In or out?

de Matos Fernandes, Carlos A.

DOI:
10.33612/diss.768055223

IN OR OUT?

THE PARADOX OF EXCLUSIONARY MECHANISMS IN KEEPING COOPERATION GOING

Carlos A. de Matos Fernandes

Provided by thesis specialist Ridderprint, ridderprint.nl
Printing: Ridderprint
Layout and design: Hans Schaapherder, persoonlijkproefschrift.nl

In or out?

The paradox of exclusionary mechanisms in keeping cooperation going

Proefschrift

ter verkrijging van de graad van doctor aan de
Rijksuniversiteit Groningen
op gezag van de
rector magnificus prof. dr. ir. J.M.A. Scherpen
en volgens besluit van het College voor Promoties.

De openbare verdediging zal plaatsvinden op

donderdag 26 oktober 2023 om 12:45 uur

door

Carlos Antonio de Matos Fernandes

geboren op 7 september 1990
te 's-Gravenhage

Promotor

Prof. dr. A. Flache

Copromotores

Dr. J. Dijkstra
Dr. D. M. Bakker

Beoordelingscommissie

Prof dr. R. P. M. Wittek
Prof dr. J. Tolsma
Prof dr. B. Simpson

Voor mijn gezin

CONTENTS

PRELUDE

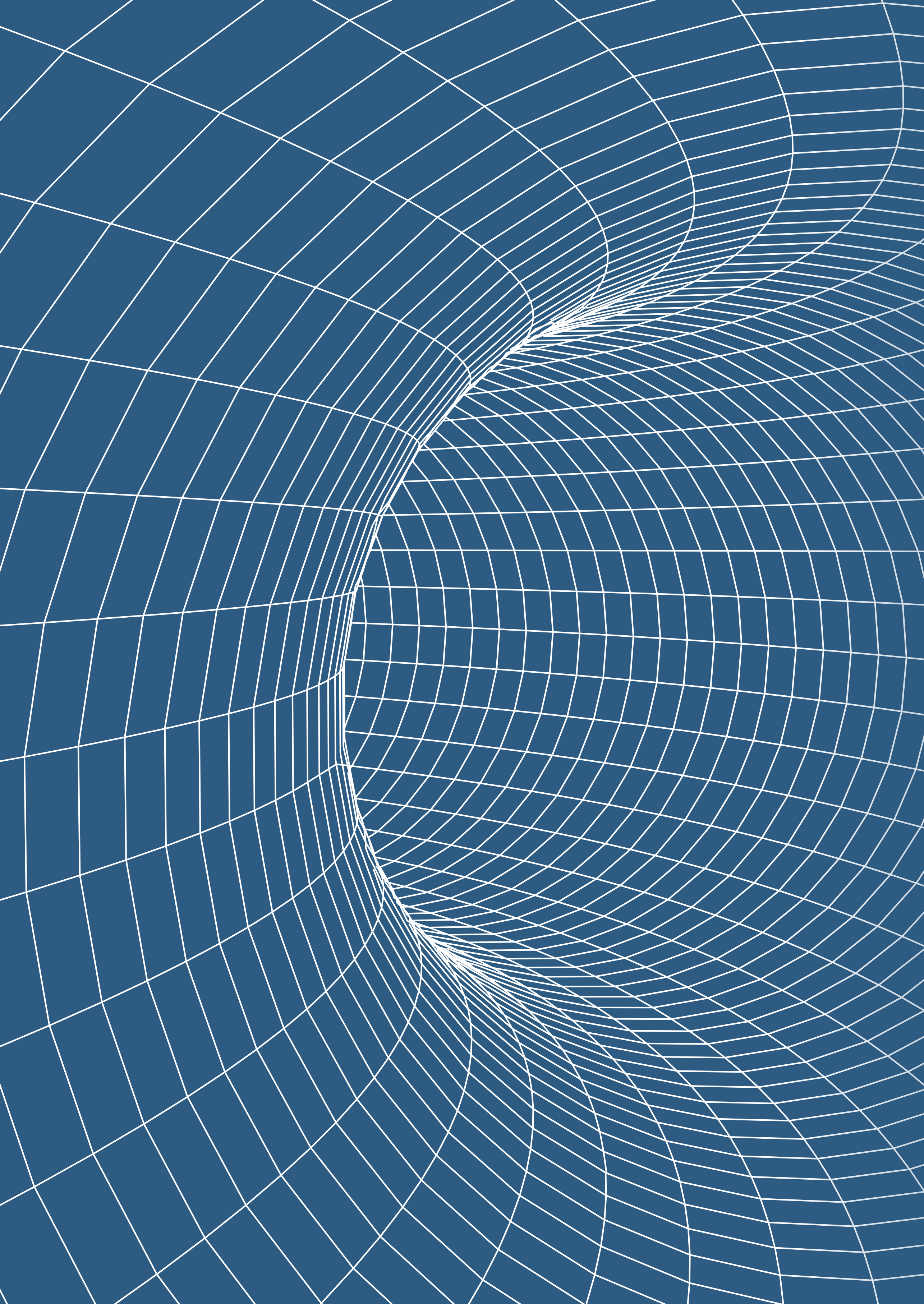
Imagine a prosocial or altruistic person who is generally inclined to behave so that it also benefits others involved. You probably think of someone you know, but let us call this person Fabio. Examples of Fabio's behavior – commonly labeled as cooperation – comprise volunteering, charitable giving, or helping someone move.

