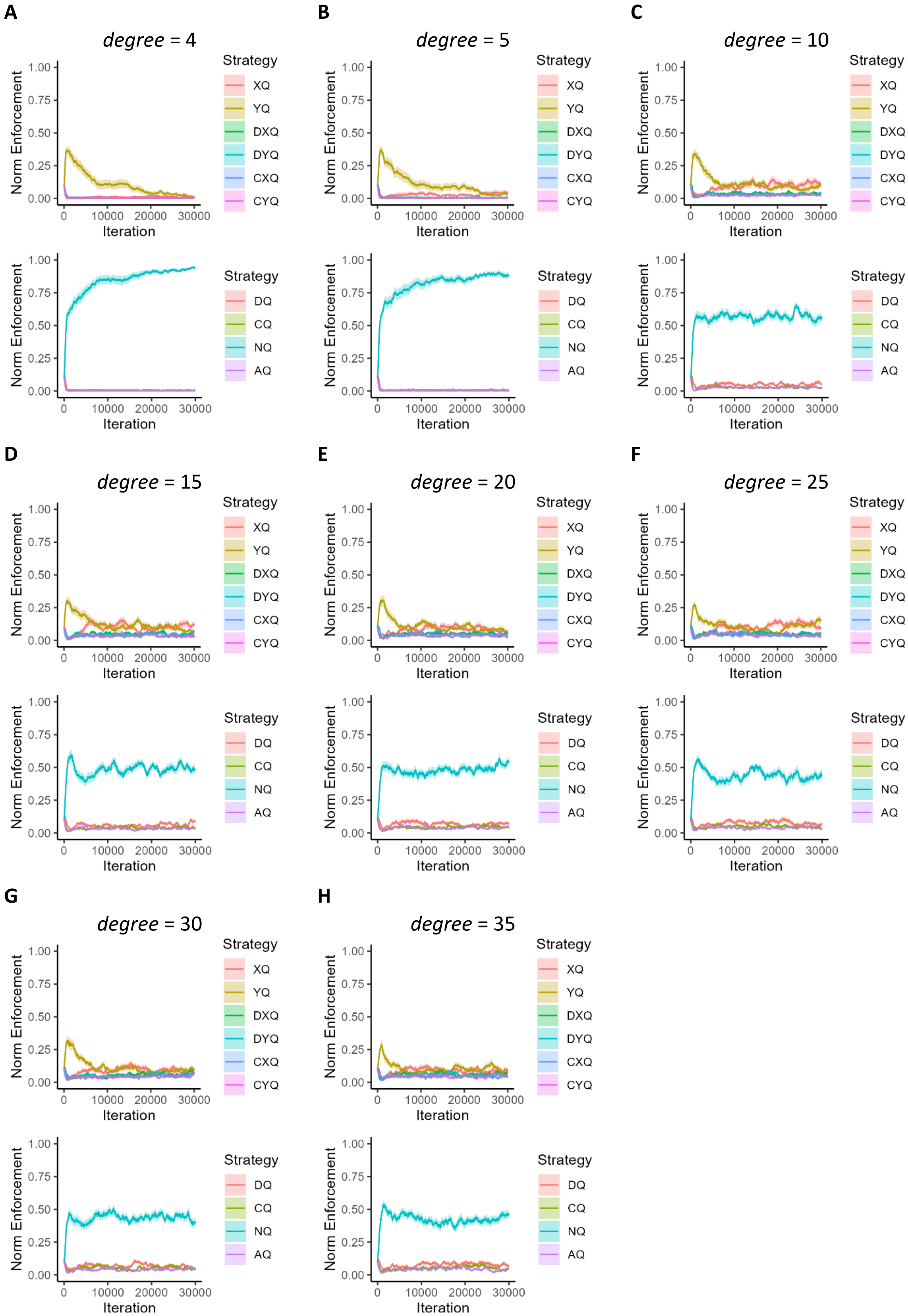
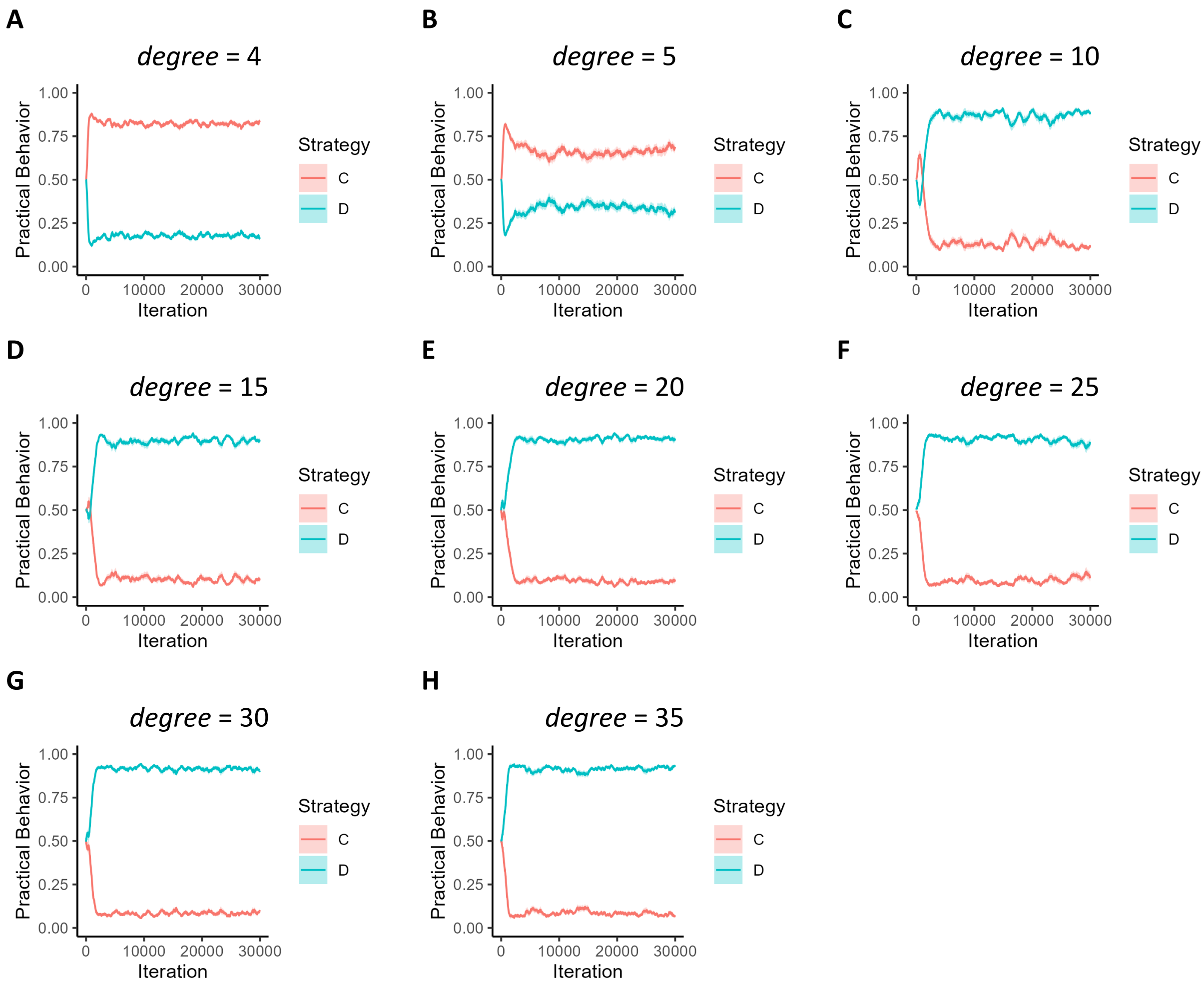


Figure 87 The Evolution of Cooperation on Random Regular Networks



2.7.2 Small-world Network

In this section, a series of small-world networks were used to test the robustness of the norm spillover effect. Small world networks have been widely used to resemble connections in real-world (Milgram, 1967; Weeden & Cornwell, 2020). The small-world network is generalized with the algorithm of Watts and Strogatz (D. J. Watts & Strogatz, 1998) using the “connected_watts_strogatz_graph()” function in NetworkX with Python. This algorithm first creates a ring over $N = 1024$ nodes. Then each node in the ring is connected with its *swK* nearest neighbors. The parameter *swK* controls the average degree of the small-world network. Then shortcuts are created by replacing some edges as follows: For each edge u-v in the network, with a probability *swP* = 0.1, replace it with a new edge u-w with uniformly random choice of existing node w (Hagberg et al., 2008). The algorithm attempts to generate a *connected* graph by repeated generation of Watts–Strogatz small-world graphs. It will stop after 100 tries.

The average degree of the small-world network *swK* was manipulated. 50 simulation runs were run under each level of degree ∈ [4, 5, 10, 15, 20, 25, 30, 35] under $\sigma = 0.15$. Same as the random regular networks, to match the interaction frequency across different network degrees, I randomly selected only a proportion of connected pairs to interact and make the total number of interactions equal to the interactions in the default model.

Figure 88 and **Figure 89** show that across all the conditions, the symbolic norm and norm enforcement evolved, supporting the norm spillover effect. However, there is no clear monotonic effect of how the average degree of a small-world network influences the strength of the symbolic norm. Compared with a random regular network, the strength of the symbolic norm is slightly higher in a small-world network under the same network degree. This result may indicate that a small-world network is more favorable for the evolution of symbolic norm than a random regular network, although the mechanism behind this difference requires further exploration.

Figure 88 The Evolution of Symbolic Norm on Small-world Networks

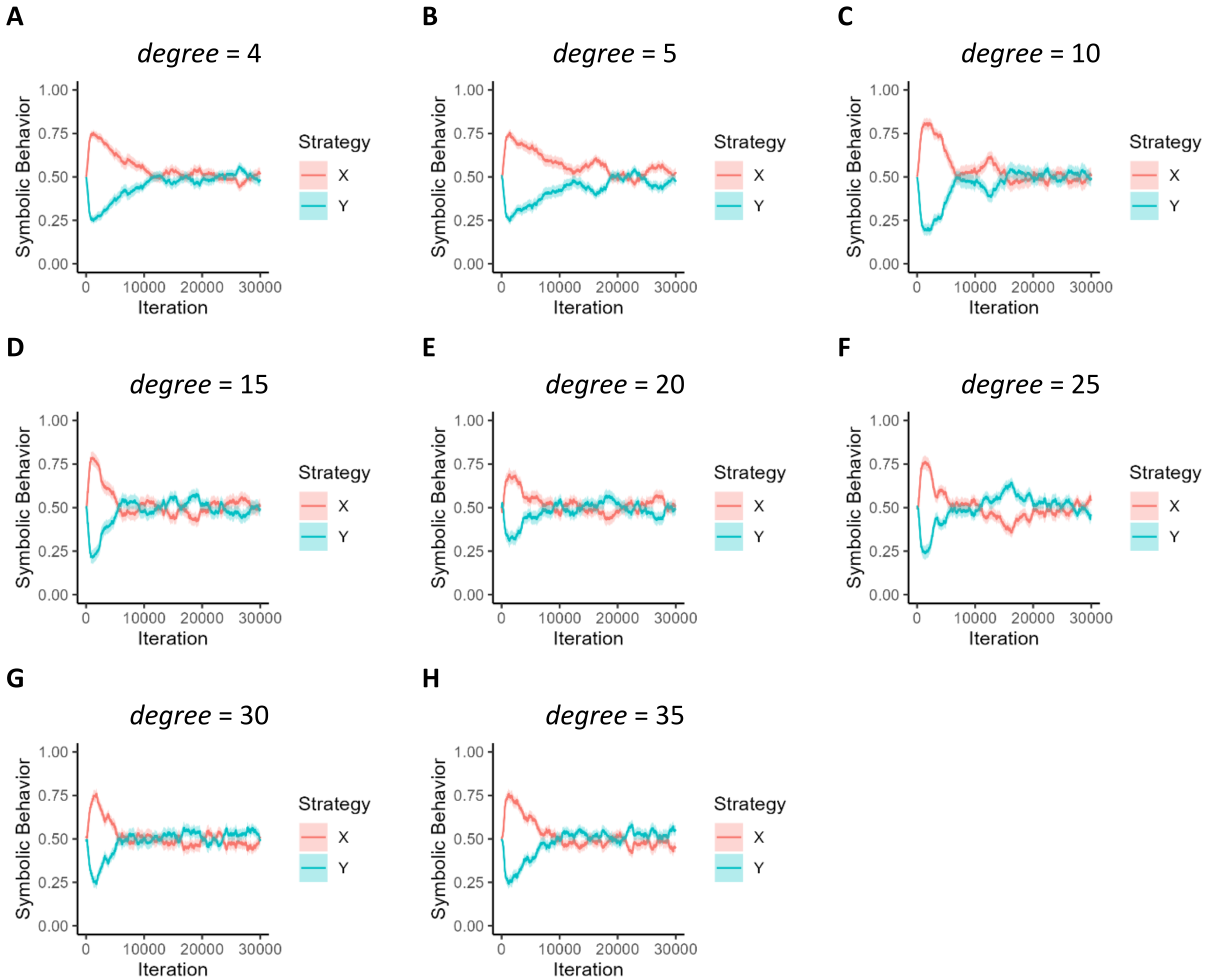
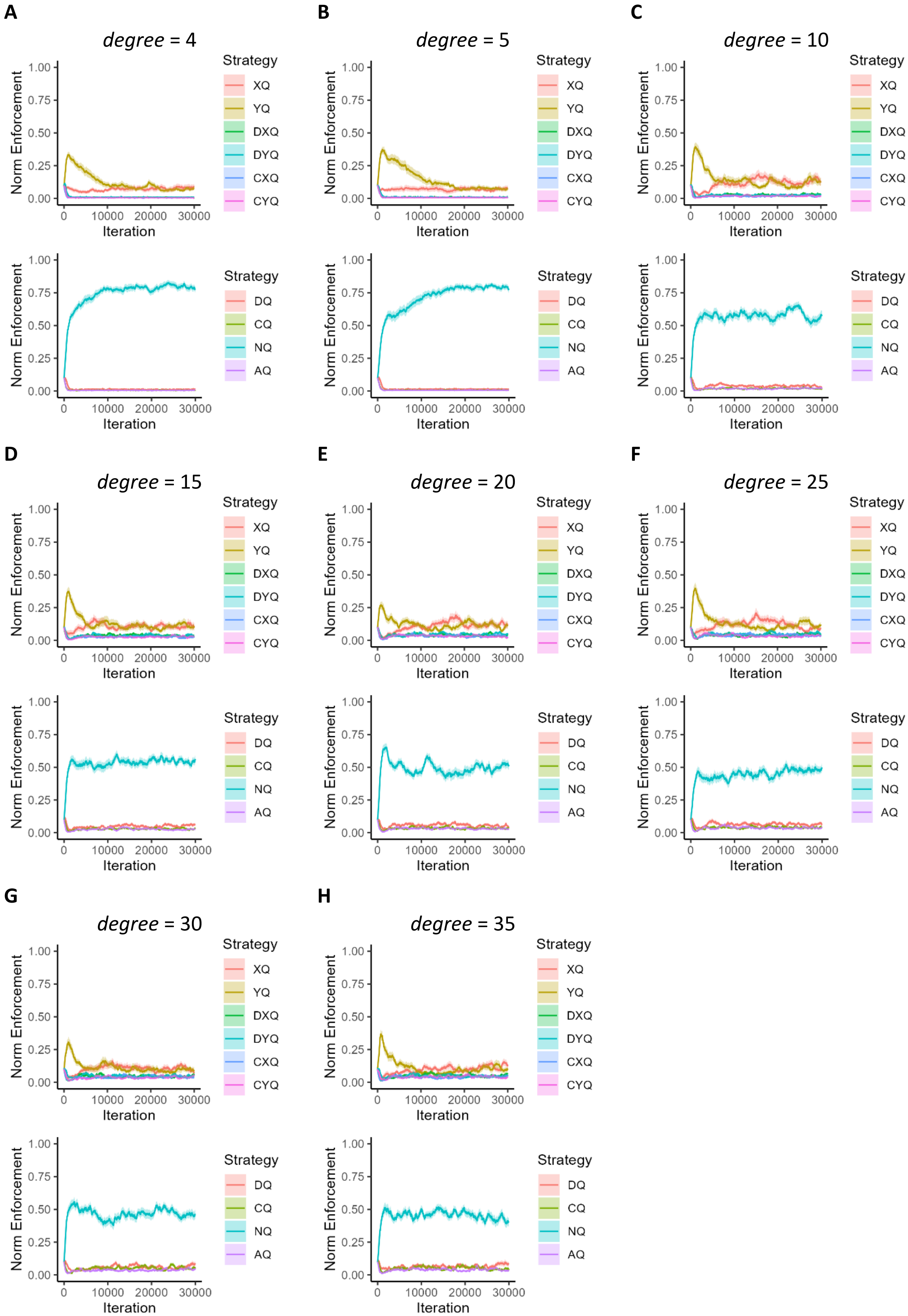


Figure 89 The Evolution of Norm Enforcement on Small-world Networks



2.7.3 Well-mixed Population

In this section, I reran the simulations in a well-mixed population. A well-mixed network can be considered as a population where everyone is connected. Well-mixed populations have been widely used in mathematical models because of its mathematical tractability. Though a mathematical model was not used in this dissertation, I tested the model in a well-mixed population to allow for future studies to compare this model with a mathematical model.