Now imagine a more proself or self-interested person who generally behaves in an individually beneficial way. We all know such a person, but I name him Carlos for now (for the sake of the example). Examples of Carlos' behavior – commonly labeled as defection – are talking at the movies, not picking up your dog's waste, or evading to pay taxes.

What happens if Carlos and Fabio are in a particular situation in which efforts from both are needed to realize benefits they cannot achieve alone? Consider organizing an event. They have limited time due to work arrangements and expertise in different areas. Thus, they need to work together to recruit volunteers, secure funding, and coordinate logistics. You need the joint production of two people to organize a successful event within the limited time frame. Fabio needs little incentive to invest time and effort since he is generally inclined to cooperate in such instances. Yet, Carlos shirks. He hopes that Fabio organizes the event (and that he, Carlos, can take credit for it). Carlos guards his self-interested needs.

Much is said and written about what Fabio can do when a mismatch in interests and needs arises. Possible solutions comprise expressing disapproval of Carlos' behavior, offering a reward for helping, or threatening Carlos with a social sanction. Whatever means Fabio might have to bring Carlos in line; if Carlos can walk off and end their relationship, all of these solutions fail. One could say that having an "out" allows more selfish people to test the system into which they are embedded. Carlos tests, for example, if his defection puts the relationship with Fabio in danger.

There is a solution. For Fabio, excluding Carlos from the pool of people to ask when such endeavors arise could have prevented the mismatch in needs and expectations. More information on Carlos' prior behavior and his inner workings would have been necessary for Fabio to make this possible, but such information was not readily available. Fabio could – and perhaps wanted to – have asked Seppe or Matteo – well-known cooperators – to help him instead. An exclusionary mechanism ideally allows those involved to prevent or act upon instances where a mismatch in interests and needs occurs or may occur. As a post hoc measure, Fabio's "bad" experience motivates him to exclude Carlos from future cooperative endeavors. This exclusion also signals to others that Carlos is not a good candidate if cooperation is required. Or so it seems.



Chapter 1

Cooperation: A clash between individual and collective interests

For even the very wise cannot see all ends.

—A quote by Gandalf in J. R. R. Tolkien's *The Fellowship of the Ring* (1954)

ABSTRACT

Chapter 1 provides an overview of the research topic, namely the cooperation problem, and introduces the context in which such problems arise. It explains what exclusion is and how exclusionary mechanisms affect cooperation dynamics, networks, and groups. The research questions in this dissertation are outlined, and the contributions of this work to the existing body of research are highlighted. The role of exclusion is specified from a theoretical perspective, and insights from empirical work are introduced step-by-step to add complexity. The data and methodology used in the research are addressed. Chapter 1 outlines the chapters presented in this thesis, with key findings summarized in grey boxes throughout Chapter 1.

1.1 SETTING THE STAGE

In the “real world,” we all work together frequently with others for mutual benefits. At times, cooperation goes smoothly, but not always: Some freeride and let others do most of the work in collective endeavors. In this dissertation, I focus on instances where cooperation in *dynamic* group and network settings flourishes and withers. Workgroups break up and new teams form, friendships end and new ones are formed, and people change their stances over time due to, for example, social influence. Accounting for the dynamics of real-life, research generally shows that cooperation fares better in some groups and networks than in others. Why?

Among a myriad of explanations, one answer to this question is that dynamics – i.e., the possibility to change network or group relationships – offer the opportunity for *exclusion*. Cooperators can exclude defectors from their group or cut ties to them in networks, thereby protecting their benefits from cooperation. Exclusion enables individuals to differentiate with whom they form a group or a network relationship. For example, if three students work together on a report and one slacks off, they may exclude the free-rider from joining their group in future endeavors. Then, the free-rider can repent and promise to cooperate in the future, or remain excluded if the free-riders are unwilling to change their behavior. Note that benefits from cooperation accrue to those included. Thus, if a student is excluded from a project team, a good grade or some bonus associated with producing a good report is not awarded to the excluded student.