A model with a well-mixed population differs from a grid-networked population in the following ways. First, at the beginning of each iteration, agents will be randomly reassigned into pairs for the interaction phase. The reassignment and the interaction phase happen four times in order to make the total number of interactions equal to that in the default model. Second, in the death-birth rule, when an agent "dies," it will be removed from and leave an empty spot in the population. When an agent "reproduces," they will create an offspring to any available empty spot in the population until the population is fully filled. 50 simulation runs were run under each level of $\sigma \in [0, 0.05, 0.1, 0.15, 0.2, 0.25]$.

Figure 90 and **Figure 91** show that the symbolic norm and norm enforcement evolved when the initial correlation between the symbolic and practical behaviors are strong enough, supporting the norm spillover effect. Compared with populations embedded in a social network, the strength of the symbolic norm is weaker in a well-mixed population. This is consistent with the findings in Supplementary Section **2.7.1 Random Regular Network** that the strength of the symbolic norm is weaker when the degree of the network is higher. Since a well-mixed population is equivalent to a population with full connection (i.e., the highest degree), it is reasonable that the symbolic norm is the weakest in a well-mixed population.

Notably, **Figure 92** shows that almost all agents become defectors in a well-mixed population, which is consistent with previous literature on the evolution of cooperation on graphs (Ohtsuki et al., 2006). Nevertheless, the symbolic norm still evolved when σ is high enough (i.e., $\sigma > 0.15$). This further supports the arguments in Section ***The Impact of Practical Behavior on the Norm Spillover Effect*** that the evolution of the symbolic norm does not require the evolution of the practical norm.

Figure 90 The Evolution of Symbolic Norm in Well-mixed Populations

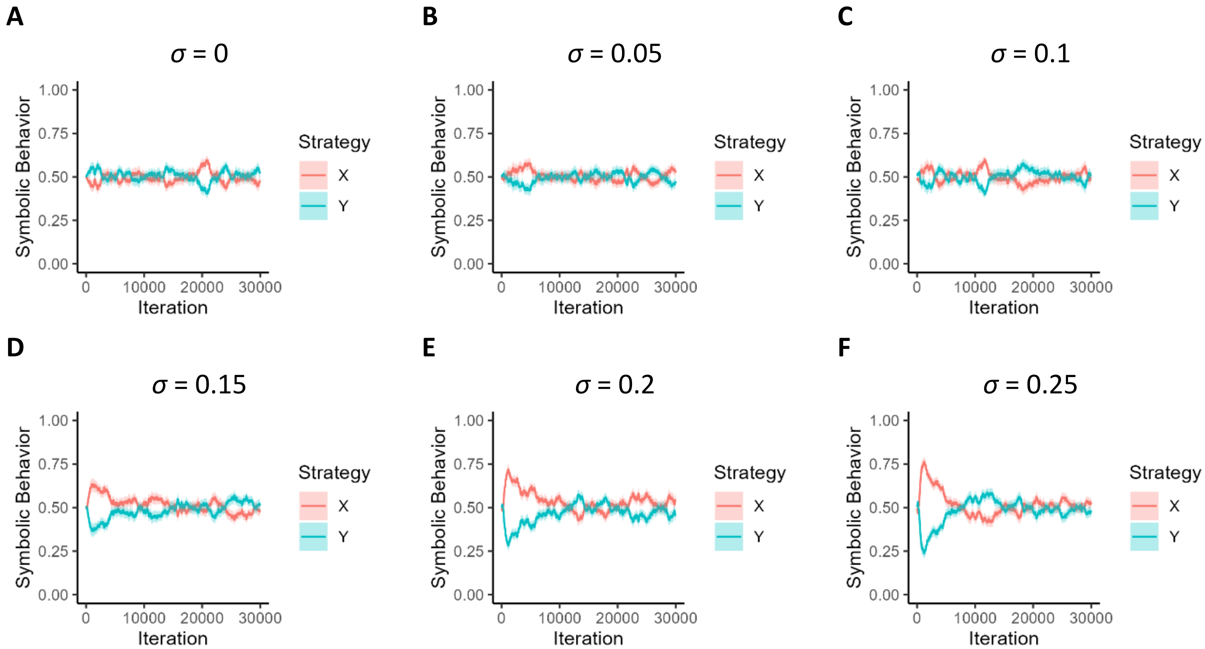


Figure 91 The Evolution of Norm Enforcement in Well-mixed Populations

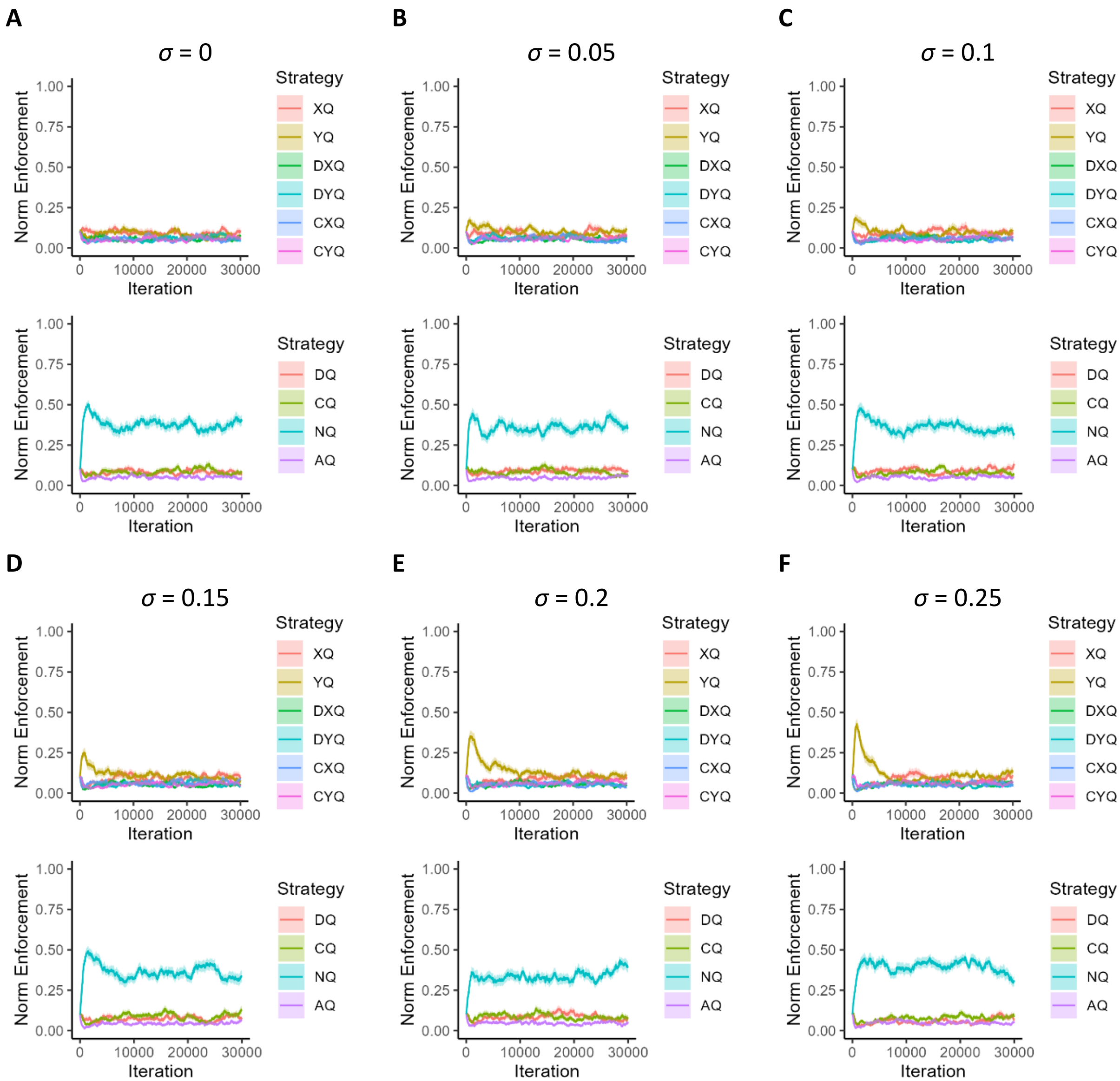
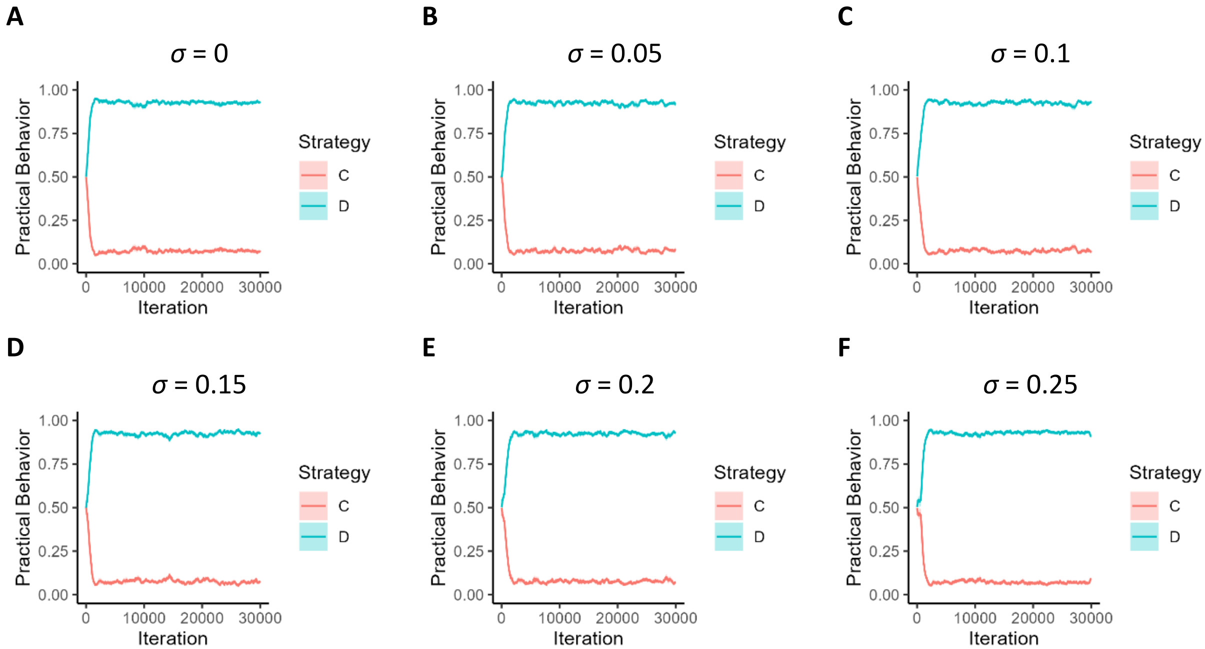


Figure 92 The Evolution of Cooperation in Well-mixed Populations



2.8 Network Mobility

In the default model, the network structure is fixed. In other words, the mobility of the population $m = 0$. In this robustness test, I manipulated the mobility of the network. In each iteration, after the interaction phase but before the evolutionary dynamic, a proportion of agents are selected to move. The proportion m represents the magnitude of mobility. If an agent is selected to move, the agent randomly selects a destination location among all the locations in the network. If there is an alive agent in that destination location, the moving agent exchanges its location with the agent in the destination. If the destination is empty, the moving agent moves to the destination and leaves an empty spot at the original location. In the simulation code, when an agent randomly selects the destination location, it is possible that the agent selects the original location and thus does not move. However, the probability of selecting the original location as the destination is only 1/1024. Thus, this special situation is neglectable in our discussion. When an agent moves, they keep their memories, and everyone keeps their memories of the moving agent. 50 simulation runs were run under each level of mobility $m \in [0.0005, 0.001, 0.002, 0.004, 0.005, 0.006, 0.008, 0.01, 0.015, 0.02, 0.05, 0.1, 0.15, 0.2]$ under $\sigma = 0.15$.

Figure 93 shows that the symbolic norm and norm enforcement evolved under all levels of network mobility, supporting the norm spillover effect. Overall, as network mobility increases, the symbolic norm becomes weaker. Previous models showed that it is harder to maintain practical norms under high social network mobility (De et al., 2015) (see **Figure 96**). The current model echoes this result and shows that it is also harder to maintain symbolic norms under high mobility.

Notably, the relationship between mobility and the strength of the symbolic norm might be non-monotonic. When mobility is low, there is a slight trend that the symbolic norm becomes stronger as mobility increases (see **Figure 94**). This is probably because when there is some mobility, agents interact with new partners occasionally. When interacting with a stranger, it is more helpful to observe their symbolic strategy and infer their cooperation strategy, compared with interacting with an old neighbor. In this case, relying on the symbolic norm is more effective when there is some mobility. However, when mobility is too high, agents interact with strangers too frequently. In this case, rather than inferring other's cooperation strategy, a better strategy is to just become a defector and keep exploiting new agents. Thus, the cooperation rate becomes very low under high mobility and agents no longer rely on the symbolic strategy to guide their cooperation decisions. The non-monotonic relationship between mobility and symbolic norm needs further exploration, which will not be discussed in this dissertation.

As for the enforcement of the symbolic norm, the non-monotonic pattern is not consistent. In general, the enforcement of the symbolic norm is weaker as mobility increase (see **Figure 95**). Overall, the results in this section show that the norm spillover effect is robust across different mobility levels.

Figure 93 The Evolution of Symbolic Norm Under Different Levels of Network Mobility

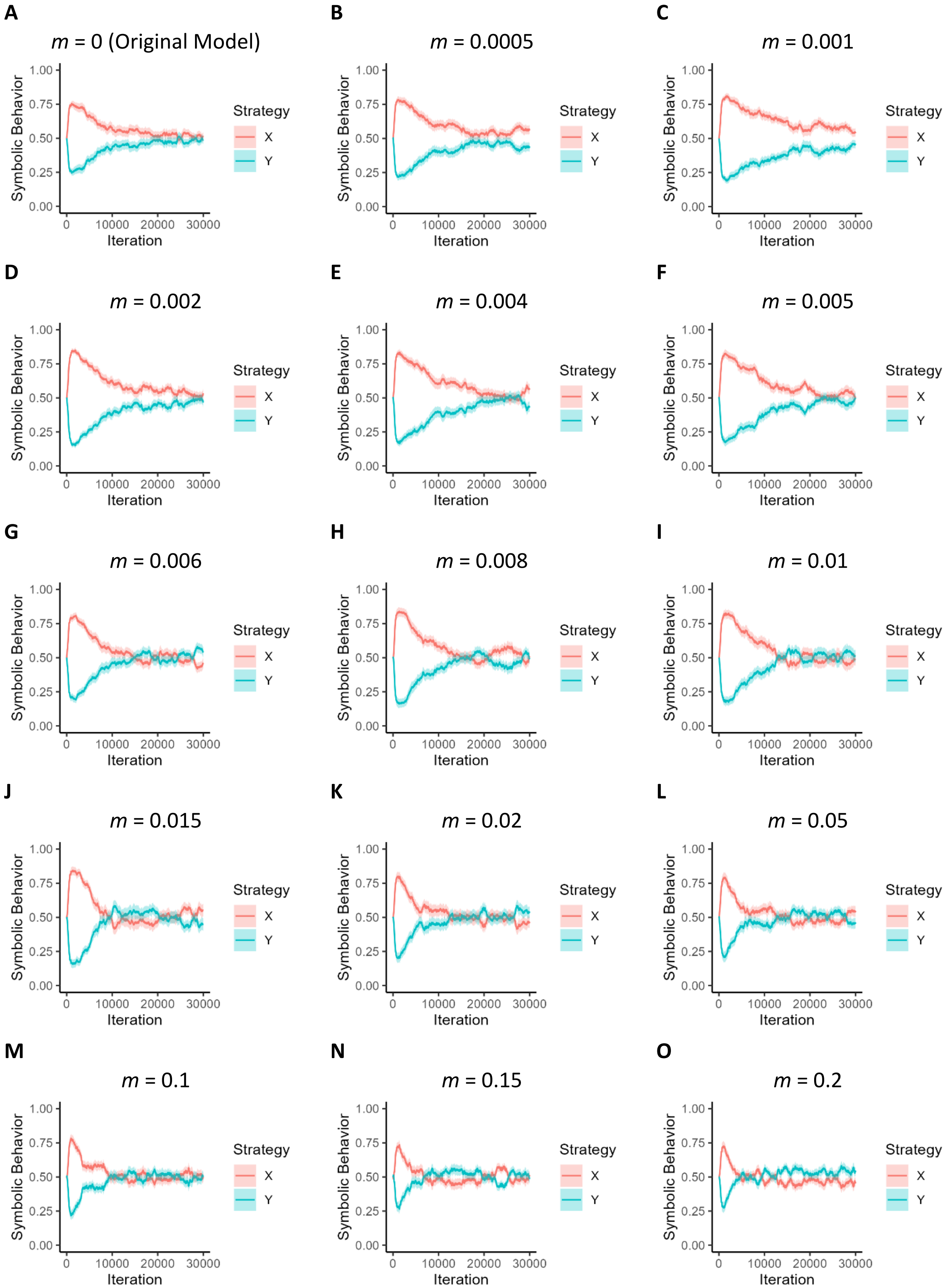


Figure 94 The Potential Non-monotonic Relationship Between Mobility and Symbolic Norm

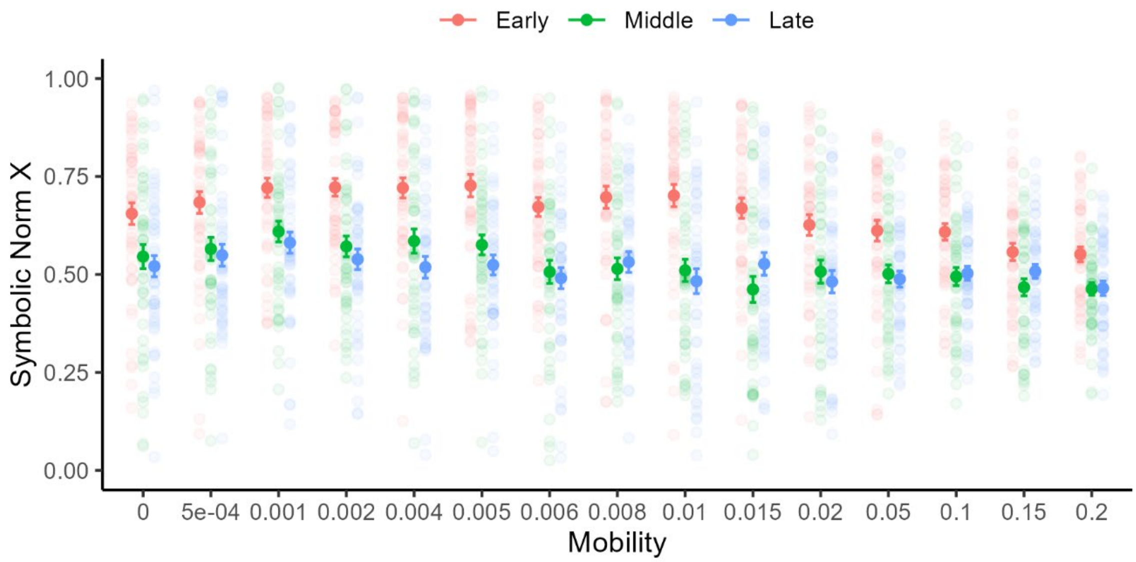


Figure 95 The Evolution of Norm Enforcement Under Different Levels of Network Mobility

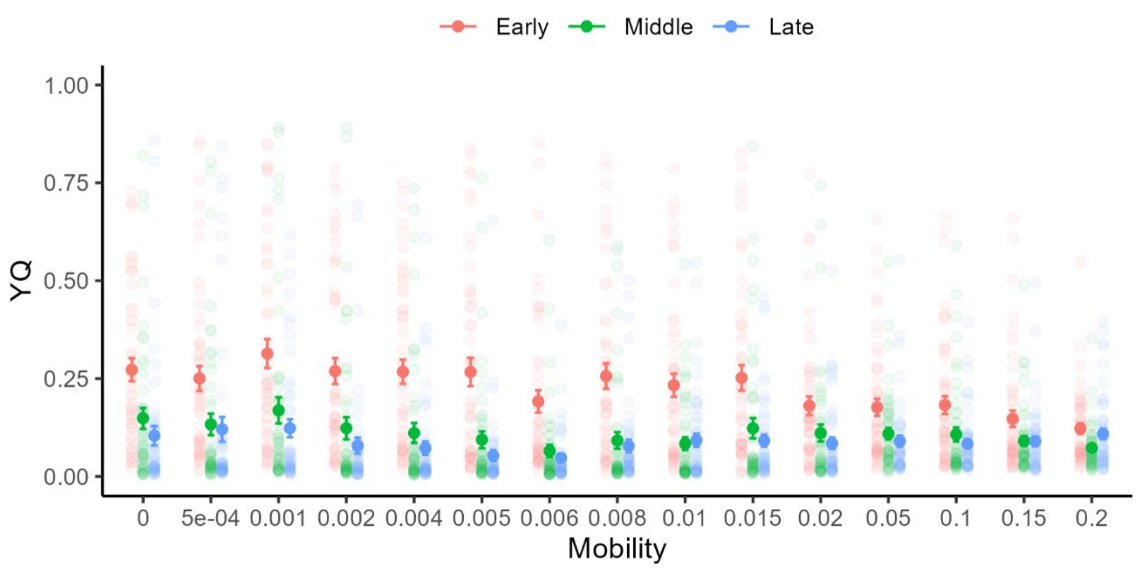
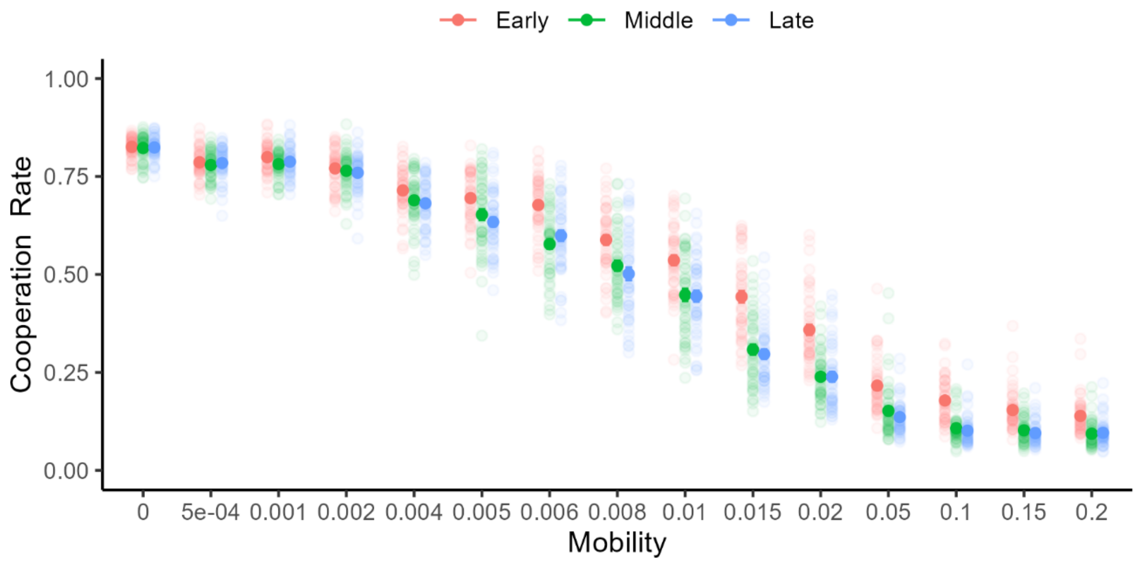


Figure 96 The Evolution of Cooperation Under Different Levels of Network Mobility



2.9 Norm Enforcement by Punishment

In the default model, agents enforce a behavior by avoiding interacting with its violators. In this robustness test, I examined another model setup where agents enforce a norm by punishing its violators.

2.9.1 Methods

There are ten options of *punishment strategies*. Specifically, they are 1) *NP*, who never punishes any agents, 2) *CP*, who punishes only cooperators, 3) *DP*, who punishes only defectors, 4) *XP*, who punishes only the agents who perform *X* in the symbolic domain, 5) *YP*, who punishes only the agents who perform *Y* in the symbolic domain, 6) *CXP*, who punishes both cooperators and X-performers, 7) *CYP*, who punishes both cooperators and Y-performers, 8) *DXP*, who punishes both defectors and X-performers, 9) *DYP*, who punishes both defectors and Y-performers, and 10) *AP*, who punishes everyone (see **Table *8***). Again, I include all the combinations to make the strategy set exhaustive.

When the norms are enforced by punishment, there is no ostracism subphase. Each pair of agents connected by the grid network interact once in the cooperation game per iteration. After the cooperation game, the two agents in each pair decide whether to punish their partner based on their own punishment strategy and their partner's cooperation and symbolic strategies. If *Agent A* decides to punish *Agent B*, *Agent A* spends a cost of λ and makes *Agent B* lose ρ in their payoff, where $\lambda < \rho$. Note that because the punishment happens after the cooperation game, each agent knows exactly whether their partner is a cooperator or defector and whether their partner is an X- or Y-performer.

The cost of the punisher λ and the cost of being punished ρ were manipulated. Six levels of λ were chosen, $\lambda \in [0.1, 0.2, 0.3, 0.5, 1, 1.5]$. For each level of λ, the ratio of ρ/λ were manipulated, $\rho/\lambda \in [3, 5, 7, 10, 12, 15]$. Thus, there were 6 * 6 = 36 combinations of $\lambda * \rho$ (see **Table 9**). Under each combination, 50 simulation runs were run under $\sigma = 0.15$.

Table 8 Punishment Strategies

Punishment strategy	Description
NP	Never punish any agents
CP	Punish cooperators
DP	Punish defectors
XP	Punish agents who perform *X* in the symbolic domain
YP	Punish agents who perform *Y* in the symbolic domain
CXP	Punish both cooperators and X-performers
CYP	Punish both cooperators and Y-performers
DXP	Punish both defectors and X-performers
DYP	Punish both defectors and Y-performers
AP	Punish everyone

Table 9 Combinations of λ and ρ

λ	ρ
0.1	$\rho \in [0.3, 0.5, 0.7, 1, 1.2, 1.5]$
0.2	$\rho \in [0.6, 1, 1.4, 2, 2.4, 3]$
0.3	$\rho \in [0.9, 1.5, 2.1, 3, 3.6, 4.5]$
0.5	$\rho \in [1.5, 2.5, 3.5, 5, 6, 7.5]$
1	$\rho \in [3, 5, 7, 10, 12, 15]$
1.5	$\rho \in [4.5, 7.5, 10.5, 15, 18, 22.5]$

2.9.2 Results

Figure 97 shows that the symbolic norm *X* evolved only when the cost of being punished (i.e., ρ value) is large enough (i.e., $\rho > 7$). When the cost of being punished is very small, punishment did not influence agents' behaviors much. As a result, the punishment did not lead to the evolution of any symbolic norm. When the cost of being punished is medium, interestingly, symbolic strategy *Y* becomes the norm. In other words, the symbolic strategy that initially co-occurs more often with *defection* becomes the norm.

Similarly, **Figure 98** shows that when ρ is high, the majority of agents are DYPs who punish defectors and Y-performers. However, when ρ is medium, the majority of agents are CXPs who punish cooperators and X-performers, a strategy that is exactly the opposite of DYP.

Why does the symbolic norm X evolve when ρ is high but the opposite norm evolve when ρ is medium? This depends on whether the punishment itself is benefiting or harming cooperation, which will be explained below.

Figure 97 The Evolution of Symbolic Norm Under Different Punishment Parameters

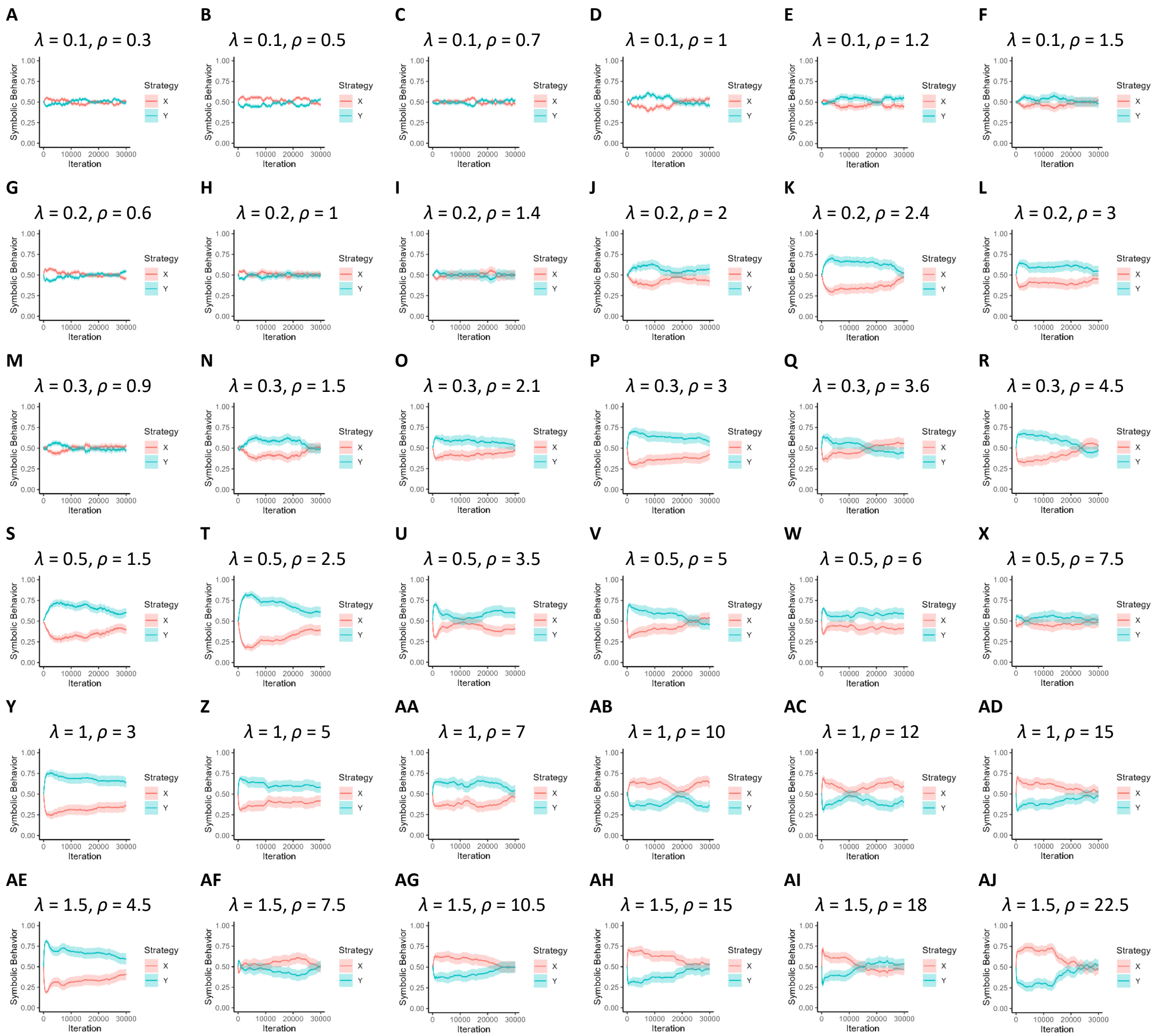
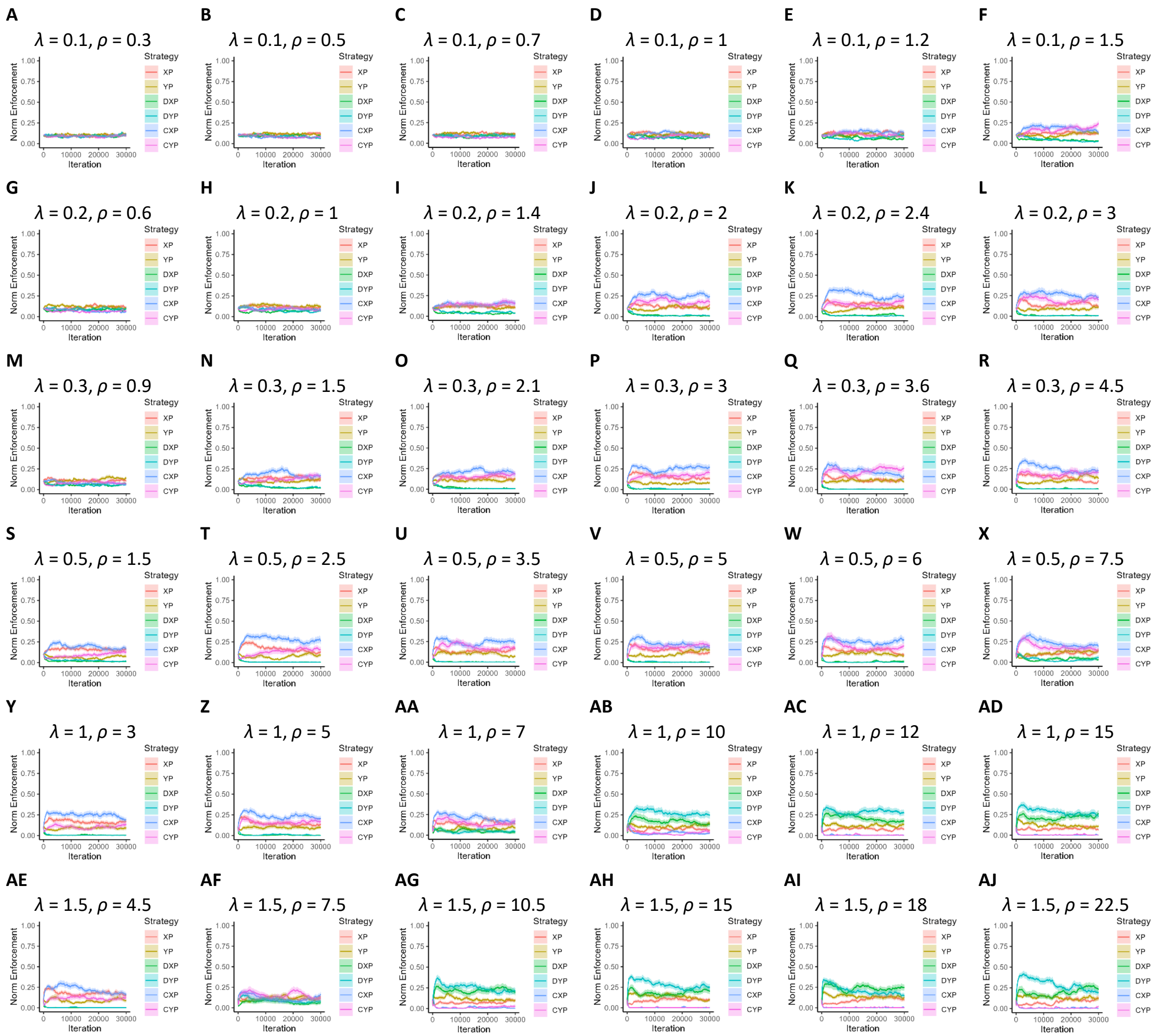


Figure 98 The Evolution of Norm Enforcement Under Different Punishment Parameters



2.9.3 The Impact of Punishment on Cooperation

This section examines the conditions under which punishment benefits vs. harms cooperation. Two sets of simulations were run. In the first set of simulations, I used a model *without* any punishment. Everything remains the same as in the punishment model except that, after the cooperation game, there was no punishment subphase. This first set of simulations aims to get a baseline cooperation rate without punishment so that it can be compared against the other conditions to examine the impact of punishment.