This dissertation underscores a pivotal distinction in the objectives that underlie exclusion mechanisms within the context of cooperative systems. At first, it protects cooperators from exploitation by defectors. In this sense, excluding defectors is fair and just. As such, exclusion provides a protective bulwark for cooperators to continue their behavior, creating an environment wherein the benefits of cooperation can be shared equitably. Cooperation is (or may be) promoted since it allows innately prosocials and cooperative individuals to cooperate reciprocally rather than having their cooperation dragged down by defectors. Similarly, exclusionary mechanisms also reduce the tendency for cooperators to resort to defection when being exposed to uncooperative behavior (since relationships with defectors are absent). Still, exclusion for the above reasons does not unveil a paradox: Does who do not want to cooperate are “kicked out.”

The second goal of exclusionary mechanisms creates a paradox: Exclusion – or the threat thereof – can bring defectors back in line and thereby increase cooperation rates. Yet, exclusion may prevent defectors from adopting a cooperative stance due to the absence of positive role models, that is, cooperators. Although exclusionary mechanisms shield cooperators from temptations to resort to defection, owing to their insulation from negative influences from defectors, the inverse holds for defectors who are insulated from influences to change their ways. We assume here that (a priori unknown) defectors sometimes may be willing to cooperate, after all, given sufficient nudging or graduated sanctioning (Tversky & Kahneman, 1974; Ostrom, 2009). For example, graduated sanctions – i.e., not punishing severely in the beginning but incrementally increasing punishments if one defects more systematically – work to foster a cooperative environment (Ostrom, 2009). Also, think of Axelrod's (1984) strategies. The criterion "forgiveness" in tit-for-tat helps a strategy be successful in realizing cooperation in an environment where occasional defectors are willing to change their way when they face retaliation. "Grim trigger" instead is a strategy that is unequivocally harsh, unforgiving, and final and – as Axelrod argues – fares less well because it is incapable of "repairing" relations damaged by occasional defection. If a non-exclusionary mechanism, or perhaps even temporary exclusion, would help (prospective) defectors to avoid defecting, then we find ourselves in a paradox. Higher levels of cooperation were possible, and cooperative potential was lost. Note that steadfast defectors are better off being excluded immediately: If teaching defectors to cooperate makes cooperators more likely to defect, the collective as a whole and cooperators, in particular, may end up worse by non-exclusion than by exclusion. In this dissertation, I identify how standard exclusion mechanisms may not create opportunities for defectors to change their ways, and the collective may be worse off in terms of cooperation achieved.

It is important to note that the classical approach to exclusion centers on *excluding defectors* ("out") and *including cooperators* ("in"). Exclusion thus comes along with a selection-of-cooperators process. Cooperators are more likely selected as partner whereas defectors are not. Here, I show that exclusion may have unintended consequences and, at times, lead to excluding innately cooperative types (such a process does not benefit the collective) or potential defectors who need a little nudging to cooperation. Furthermore, I point to the paradox of exclusion in which exclusion increases the chances that defectors fail to become cooperators again rather than deterring them from persistent defection. A consequence of this is that the exclusionary mechanism intended to promote cooperation may do the opposite.

For example, if a student contemplates defecting, the student is ideally motivated to cooperate to avoid being excluded. So, this (threat of) exclusion forces a (potential) defector – or defectors in general – to choose: Either (i) they mend their ways to be considered for a cooperative group or network relation in the future, or (ii) they do not and remain excluded and are forced to connect to similarly defecting others or remain isolated. The paradoxical nature of exclusion becomes apparent if defectors go for the latter. An exclusionary mechanism can backfire if there is a chance that defectors could mend their ways under the threat of exclusion or after being excluded but fail to do so because they connect to, for example, like-minded others. Then, cooperation levels are less-than-ideal because defectors keep each other from cooperating. Resulting in the loss of cooperative potential. The story of exclusion is thus not all that clear. Simply put, the mere option of an "out" tells half the story. In this dissertation, I investigate under which conditions cooperation can arise and be sustained through exclusion but also charter when exclusion fails to do what it intends to; that is, to promote cooperation.

Finally, thwarting the contagion of undesirable behavior – such as criminal behavior or defection – surfaces as an additional rationale for the practice of exclusion: Separating the “good” from the “bad.” By drawing parallels between exclusionary mechanisms and the concept of imprisonment, we elucidate the connection between protecting the community, fair retribution, and preventing the spread of criminal behavior on the one hand and reformative motives (“rehabilitation”) on the other. This, in turn, reveals a linkage with historical debates encompassing the tenets of protecting the collective versus the objectives of rehabilitation of the defectors. The issue is not only the fairness of exclusion, but also the protection exclusionary mechanisms provide cooperators from exploitation. In that sense, there is a clear parallel with prison sentences and other legal punishments. To what extent is punishment “reformatory,” and to what extent is it “retributive?” Both rationales for exclusionary punishments exist. Ultimately, this dissertation steers away from punishments in a legal manner and navigates this paradox of exclusion in informal contexts.