In the second set of simulations, I used a model in which agents only punish others based on their cooperation behavior, but not symbolic strategy. In this set of simulations, only four punishment strategies were used, 1) *NP*, who never punishes any agents, 2) *CP*, who punishes only cooperators, 3) *DP*, who punishes only defectors, and 10) *AP*, who punishes everyone. To match the proportions of the strategies in the original punishment model, I set the initial proportions of DP, CP, NP, and AP as 3/10, 3/10, 3/10, and 1/10, respectively. The logic of this setup is that 1) the new proportion of DP equals to the sum of the proportions of DP, DXP, and DYP in the original punishment model, 2) the new proportion of CP equals to the sum of the proportions of CP, CXP, and CYP in the original punishment model, 3) the new proportion of NP equals to the sum of the proportions of NP, XP, and YP in the original punishment model, and 4) the new proportion of AP equals to the proportion of AP in the original punishment model. Similarly, when mutation happens, the probability that an agent's strategy mutates to DP, CP, NP, and AP equals to 3/10, 3/10, 3/10, and 1/10, respectively. As in the original punishment, I used 6 * 6 = 36 combinations of $\lambda * \rho$ (see **Table 9**). Under each combination, 50 simulation runs were run under σ = 0.15.

When this second set of simulations is compared against the first set of simulations, it examines how practical punishment *alone* influences cooperation. When this second set of simulations is compared against the original punishment model, it examines how symbolic punishment influences cooperation.

Figure 99 shows the baseline cooperation rate without punishment. **Figure 100** shows the cooperation rate with only practical punishment under different values of λ and ρ. A comparison between **Figure 99** and **Figure 100** shows that when the cost of being punished ρ is medium, the existence of punishment *harms* cooperation. In fact, holding λ (i.e., the cost of the punisher) constant, as ρ gets higher, punishment based on practical behavior is increasingly harmful to cooperation.

Such a result is counterintuitive but is consistent with previous literature. Literature on punishment indicated that although prosocial punishment (i.e., punishing defectors) is often believed to sustain cooperation, once defectors are also allowed to punish cooperators, antisocial punishment (i.e., punishing cooperators) will evolve. In that case, punishment will not benefit cooperation (Rand & Nowak, 2011). **Figure 101** further confirms this literature. **Figure 101** shows that when ρ is medium, the majority of agents are anti-social punishers who punish cooperators (i.e., CPs).

However, when ρ is high (i.e., $\rho > 10$), the pattern is reversed. When the cost of being punished ρ is high enough, most agents become prosocial punishers (i.e., DPs) and punish defectors (see **Figure 101**). In this case, punishment starts to benefit cooperation (see the comparison between **Figure 99** and **Figure 100**).

Figure 99 The Evolution of Cooperation Without Punishment

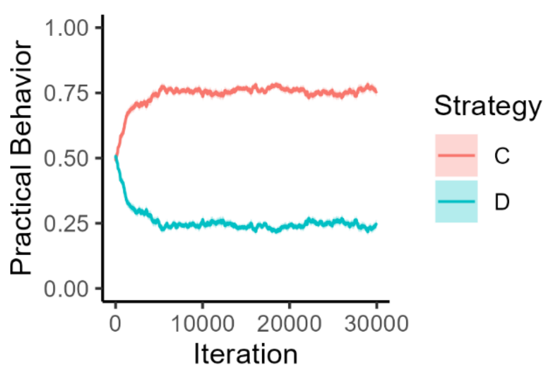


Figure 100 The Evolution of Cooperation Under Only Practical Punishment

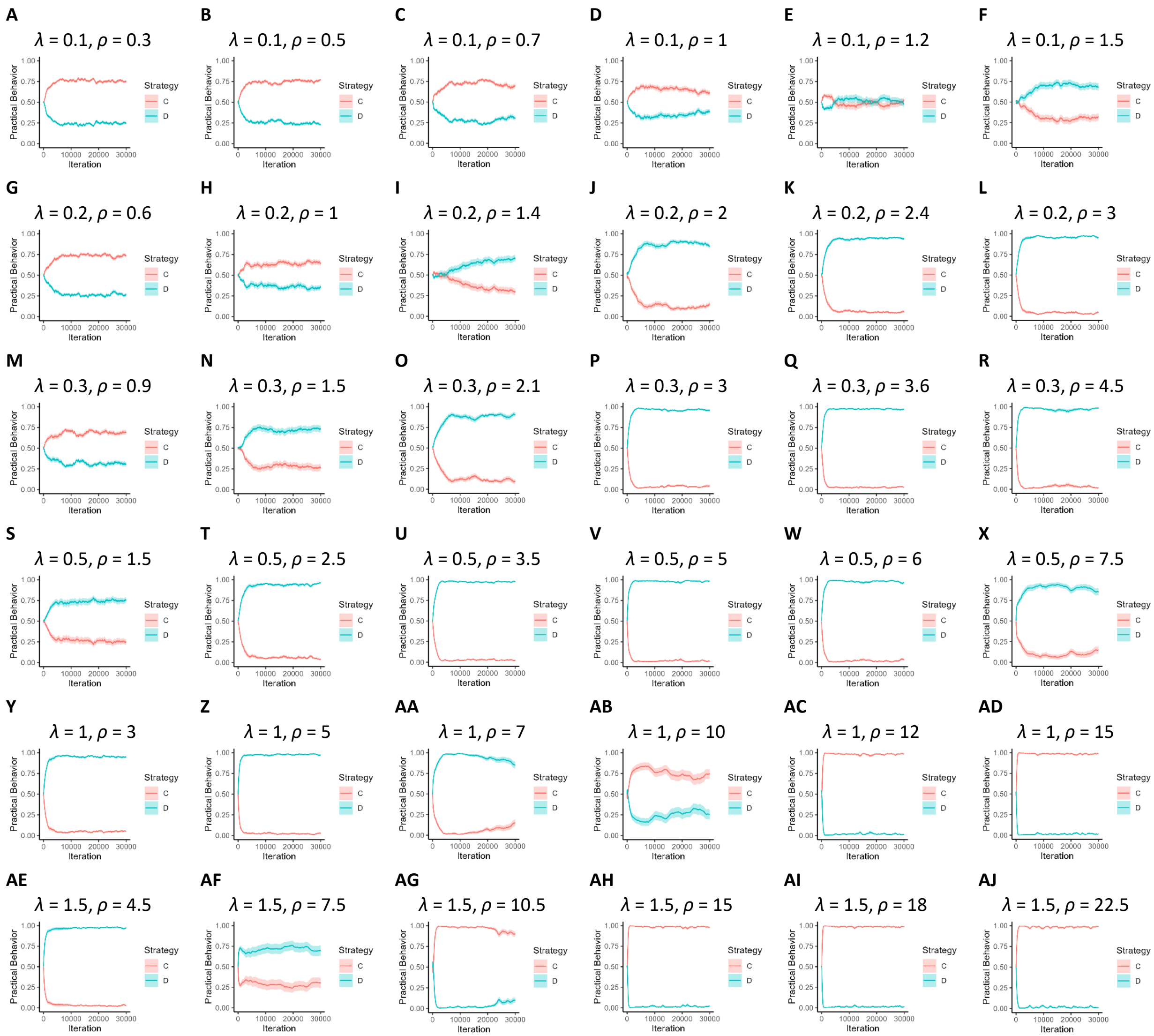
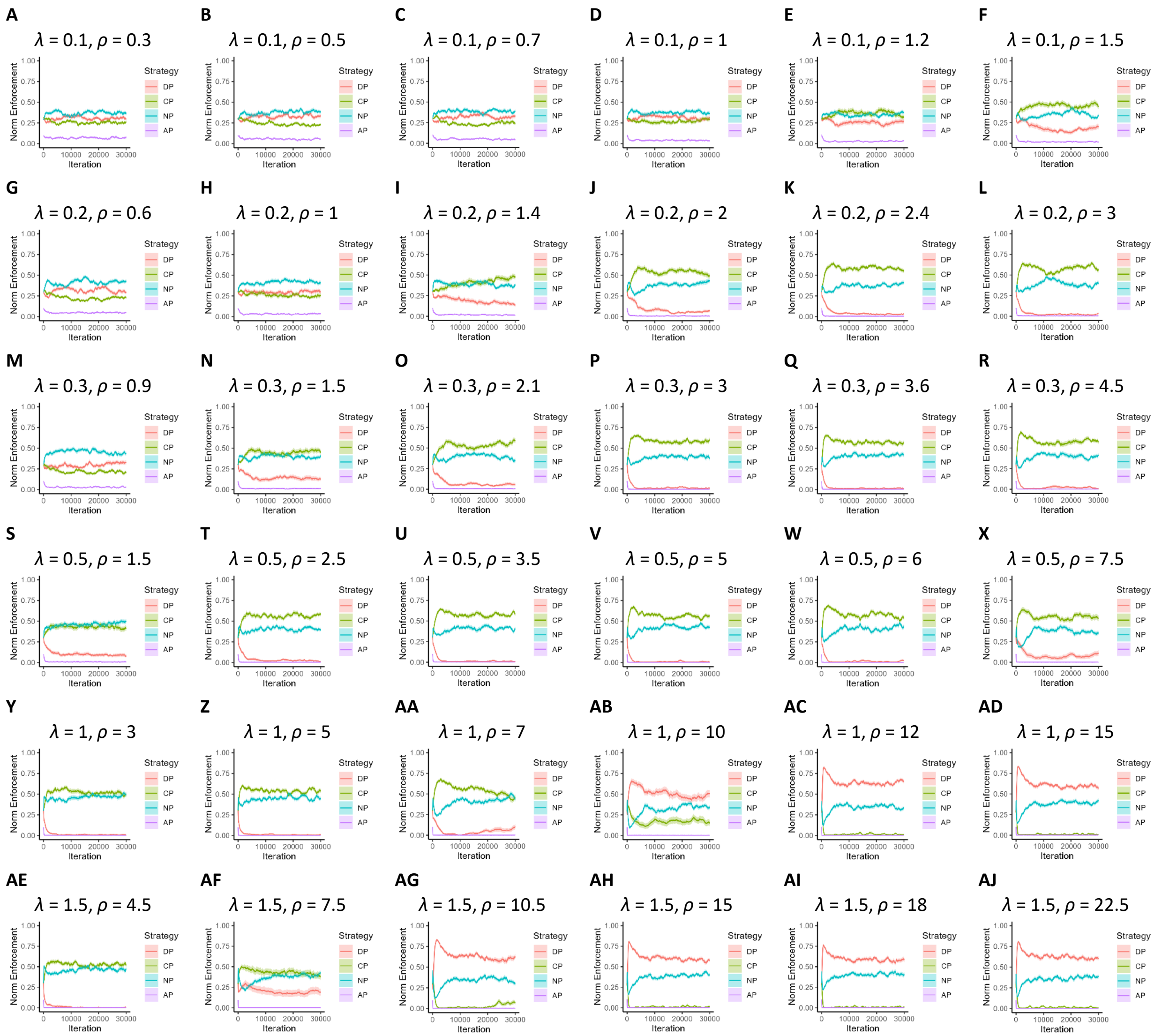


Figure 101 The Evolution of Pro- and Anti-social Punishment Under Only Practical Punishment



2.9.4 The Mechanism of Norm Spillover in the Punishment Model

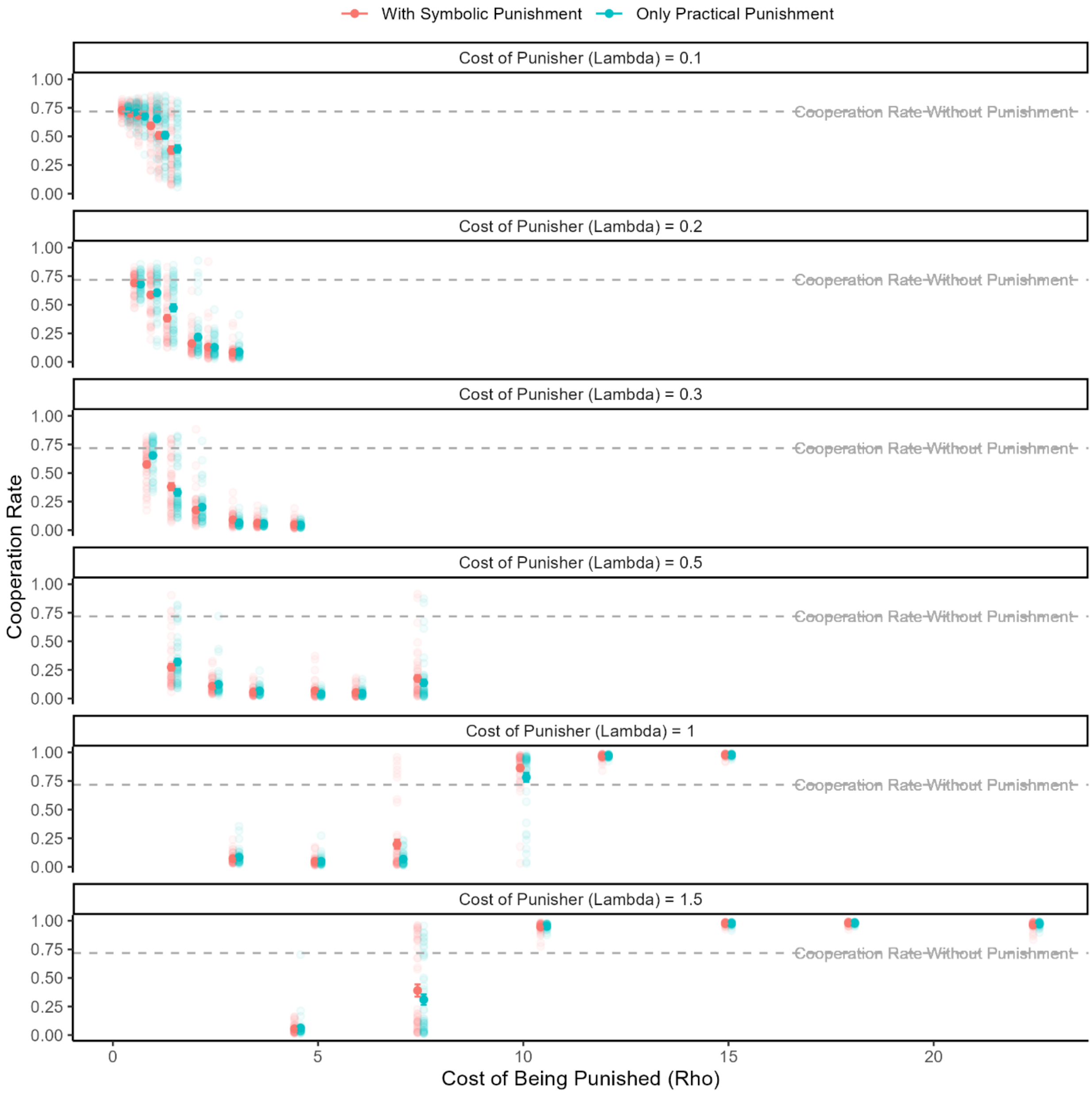
To summarize, when symbolic strategy *X* co-occurs more with cooperation and symbolic strategy *Y* co-occurs more with defection in a punishment model, whether *X* or *Y* evolves depends on whether punishment itself is benefiting or harming cooperation. If punishment benefits cooperation, as in the ostracism model, the symbolic norm *X* will evolve. However, if punishment itself harms cooperation, the opposite symbolic norm *Y* will evolve.

Furthermore, whether punishment benefits or harms cooperation depends on how severe the consequence of being punished is. Punishment can only benefit cooperation when the cost of being punished (i.e., ρ value) is high enough.

Figure 102, **Figure 103**, and **Figure 104** summarizes the cooperation rates in the early, middle, and late phases of the simulations respectively. Each figure shows the cooperation rate as a function of λ (i.e., the cost of the punisher) and ρ (i.e., the cost of being punished) and examines the impact of punishment on cooperation. The red points show the cooperation rates in the original punishment model with punishment based on both symbolic and practical behaviors (i.e., with all the ten punishment strategies). The green points show the cooperation rates in the model when there is only punishment based on practical behavior (i.e., with only DP, CP, NP, and AP). The dotted gray line shows the baseline cooperation rate in the model with no punishment. When a point is above the gray line, it means punishment is benefiting cooperation in that condition. When a point is below the gray line, it means punishment is harming cooperation in that condition.

Figure 102, **Figure 103**, and **Figure 104** show that punishment only benefits cooperation when the cost of being punished ρ is high enough. Consistent with these results, **Figure 105**, **Figure 106**, and **Figure 107** further show that the symbolic norm X evolves when ρ is high, the symbolic norm Y evolves when ρ is medium, and the symbolic strategy gets closer to random when ρ is small.

Figure 102 Early Phase Cooperation Rate Under No vs. Practical vs. Symbolic Punishment



Note. Cooperation rates in the early phases of the simulations. In this figure and the two figures below, the red points show the cooperation rates in the original punishment model with punishment based on both symbolic and practical behaviors (i.e., with all the ten punishment strategies). The green points show the cooperation rates in the model when there is only punishment based on practical behavior (i.e., with only DP, CP, NP, and AP). The dotted gray line shows the baseline cooperation rate in the model with no punishment. This figure shows cooperation rate as a function of λ (i.e., the cost of the punisher) and ρ

(i.e., the cost of being punished). When a point is above the gray line, it means punishment is benefiting cooperation in that condition. When a point is below the gray line, it means punishment is harming cooperation in that condition. Punishment only benefits cooperation when the cost of being punished ρ is high enough.

Figure 103 Middle Phase Cooperation Rate Under No vs. Practical vs. Symbolic Punishment

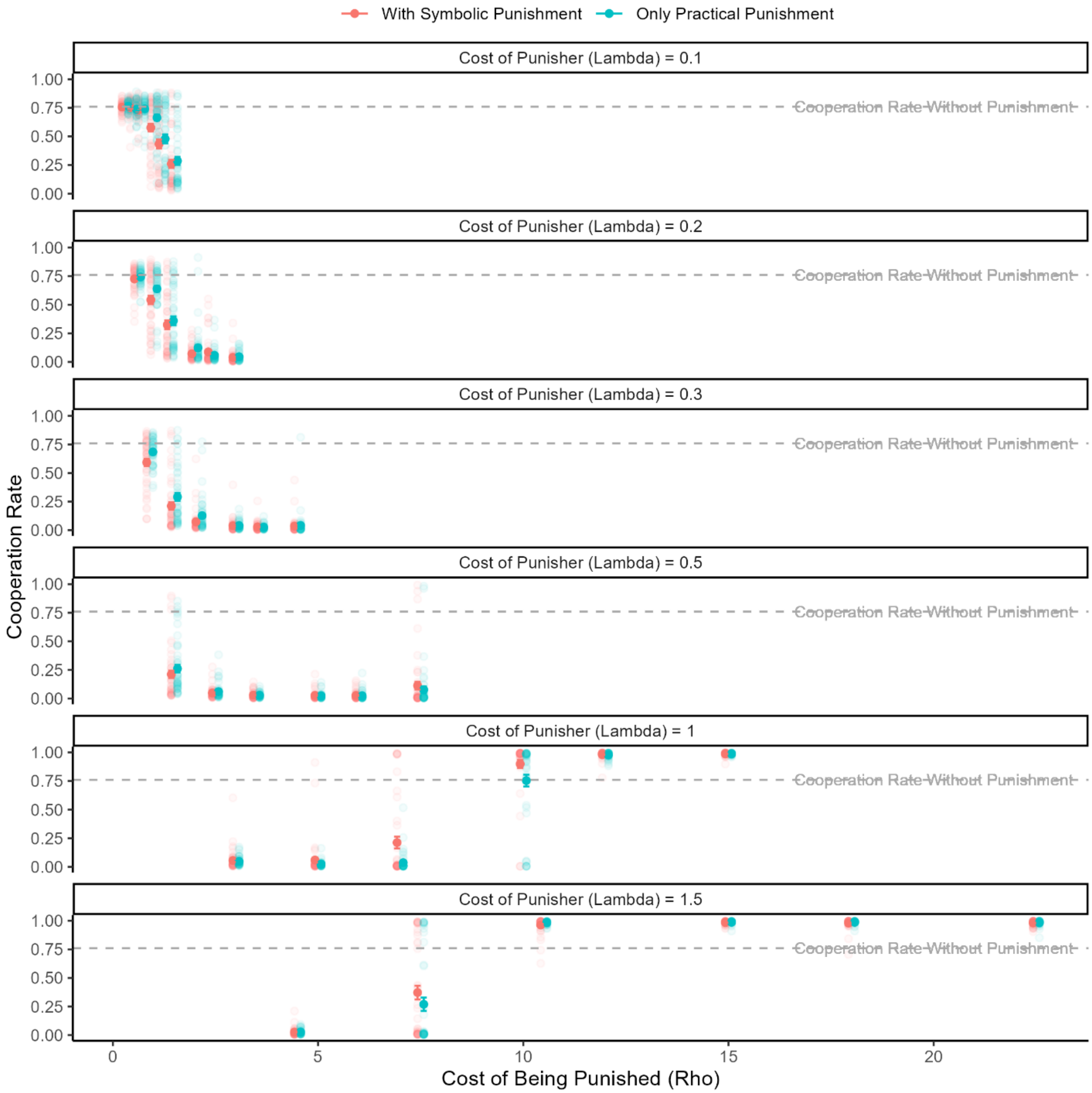


Figure 104 Late Phase Cooperation Rate Under No vs. Practical vs. Symbolic Punishment

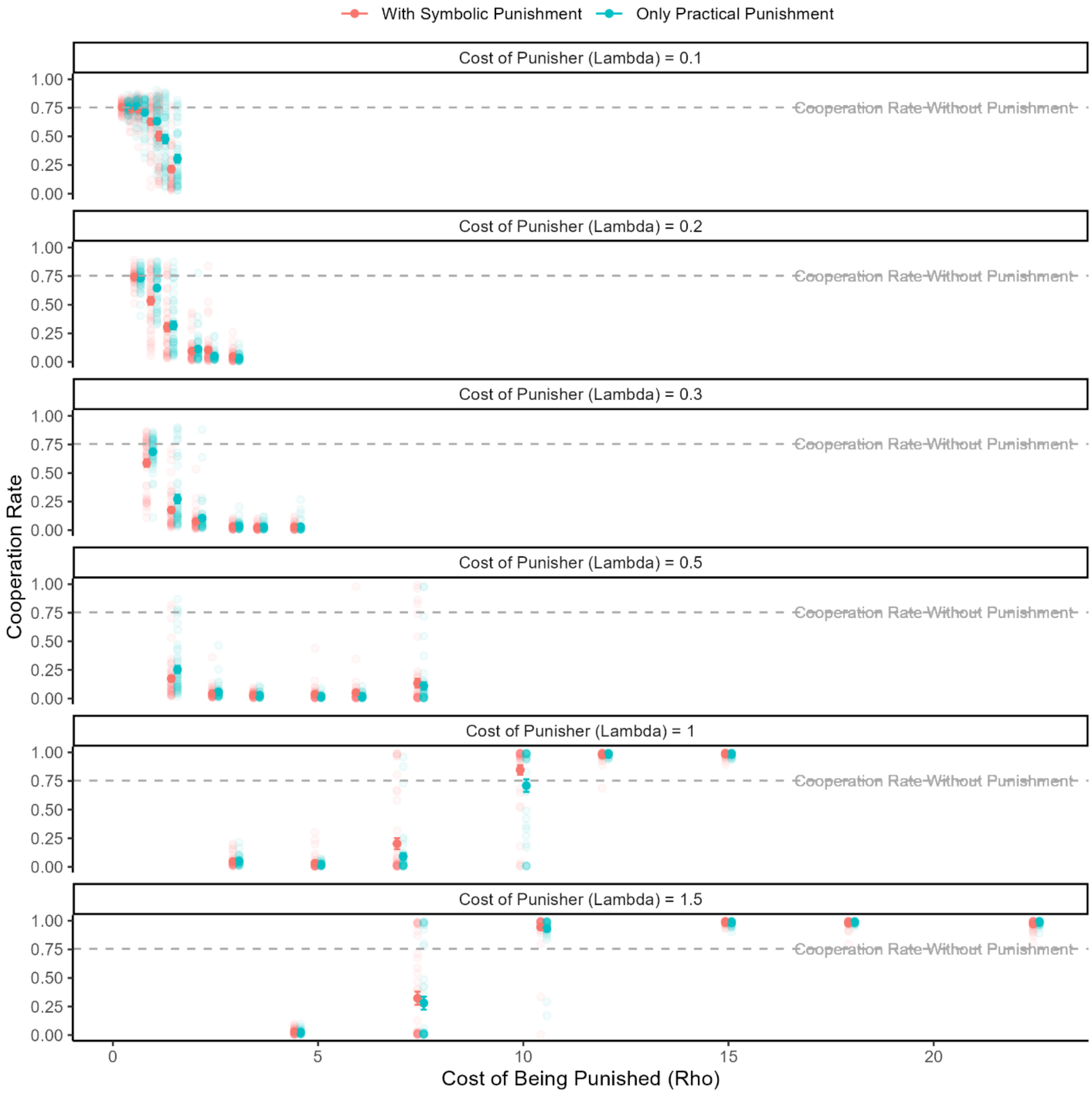
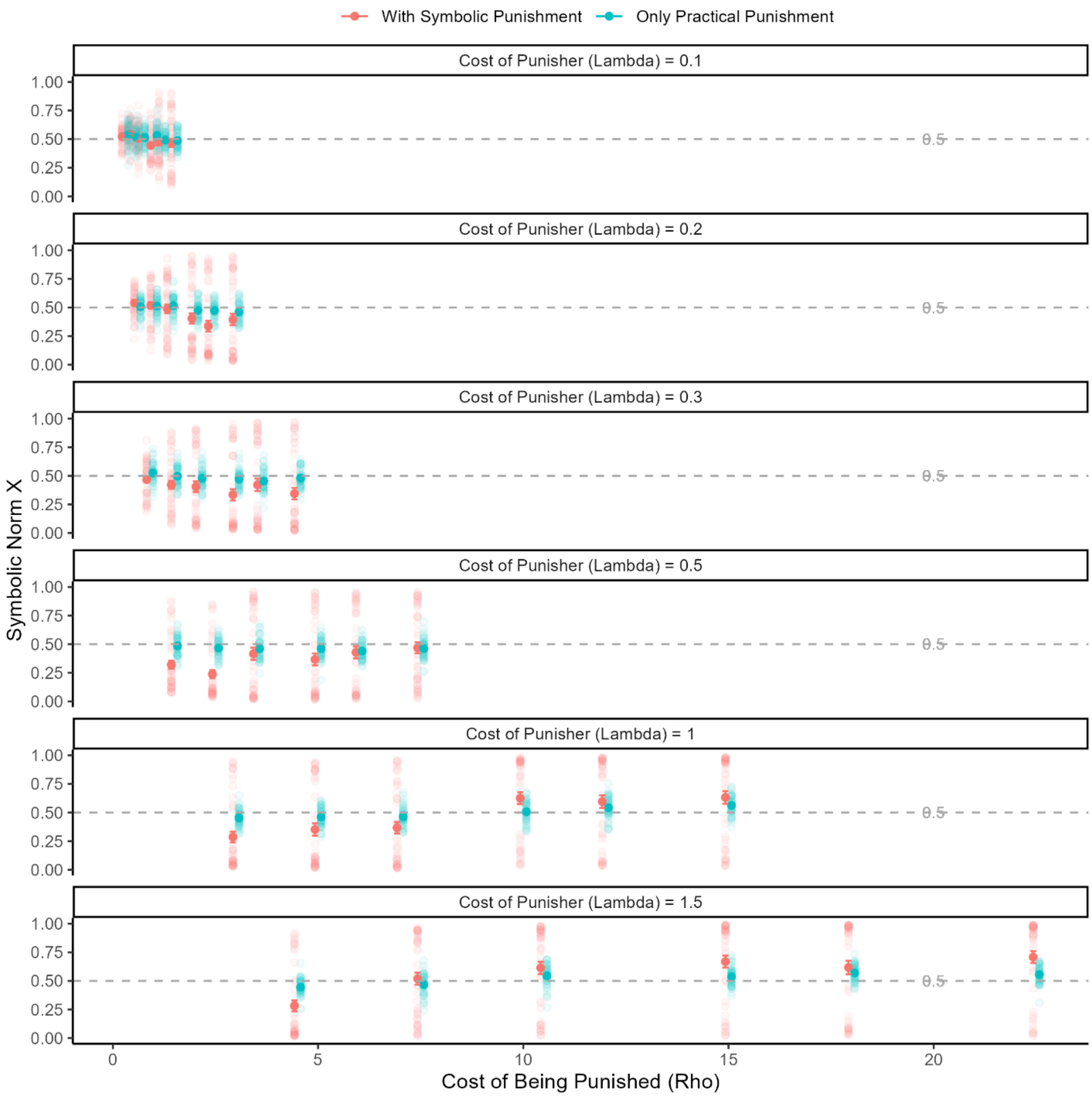


Figure 105 Early Phase Symbolic Norm Under Different Punishment Parameters



Note. The proportions of X-performers in the early phases of the simulations. In this figure and the two figures below, the red points show the proportions of X-performers in the original punishment model with punishment based on both symbolic and practical behaviors (i.e., with all the ten punishment strategies). The green points show the proportions of X-performers in the model when there is only punishment based on practical behavior (i.e., with only DP, CP, NP, and AP). The dotted gray line is set at 0.5, which represents a random symbolic norm. This figure shows the evolution of symbolic norms as a function of λ

(i.e., the cost of the punisher) and ρ (i.e., the cost of being punished). When a point is above the gray line, it means the symbolic norm X evolves in that condition. When a point is below the gray line, it means the symbolic norm Y evolves in that condition. All the green points fall around 0.5 because symbolic norms will not evolve when there is no symbolic punishment. When there is symbolic punishment (i.e., the red points), the symbolic norm X evolves when ρ is high, the symbolic norm Y evolves when ρ is medium, and the symbolic strategy gets closer to random when ρ is small.

Figure 106 Middle Phase Symbolic Norm Under Different Punishment Parameters

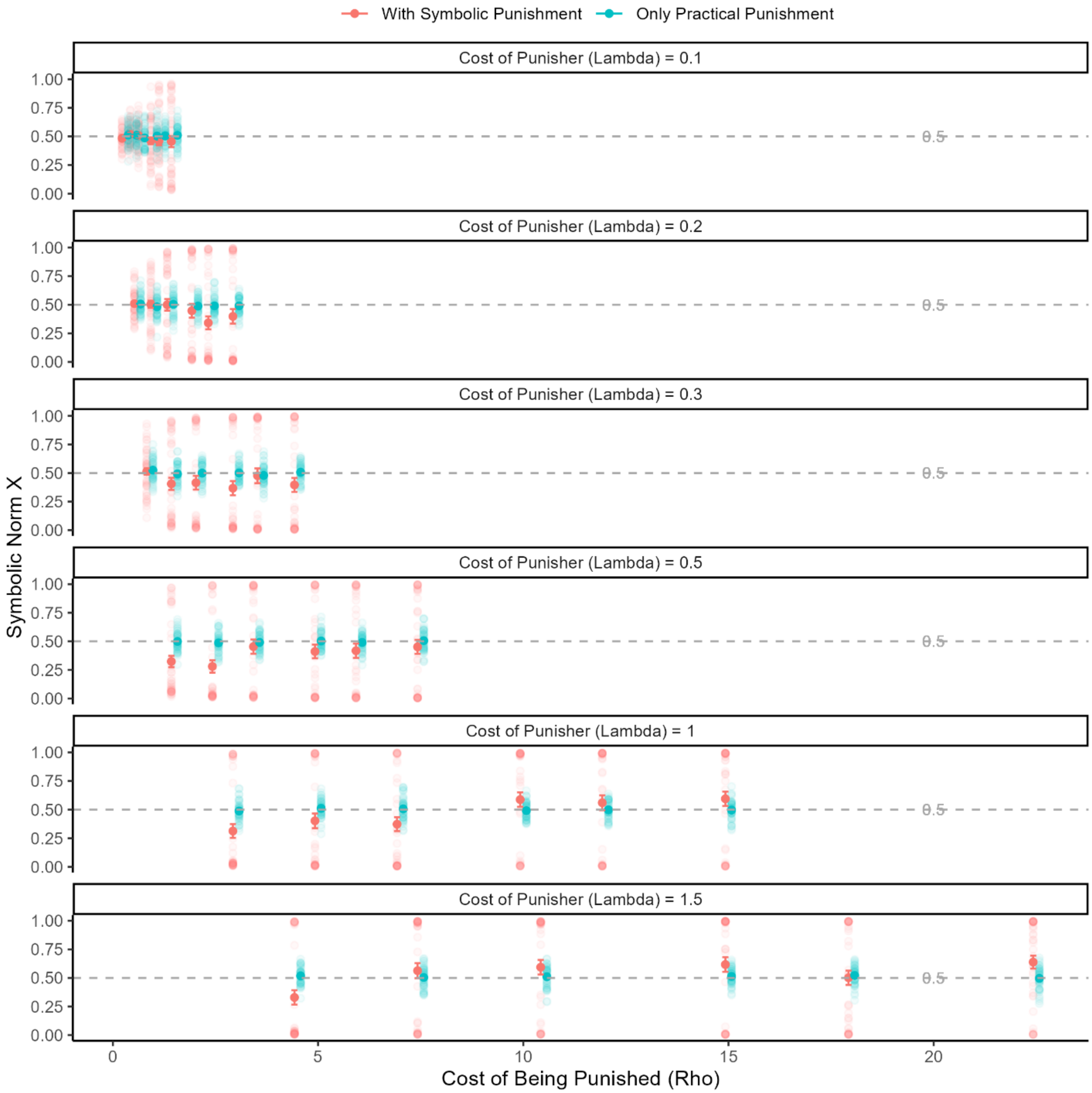
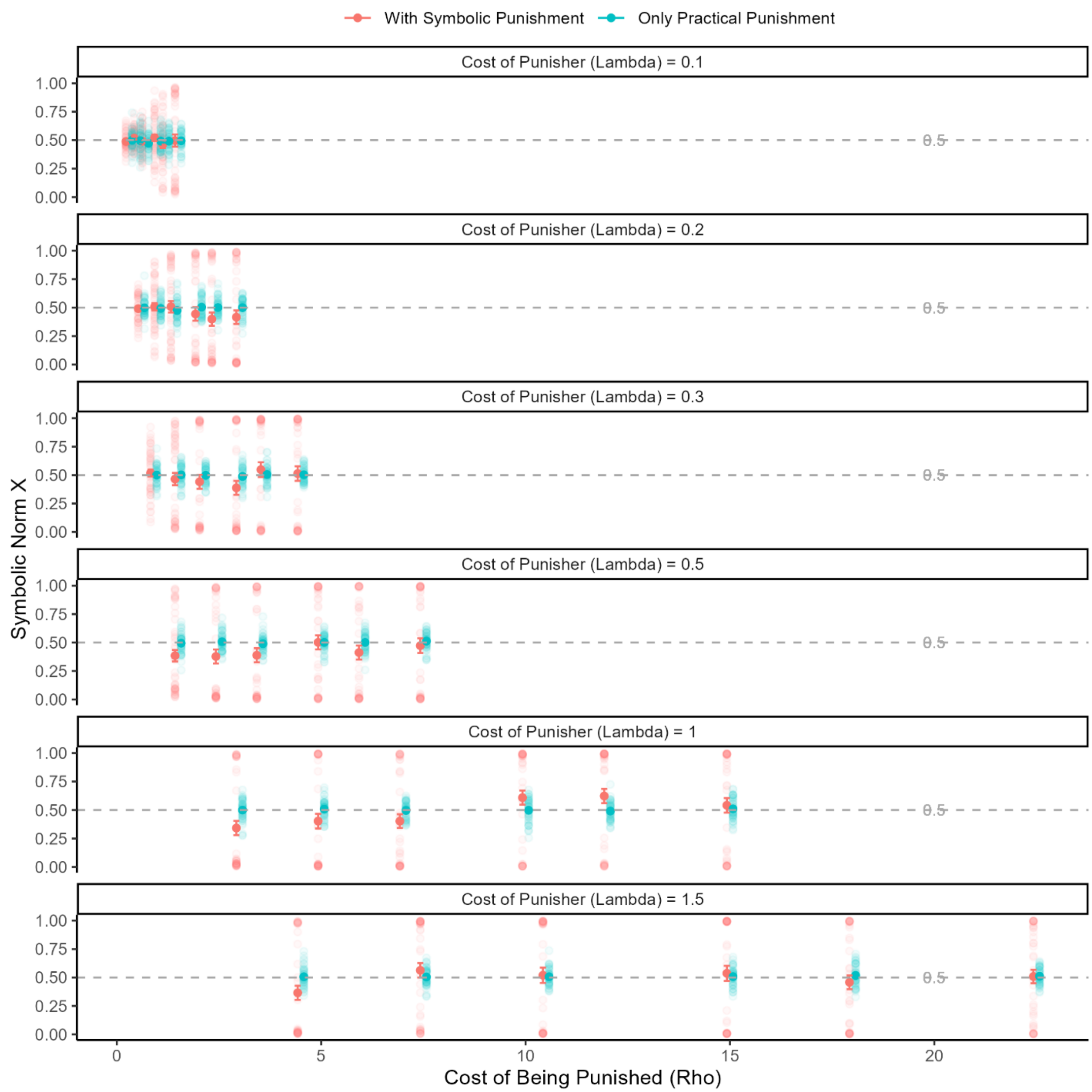


Figure 107 Late Phase Symbolic Norm Under Different Punishment Parameters



3 Supplementary Methods for the Empirical Experiment

3.1 The Age Composition of Participants

Table 10 The Age Composition of Participants

Age	Count
18-24 years old	89
25-34 years old	218
35-44 years old	117
45-54 years old	71
55-64 years old	38
65+ years old	18

3.2 The Selection of the Four Faces in the Partner Selection Task

The four faces in the partner selection section were chosen from the White male faces in the 10k US Adult Faces Database (Bainbridge et al., 2013). All the faces are White males to decrease variability. Same as the faces in the 28 profiles, only faces that meet the following requirements were chosen: 1) the eyes are gazing the front, 2) the emotion is happy, 3) the picture is a good profile picture, 4) the image quality is larger than 3.5 on a 1-5 scale, 5) the face is facing the front, 6) the person is not famous, 7) the age of the figure is at least 20, and 8) the participants passed the attention check when giving these ratings. After this exclusion, 349 faces remain. The four most similar faces were chosen from them via the following procedure.

Bainbridge et al. (2013) measured many traits of these faces. 56 traits were taken as variables to decide the similarities between faces. These traits include 1) amount of emotion, 2) emotion, 3) eye direction, 4) face direction, 5) facial hair, 6) catch question, 7) friendly, 8) makeup, 9) gender, 10) potential to be a movie star, 11) appropriate for profile picture, 12) image quality, 13) race, 14) memorable, 15) the speed at which the expression is happening, 16) the amount of teeth being showed, 17) atypical, 18) boring, 19) calm, 20) cold, 21) common, 22) confident, 23) egotistic, 24) emotion unstable, 25) forgettable, 26) intelligent, 27) introverted, 28) kind, 29) responsible, 30) trustworthy, 31) unattractive, 32) unemotional, 33) unfamiliar, 34) unfriendly, 35) unhappy, 36) weird, 37) aggressive, 38) attractive, 39) caring, 40) emotion stable, 41) emotional, 42) familiar, 43) friendly, 44) happy, 45) humble, 46) interesting, 47) irresponsible, 48) mean, 49) memorable, 50) normal, 51) sociable, 52) typical, 53) uncertain, 54) uncommon, 55) unintelligent, and 56) untrustworthy.

The Euclidean distances between all pairs of faces were computed. Based on these Euclidean distances, a hierarchical clustering analysis was done. Since there were 349 faces and four faces needed to be chosen, the clustering tree was cut into 87 (= 349 ÷ 4) clusters. The first cluster with four faces was chosen. The four faces in this cluster were used as the profile faces in the partner selection section.

3.3 The Strength of Social Norms in the Participant's Workplace

The strength of social norms in the participant's own workplace is measured by the workplace tightness scale, adapted from Gelfand et al. (2011).

1. There are many social norms that people are supposed to abide by in my workplace.
2. In my workplace, there are very clear expectations for how people should act in most situations.
3. People agree upon what behaviors are appropriate versus inappropriate in most situations in my workplace.
4. People in my workplace have a great deal of freedom in deciding how they want to behave in most situations.
5. In my workplace, if someone acts in an inappropriate way, others will strongly disapprove.
6. People in my workplace almost always comply with social norms.

3.4 The Strategy that Participants Used When Answering the Questions

The following questionnaire adapted from Haslam et al. (1996) is used to examine the strategies that the participants used when answering the questions.

1. When I watched the profiles, I was trying to identify on what basis people who prefer to meet at the start of the day differed from people who prefer the end of the day.

2. When I watched the profiles, I was trying to identify on what basis people with high ratings differed from people with low ratings.
3. My judgments were based on the last few profiles I saw as I remembered those ones best.
4. When I watched the profiles, I was trying to make sure I remembered whether each person prefers to meet at the start or the end of the day.
5. When I watched the profiles, I was trying to make sure I remembered how many high-rating and low-rating people there were among those who prefer the start or the end of the day.
6. I didn't pay too much attention as I watched the profiles.
7. When I was watching the profiles, I was trying to form a general impression of the people who prefer to meet at the start vs. end of the day.
8. When I was watching the profiles, I was trying hard to process as much information as possible about those who prefer to meet at the start vs. end of the day.
9. I found this experiment interesting.

4 Supplementary Results for the Empirical Experiment

4.1 Perceived Direction Function

A one-way ANOVA was done on the perceived direct function of the symbolic norm across the five conditions. The difference between the five conditions is significant ($F(4, 535) = 40.34$, $p < 0.001$). Although the difference between the start-strong and start-weak conditions is not significant ($p = 0.269$), the difference between the start-strong and no correlation conditions is significant ($p < 0.001$). The difference between the start-weak and no correlation conditions is not significant ($p = 0.400$). The difference between the no correlation and end-weak conditions is significant ($p = 0.002$). The difference between the end-weak and end-strong conditions is significant ($p = 0.002$).

4.2 Signaling Efficacy of the Symbolic Norm

4.2.1 Factor Analysis Across the 20 Traits

Exploratory factor analyses were done over the ratings on the 20 traits. Since the same question was measured once for the start of the day and once for the end of the day, and the difference between the start and end of the day was also computed for each of the 20 traits, three factor analyses were done on these three variables, respectively. These three factor analyses all show that the 20 traits fall on the same factor (see **Figure *108***, **Figure *109***, and **Figure *110***).

Figure 108 Scree Plot for the 20 Traits for the Start of the Day

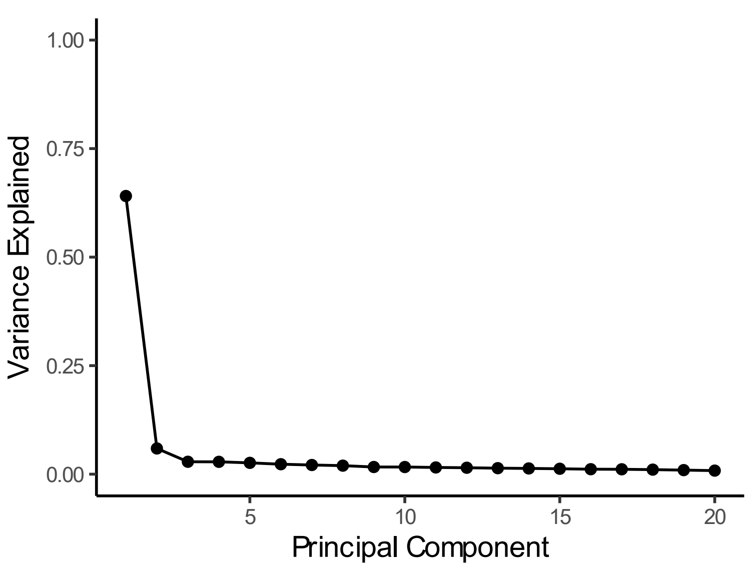


Figure 109 Scree Plot for the 20 Traits for the End of the Day

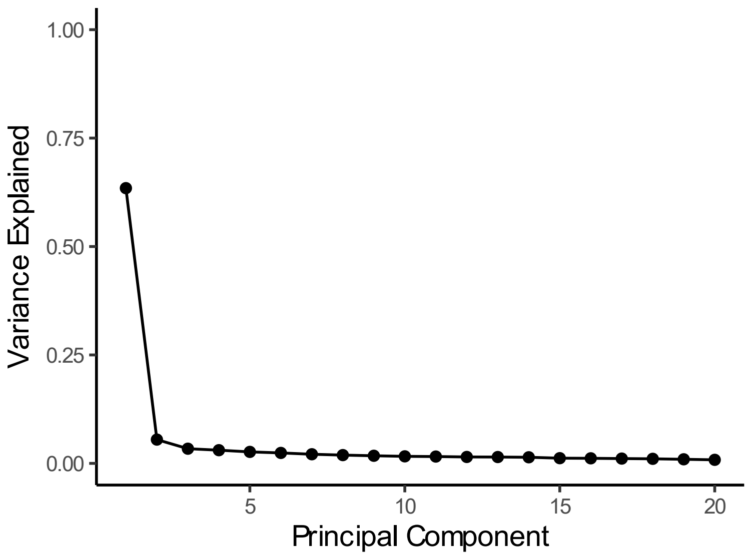


Figure 110 Scree Plot for the 20 Traits for the Difference Between the Start and the End of the Day

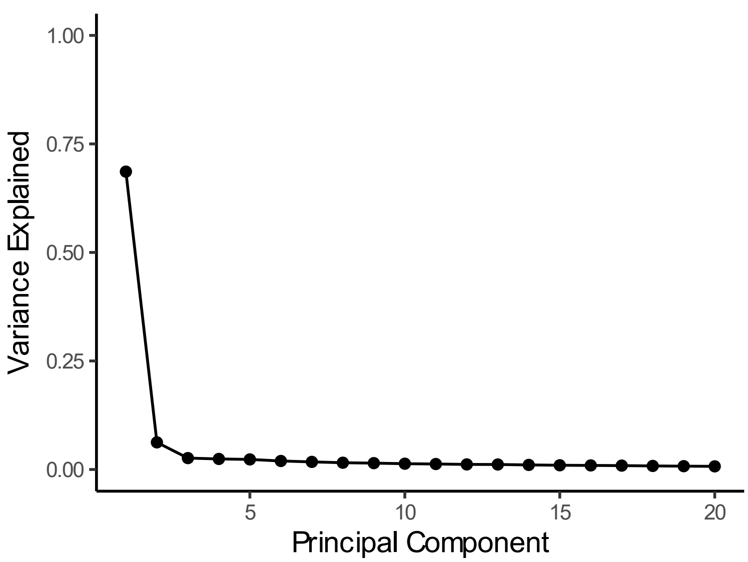


Table 11 Loadings of the 20 Traits

Original dimension	Variable	Questions for the start of day	Questions for the end of day	Difference between the start and the end of day (end - start)
Warmth	Sociable	0.722	0.676	0.707
Warmth	Sincere	0.752	0.759	0.772
Warmth	Warm	0.740	0.726	0.767
Warmth	Good natured	0.748	0.733	0.734
Warmth	Helpful	0.859	0.857	0.892
Competence	Persistent	0.773	0.781	0.802
Competence	Determined	0.807	0.819	0.837
Competence	Industrious	0.759	0.782	0.802
Competence	Skillful	0.800	0.799	0.821
Competence	Intelligent	0.779	0.731	0.777
Conformity	Compliant	0.764	0.758	0.792
Conformity	Agreeing	0.782	0.768	0.799
Conformity	Cooperative	0.836	0.814	0.858
Conformity	Obedient	0.683	0.748	0.747
Conformity	Accommodating	0.707	0.771	0.768
Commitment	Loyal to this organization	0.842	0.800	0.852
Commitment	Glad to join this organization	0.800	0.815	0.848
Commitment	Stick with this organization	0.810	0.811	0.869
Commitment	Care about the fate of this organization	0.847	0.815	0.877
Commitment	Make efforts to help this organization	0.860	0.850	0.898

4.2.2 Results When the Four Dimensions of Traits Are Analyzed Separately

The perceived correlation is positively correlated with the perceived warmth ($r(470) = 0.64$, $p < 0.001$), competence ($r(470) = 0.63$, $p < 0.001$), conformity ($r(474) = 0.68$, $p < 0.001$), and commitment ($r(477) = 0.68$, $p < 0.001$) in those who meet at the end (as against start) of the day. This is consistent with the results when the four dimensions are added up as a single factor.

4.3 Partner Selection Based on Symbolic Norm

4.3.1 Factor Analysis Across the Five Partner Selection Items

Exploratory factor analyses were done over the five ratings on the partner selection items. For the task in which the partner's practical behavior is known, three factor analyses were done, one on items for the start of the day, one on items for the end of the day, and one on different scores between the start and the end of the day. All factor analyses show a single factor (see **Figure *111***; see **Table *12*** for the loadings).

Figure 111 Scree Plots for Partner Selection Items When Partner's Practical Behavior is Known

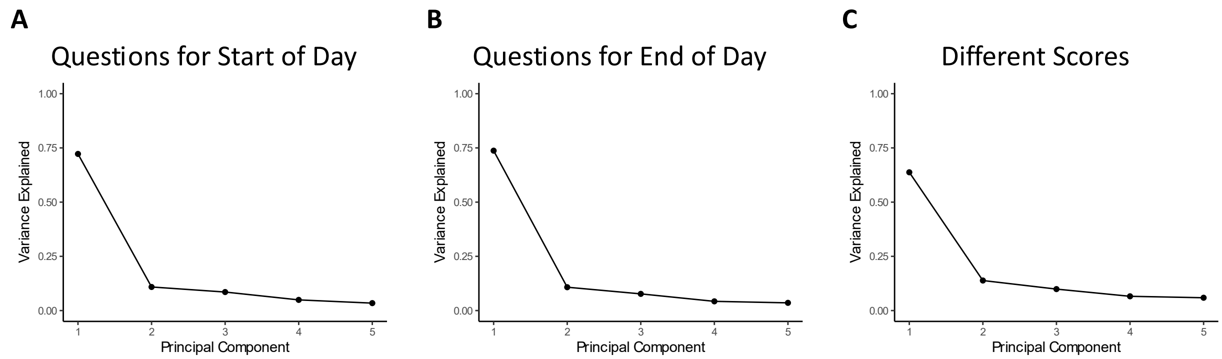


Table 12 Loadings of Partner Selection Items When Partner's Practical Behavior is Known

Item	Questions for start of day	Questions for end of day	Difference scores (end - start)
I want to work with him.	0.901	0.905	0.855
I want him to join my team.	0.914	0.903	0.836
I trust him.	0.729	0.769	0.461
I like him.	0.740	0.775	0.575
I feel similar to him.	0.706	0.718	0.700

For the task in which the partner's practical behavior is unknown, three factor analyses were done, one on items for the start of the day, one on items for the end of the day, and one on different scores. Again, all factor analyses show a single factor (see **Figure 112**; see **Table *13*** for the loadings).

Figure 112 Scree Plot*s* for Partner Selection Items When Partner's Practical Behavior is Unknown

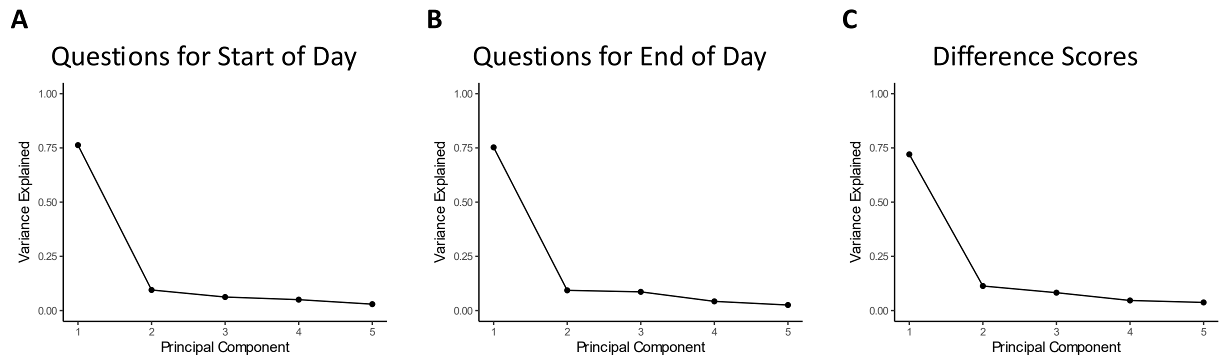


Table 13 Loadings of Partner Selection Items When Partner's Practical Behavior is Unknown

Item	Questions for start of day	Questions for end of day	Difference scores (end - start)
I want to work with him.	0.913	0.918	0.913
I want him to join my team.	0.934	0.936	0.910
I trust him.	0.799	0.728	0.666
I like him.	0.755	0.743	0.665
I feel similar to him.	0.752	0.762	0.724

4.3.2 ANOVA Analysis on Partner Selection Across Five Conditions

I also conducted two ANOVAs on partner selection across the five conditions, once on the task with known practical behavior and once on the task with unknown practical behavior. ***Figure 113*** shows that even when the partner's practical behavior is known, partner selection is still different across the five conditions ($F(4, 535) = 10.05$, $p < 0.001$). Post hoc tests show that although not all pairwise comparisons show significant difference, when meeting at the end of the day co-occurs more often with cooperative behavior, participants are generally more likely to select partners who meet at the end of the day. The differences between the following pairs of conditions are significant: start-strong and end-weak ($p = 0.003$), start-strong and end-strong ($p < 0.001$), start-weak and end-weak ($p < 0.001$), start-weak and end-strong ($p < 0.001$), and no correlation and end-strong ($p = 0.041$).

Figure 113 Partner Selection Across the Five Conditions When Practical Behavior is Known

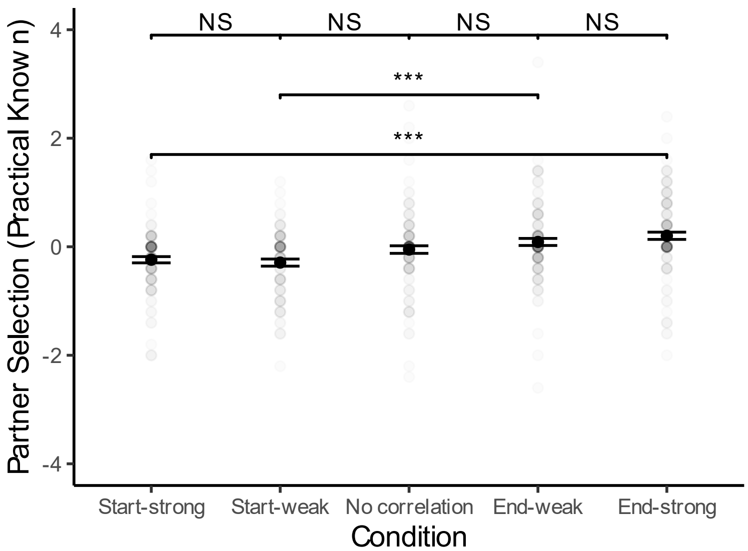
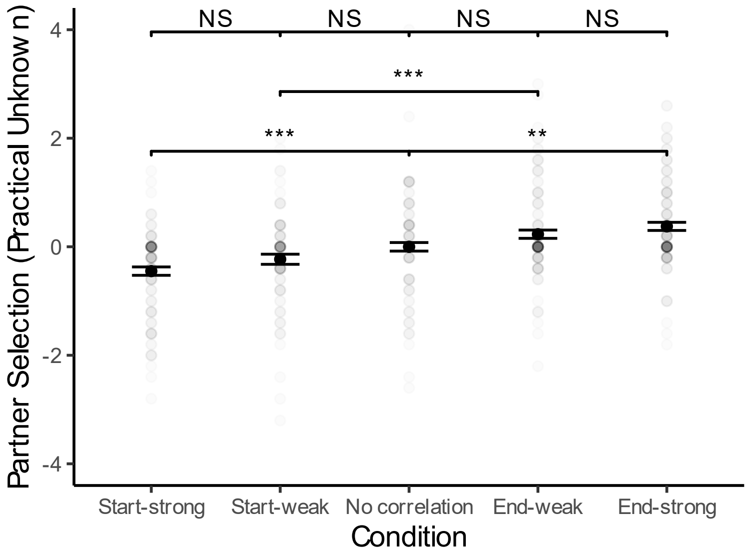


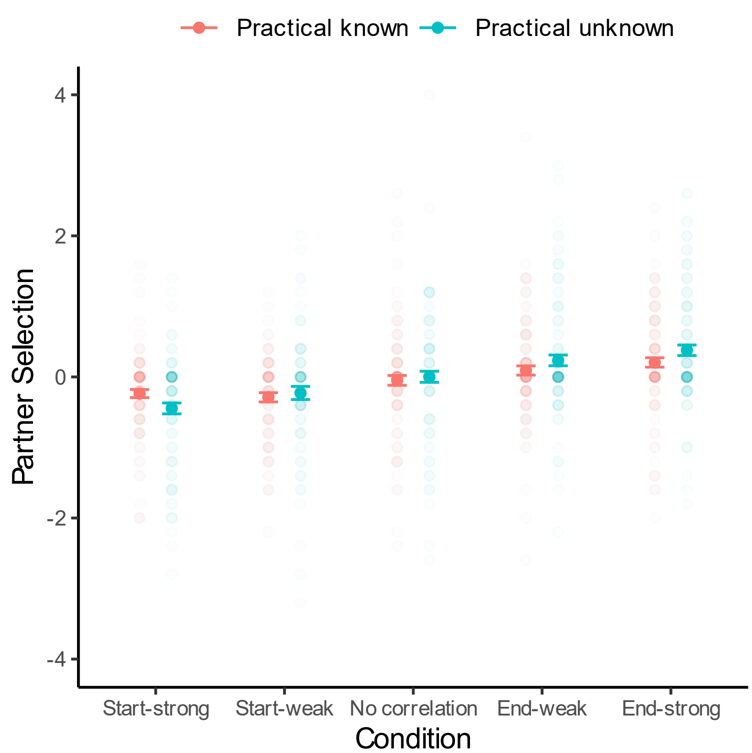
Figure 114 shows that when the partner's practical behavior is unknown, partner selection is also influenced by the correlation between the two behaviors ($F(4, 538) = 17.67$, $p < 0.001$). Post hoc tests show that when meeting at the end of the day co-occurs more often with cooperative behavior, participants are more likely to select partners who meet at the end of the day. The differences between the following pairs of conditions are significant: start-strong and no correlation ($p < 0.001$), start-strong and end-weak ($p < 0.001$), start-strong and end-strong ($p < 0.001$), start-weak and end-weak ($p < 0.001$), start-weak and end-strong ($p < 0.001$), and no correlation and end-strong ($p = 0.005$).

Figure 114 Partner Selection Across the Five Conditions When Practical Behavior is Unknown



Finally, I conducted a 5 (correlation: start-strong, start-weak, no correlation, end-weak, end-strong) * 2 (task: known practical behavior, unknown practical behavior) two-way ANOVA and examined how knowing the practical behavior of the partner influences partner selection. The interaction is significant ($F(4, 529) = 4.548$, $p = 0.008$). As shown in ***Figure 115***, the impact of symbolic norm on partner selection is larger when the partner's practical behavior is unknown.

Figure 115 Comparing Partner Selection Between Tasks When Practical Behavior is Known vs. Unknown



4.4 Perceived Norm Enforcement From Others

4.4.1 Factor Analysis Across the Five Perceived Norm Enforcement Items

Exploratory factor analyses were done over the five ratings on the perceived norm enforcement items. For the task in which the partner's practical behavior is known, three factor analyses were done, one on items for the start of the day, one on items for the end of the day, and one on different scores. All factor analyses show a single factor (see **Figure 116**; see **Table *14*** for the loadings).

Figure 116 Scree Plots for Perceived Norm Enforcement Items When Practical Behavior is Known

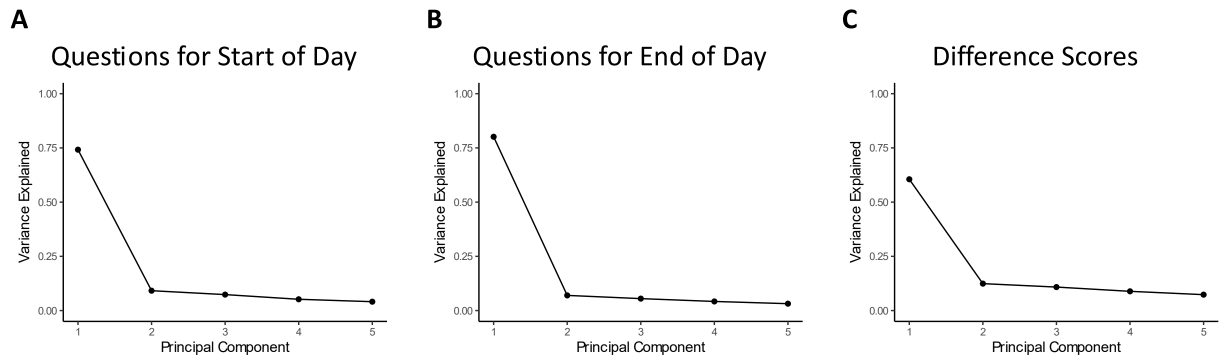


Table 14 Loadings of Perceived Norm Enforcement Items When Partner's Practical Behavior is Known

Item	Questions for start of day	Questions for end of day	Difference scores (end - start)
I think others in this company want to work with him.	0.866	0.913	0.769
I think others in this company want him to join their teams.	0.892	0.911	0.784
I think others in this company trust him.	0.796	0.844	0.652
I think others in this company like him.	0.842	0.861	0.653
I think others in this company feel similar to him.	0.692	0.792	0.659

For the task in which the partner's practical behavior is unknown, three factor analyses were done, one on items for the start of the day, one on items for the end of the day, and one on different scores. Again, all factor analyses show a single factor (see **Figure *117***; see **Table *15*** for the loadings).

Figure 117 Scree Plot*s* for Perceived Norm Enforcement Items When Practical Behavior is Unknown

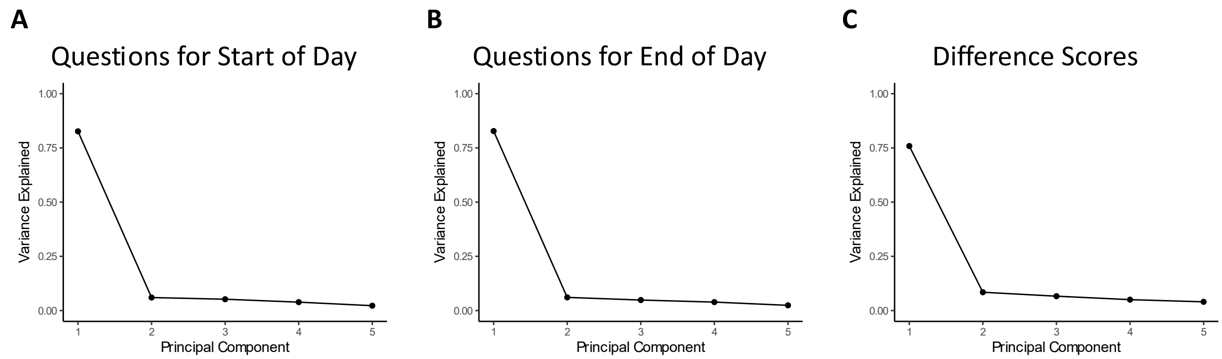


Table 15 Loadings of Perceived Norm Enforcement Items When Partner's Practical Behavior is Unknown

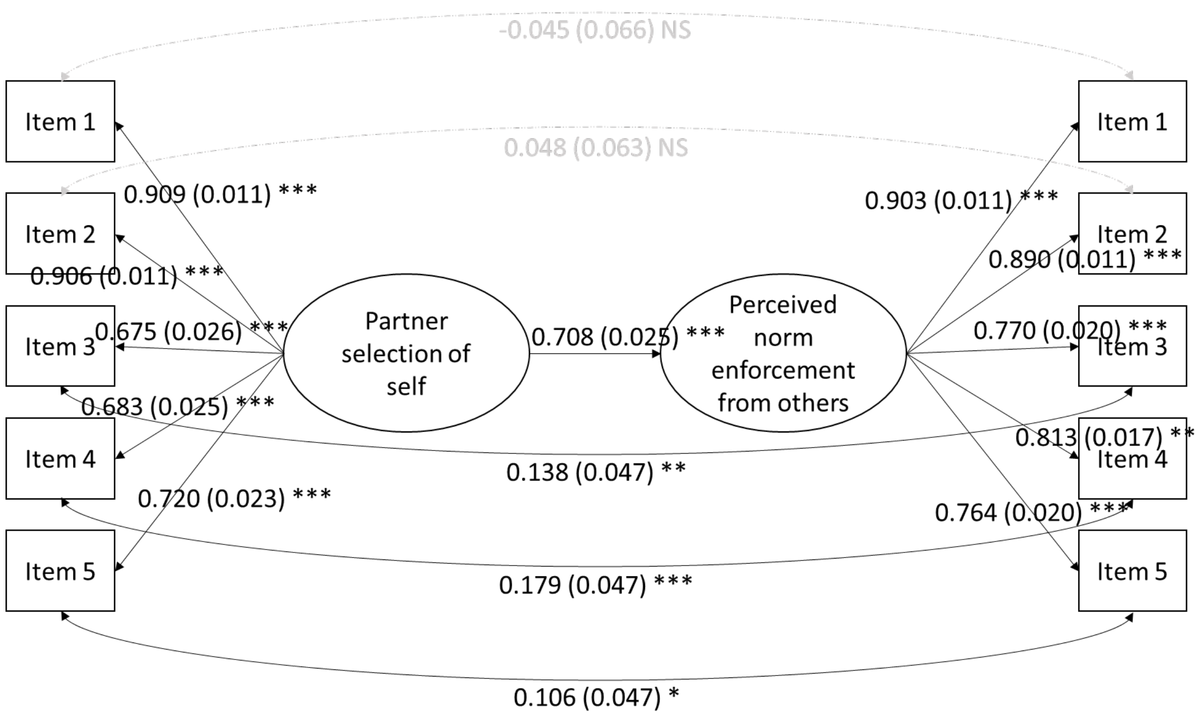
Item	Questions for start of day	Questions for end of day	Difference scores (end - start)
I think others in this company want to work with him.	0.935	0.937	0.900
I think others in this company want him to join their teams.	0.934	0.926	0.892
I think others in this company trust him.	0.843	0.832	0.772
I think others in this company like him.	0.866	0.861	0.820
I think others in this company feel similar to him.	0.833	0.860	0.759

4.4.2 Latent Variable Pathway Analysis Between Partner Selection and Perceived Norm Enforcement

A latent variable pathway model was used to examine the relationship between partner selection and perceived norm enforcement while controlling for the covariance between similar items. The model in ***Figure 118*** was used. The items that used the same language were allowed to covary (i.e., Item 1 in partner selection and Item 1 in perceived norm enforcement, etc.). After controlling for these covariances, one's own partner selection still strongly predict their perceived norm enforcement from others (standardized pathway coefficient $b = 0.708$, $p < 0.001$). The fit of this model is good (RMSEA = 0.092; CFI = 0.966; SRMR = 0.045).

Note that although the direction of the pathway in ***Figure 118*** is from one's own partner selection toward perceived norm enforcement from others, it does not mean that there is a causal relationship between the two latent variables. In fact, the coefficient will be the same if the pathway is in the opposite direction.

Figure 118 Latent Variable Pathway Model of Partner Selection and Perceived Norm Enforcement



4.4.3 ANOVA Analysis on Perceived Norm Enforcement Across Five Conditions

I also conducted two ANOVAs on perceived norm enforcement across the five conditions, once on the task with known practical behavior and once on the task with unknown practical behavior. ***Figure 119*** shows that perceived norm enforcement from others is significantly different across the five conditions when the partner's practical behavior is known ($F(4, 539) = 3.15$, $p = 0.014$). Post hoc tests only show significant difference between the start-strong and the end-strong conditions ($p = 0.007$), but the pattern is generally as expected. When meeting at the end of the day co-occurs more often with cooperative behavior, participants are more likely to select partners who meet at the end of the day.

Figure 119 Perceived Norm Enforcement Across the Five Conditions When Practical Behavior is Known

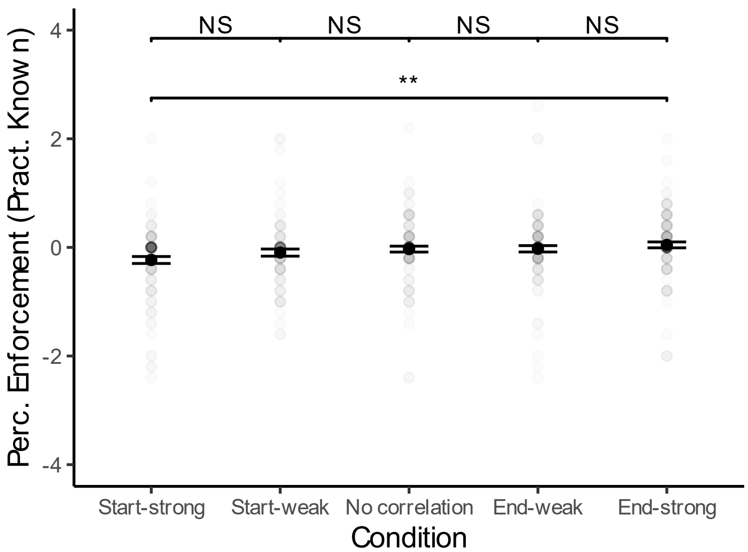
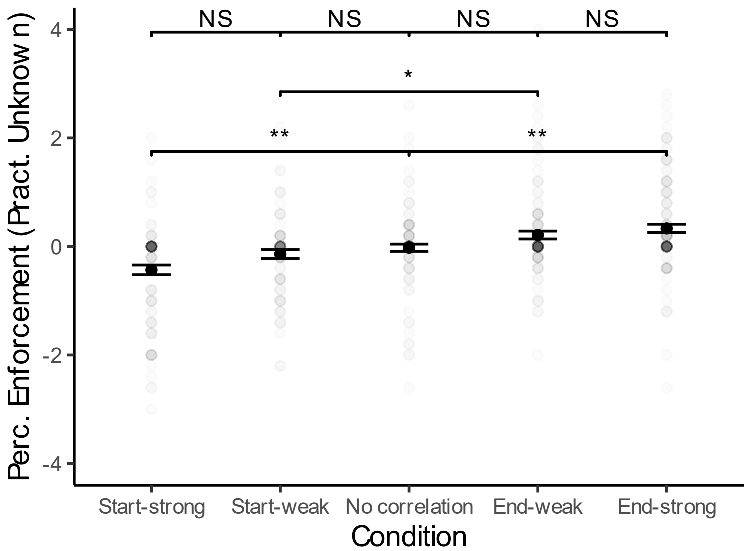


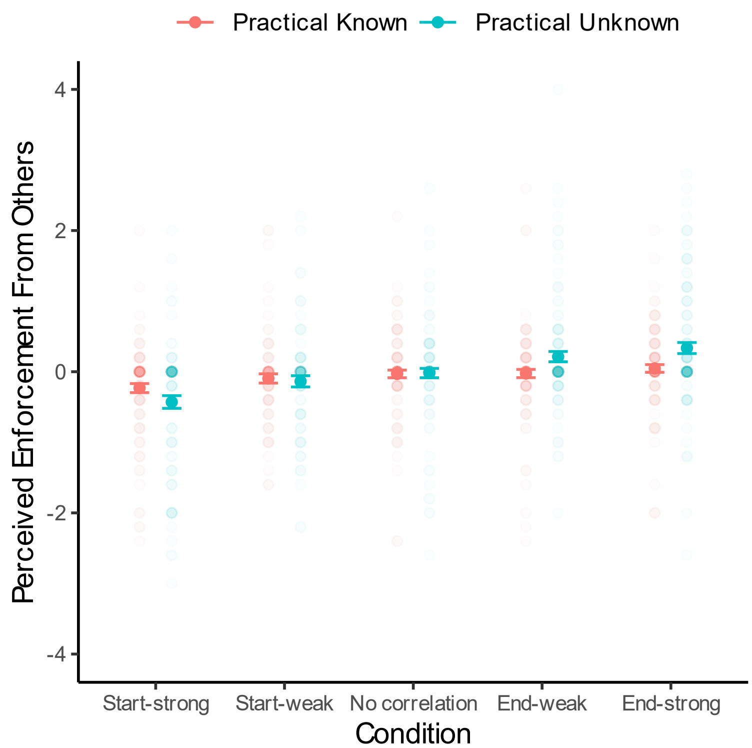
Figure 120 shows that perceived norm enforcement from others is significantly different across the five conditions when the partner's practical behavior is known ($F(4, 538) = 15.16$, $p < 0.001$). Post hoc tests show that when meeting at the end of the day co-occurs more often with cooperative behavior, participants perceive that other people are more likely to select partners who meet at the end of the day. The differences between the following pairs of conditions are significant: start-strong and no correlation ($p = 0.002$), start-strong and end-weak ($p < 0.001$), start-strong and end-strong ($p < 0.001$), start-weak and end-weak ($p = 0.019$), start-weak and end-strong ($p < 0.001$), and no correlation and end-strong ($p = 0.007$).

Figure 120 Perceived Norm Enforcement Across the Five Conditions When Practical Behavior is Unknown



I also conducted a 5 (correlation: start-strong, start-weak, no correlation, end-weak, end-strong) * 2 (task: known practical behavior, unknown practical behavior) two-way ANOVA and examined how knowing the practical behavior of the partner influences perceived other's partner selection. The interaction effect is significant ($F(4, 531) = 7.493$, $p = 0.016$). As shown in **Figure 121**, the impact of symbolic norm on perceived other's partner selection is larger when the partner's practical behavior is unknown.

Figure 121 Comparing Perceived Norm Enforcement When Practical Behavior is Known vs. Unknown



4.4.4 Comparing Partner Selection for Self and Others

Finally, I conducted a 2 (perspective: self, others) * 5 (correlation: start-strong, start-weak, no correlation, end-weak, end-strong) * 2 (task: known practical behavior, unknown practical behavior) three-way ANOVA to compare the difference between one's own partner selection and their perceived partner selection of others. The three-way interaction is not significant ($F(4, 516) = 1.97$, $p = 0.099$). The interaction between perspective and condition is not significant ($F(4, 516) = 2.05$, $p = 0.086$). The interaction between perspective and task is not significant ($F(1, 516) = 0.278$, $p = 0.598$). The main effect of perspective is not significant ($F(1, 516) = 0.109$, $p = 0.742$). Such results show that participants did not give significantly different patterns of answers when they think of their own vs. other's partner selection.

4.5 General Perception of the Symbolic Norm

4.5.1 Factor Analysis Across the Five Perceived General Norm Items

Three exploratory factor analyses were done over the five items on perceived general norm, one on items for the start of the day, one on items for the end of the day, and one on different scores. All factor analyses show a single factor (see ***Figure 122***; see ***Table 16*** for the loadings).

Figure 122 Scree Plots for the Five Perceived General Norm Items

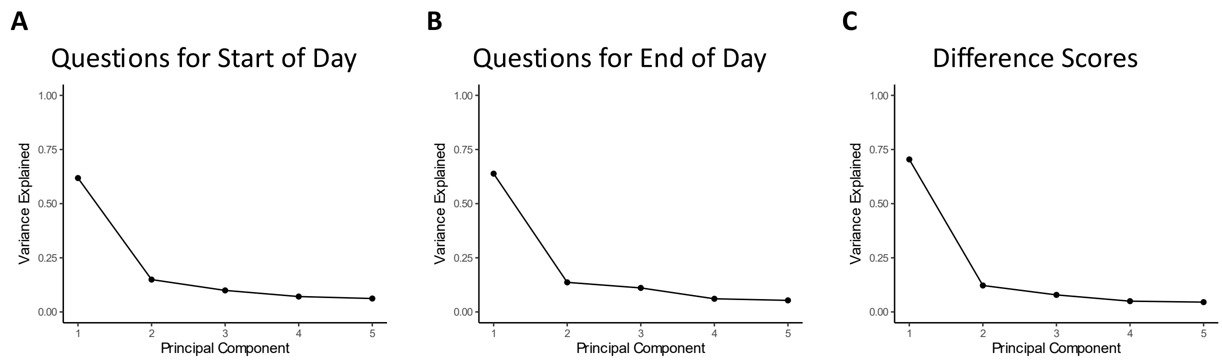


Table 16 Loadings of Items on Perceived General Norm

Item	Questions for start of day	Questions for end of day	Difference scores (end - start)
People in this company usually schedule meetings at the start/end of the day.	0.683	0.691	0.705
People in this company should schedule meetings at the start/end of the day.	0.754	0.763	0.810
People in this company will expect me to schedule meetings at the start/end of the day.	0.781	0.814	0.809
People in this company will disapprove if I schedule meetings at the end/start of the day.	0.507	0.509	0.585
I will schedule meetings at the start/end of the day because I expect that others prefer to meet at the start/end of the day.	0.832	0.848	0.922

4.5.2 Statistics on Perceived Descriptive Norm

This section shows the statistics for the perceived descriptive norm across the five conditions. There are two measures of perceived descriptive norm. The first measure is the first item in the perceived general norm section: People in this company usually schedule meetings at the start/end of the day. The second measure is the perceived prevalence of the behavior. At the beginning of the study, participants were asked to recall how many colleagues prefer to meet at the start of the day and at the end of the day, respectively. The difference score (end - start) is used as the measure of perceived descriptive norm. The larger the difference is, the more participants perceive the descriptive norm to be meeting at the end of the day. **Table 7** shows the statistics of these two measures.

Table 17 Means and Standard Errors for Perceived Descriptive Norm

Condition	Item 1 in general perception of norm	Perceived prevalence
Start-strong	-1.08 (0.165)	-1.67 (0.470)
Start-weak	-0.79 (0.171)	-2.16 (0.548)
No correlation	-0.48 (0.151)	-1.43 (0.492)
End-weak	-0.13 (0.159)	-0.68 (9,485)
End-strong	0.36 (0.169)	0.84 (0.475)

For the post hoc tests for Item 1, the differences between the following pairs of conditions are significant: start-strong and end-weak ($p < 0.001$), start-strong and end-strong ($p < 0.001$), start-weak and end-weak ($p = 0.047$), start-weak and end-strong ($p < 0.001$), and no correlation and end-strong ($p = 0.002$).

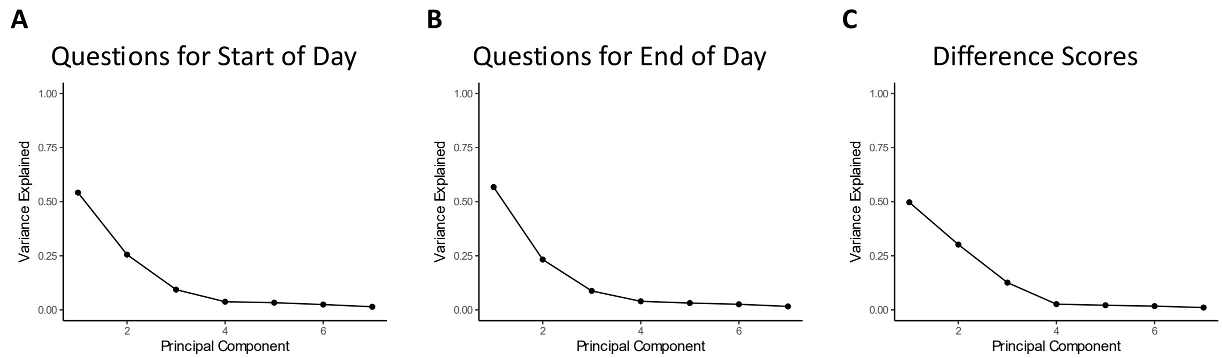
For the post hoc tests for perceived frequency, the differences between the following pairs of conditions are significant: start-strong and end-strong ($p = 0.003$), start-weak and end-strong ($p < 0.001$), and no correlation and end-strong ($p = 0.007$).

4.5.3 Factor Analysis Among Item 3-4 and Perceived Norm Enforcement Items

I proposed that Items 3-4 could be used as a measure of perceived norm enforcement. To examine the convergence between Items 3-4 and other items on perceived norm enforcement, three exploratory factor analyses were done over Items 3-4 and the five items. The three factor analyses include one on items for the start of the day, one on items for the end of the day, and one on different scores.

Figure 123 Plot A shows that, for items for the start of the day, the items load on two factors. **Figure 123** Plot B shows that, for items for the end of the day, the items load on two factors. **Figure 123** Plot C shows that, for items for difference scores, the items load on three factors. ***Table 18*** shows the loadings of these items. Direct oblimin method was used for rotation because the factors are supposed to be correlated. According to the loadings, the five items in the original partner selection task load on the same factor; Items 3-4 in the perceived general norm section load on another factor.

Figure 123 Scree Plots for the Seven Perceived Norm Enforcement Items



Note. Scree plot for the seven perceived general norm items. Plot A is the factor analysis over the questions for the start of the day. Plot B is the factor analysis over the questions for the end of the day. Plot C is the factor analysis over the difference scores. Plots A and B show there are two factors. Plot C shows there are three factors.

Table 18 Loadings of Items on Perceived Norm Enforcement

Item	Questions for start of day		Questions for end of day		Difference scores (end - start)		
	Factor 1	Factor 2	Factor 1	Factor 2	Factor 1	Factor 2	Factor 3
I think others in this company want to work with him.	**0.936**	-0.008	**0.912**	-0.027	**1.037**	-0.052	0.271
I think others in this company want him to join their teams.	**0.937**	-0.014	**0.909**	-0.040	**0.951**	0.001	0.122
I think others in this company trust him.	**0.842**	0.001	**0.779**	-0.015	**0.786**	0.004	-0.235
I think others in this company like him.	**0.863**	0.018	**0.831**	-0.025	**0.847**	-0.024	-0.104
I think others in this company feel similar to him.	**0.840**	-0.030	**0.761**	-0.005	**0.868**	0.008	0.050
People in this company will expect me to schedule meetings at the start/end of the day.	0.058	**0.484**	0.243	**0.443**	0.126	**0.507**	0.021
People in this company will disapprove if I schedule meetings at the end/start of the day.	-0.028	**1.003**	-0.103	**1.040**	-0.067	**1.014**	-0.017

4.6 Exploratory Analysis

4.6.1 Preference for Meeting Time

In real life, participants tend to be night people ($M(SE)$ = 3.40(0.061) on a 1-5 scale; t-test from 3: $t(549) = 6.53$, $p < 0.001$). In contrast, among participants who work full-time or part-time, in their workplace, they usually meet at the start of the day ($M(SE)$ = 2.51(0.051); t-test from 3: $t(383) = -9.76$, $p < 0.001$), prefer to meet at the start of the day ($M(SE)$ = 2.43(0.067); t-test from 3: $t(382) = 2.43$, $p < 0.001$), and think their colleagues prefer to meet at the start of the day ($M(SE)$ = 2.46(0.054); t-test from 3: $t(383) = -10.02$, $p < 0.001$). In the experiment, when participants do not see correlation between meeting time and cooperative behavior (i.e., no correlation group), they tend to schedule meetings at the start of the day ($M(SE)$ = 2.5(0.059); t-test from 3: $t(115) = -4.34$, $p < 0.001$). These results indicate that there is a preference and norm for meeting at the start of the day in real life. Nevertheless, when participants see a strong correlation between meeting at the end of the day and cooperation in a specific company, the experimental manipulation can reverse the real-world preference.

Participants' original lifestyle and meeting time preference influence their behavioral intention in the experiment. When participants do not see correlation between meeting time and cooperation in the experiment (i.e., no correlation group), night people tend to meet at the end of the day ($r(113) = 0.35$, $p < 0.001$). Participants are also more likely to meet at the end of the day if, in their workplace in real life, they prefer to meet at the end of the day ($r(76) = 0.65$, $p < 0.001$) and think their colleagues prefer to meet at the end of the day ($r(76) = 0.52$, $p < 0.001$), although the correlation between their usual meeting time in real life and their meeting time in the experiment is not significant ($r(76) = 0.19$, $p = 0.094$). Among all the participants, after controlling for the perceived correlation (i.e., ϕ), participants' behavioral intention is significantly predicted by whether they are a morning or night person ($b = 0.20$, $p < 0.001$), their usual meeting time in real life ($b = 0.30$, $p < 0.001$), their preferred meeting time in real life ($b = 0.49$, $p < 0.001$) and their colleagues' preferred meeting time in real life ($b = 0.36$, $p < 0.001$).

4.6.2 Moderation of Tightness and Income

I examined whether the relationship between the perceived correlation (i.e., ϕ) and behavioral intention is moderated by the workplace tightness in participants' real life. The hypothesis is that if a participant is embedded in a tight environment in daily life, they are more likely to adopt the symbolic norm (Gelfand et al., 2011). However, the moderating effect of tightness is not significant ($b = 0.10(0.20)$, $t = 0.49$, $p = 0.626$).

I also examined whether the relationship between the perceived correlation (i.e., ϕ) and behavioral intention is moderated by the income of the participants. The income was measured as the participant's total household income before taxes during the past 12 months. There were six levels: less than $25,000, $25,000-$49,999, $50,000-$74,999, $75,000-$99,999, $100,000-$149,999, and $150,000 or more. These six levels were recoded as 1-6 and used as a moderator for the relationship between perceived correlation and behavioral intention. However, the moderating effect of income is not significant, either ($b = 0.024(0.082)$, $t = 0.297$, $p = 0.767$).