1.1.1 Building toward the main research question

I focus on exclusionary mechanisms for three reasons: (*i*) the boundaries under which exclusion promotes cooperation are not well known; (*ii*) the relationship between exclusion and inequality – a prevailing, dominant, and very important societal issue – makes it a topical issue; and (*iii*) there is much unknown about whether, how, why, and under which conditions exclusion works in empirically realistic settings as a solution to the problem of cooperation. Exclusion is mostly studied in highly controlled settings.

As a premise for the three reasons mentioned above, I first want to highlight that I build on the principle of decreasing abstraction (Lindenberg, 1992). The method of decreasing abstraction entails first beginning with a very simple, tractable, and unrealistic model to understand the mechanism under consideration. The next step in this process is to incorporate empirical relatedness. At the same time, explaining the model and mechanisms giving rise to the outcome remains essential, even with added complexity. In my view, I extend existing research into exclusion and cooperation and move the pointer more toward empirical relatability instead of increasing abstraction. I thus aim to maintain relatedness to the empirical context while guarding analytical tractability, as I show in this dissertation in various chapters. To be clear, this dissertation utilizes the insights that the method of decreasing abstraction provides. I stress that the studies conducted in this dissertation are already a decreasing abstraction step in this line of research.

1.1.1.1 Exclusion affects the scope of other mechanisms promoting cooperation

The first reason is that there are already many heralded mechanisms that catalyze cooperation: Think of the role of reputations, norms, signals, communication, and sanctions (Axelrod, 1984; Baldassari, 2015; Fehr & Gintis, 2007; Rand & Nowak, 2013; Simpson & Willer, 2015). They all give rise to cooperation under certain conditions. But the problem with those mechanisms is that exclusion affects their effectiveness. Present relationships between individuals in networks or groups play a crucial role in facilitating the mechanisms mentioned here (Simpson & Willer, 2015). Yet, the scope of said mechanisms is limited if a connection is absent, for example, due to exclusion. For example, a relationship must be present if peers want to sanction each other for their defection informally. Likewise, reputations are more readily established when individuals are connected to others, either directly or indirectly. Studying the effectiveness of said mechanisms relies on present relationships between individuals, but those ties may be absent due to exclusion. The mere possibility of engaging in exclusion thus has effects beyond the scope of exclusion itself. Hence, the focus on exclusion is motivated by the dynamic nature of real-life cooperative endeavors. The mechanisms listed earlier may fall flat in dynamic circumstances. Given that we know that many groups and networks are dynamic, I first need to understand the scope of these dynamics. The next step is inspecting how exclusion and other mechanisms aiming to promote cooperation – think of sanctions or communication schemes – in conjugate can foster cooperation. I leave this question for future research and point to the need to understand how exclusionary processes affect cooperation levels.

1.1.1.2 Inequality and exclusion: Do they go hand in hand?

The second reason why I study exclusion relates to the classical problem of the pervasiveness of inequality in society. Sociological research covers a myriad of societal problems, comprising, for example, crime, corruption, gender inequality, "culture wars," polarization, and secularization. The seeds of many societal problems can be found in inequality. Too much inequality is generally considered dysfunctional for society (Piketty, 2014; Scheffer et al., 2017; Stiglitz, 2012; Wilkinson & Pickett, 2009). A consequence of exclusionary mechanisms in groups and networks is that they may foster persistent inequality between cooperators and defectors. On the one hand, most cooperators connected to similar others may benefit from their cooperative efforts, while other, more defecting types, do not. As benefits from cooperation accrue to mostly cooperators who exclude their defecting counterparts, defectors may fall behind. As such, one "bad" choice may lead to systemic inequality between cooperation and rehabilitated defectors. Namely, defectors cannot receive the benefits of past cooperative endeavors. On the other hand, defection could also be a reason for the continuation of inequality if defectors "get away" with the benefits they reap from exploiting the effort put in by cooperators. But if exclusion puts defectors in their cocoon where defection prevails, wielding them back into cooperation may be hard. As such, the paradoxical nature of exclusion can have a detrimental effect on the overall cooperative environment. It is thus important to consider the positive and negative societal effects of exclusion when studying the dynamics of cooperation. This dissertation studies under which conditions exclusionary mechanisms, active in networks and groups, promote unequal access to benefits arising out of cooperation.

1.1.1.3 Lack of research on exclusion in less abstract environments

The third reason is that exclusionary mechanisms surround us since we are all embedded in one or many dynamic environments. Yet, exclusion is mostly studied in controlled settings even though the mechanism is often implemented. Examples of exclusion include school admissions procedures, selection in teams, trials in US sports, social mobility, the bi-annual transfer market in football, employee recruitment, and organization project team formation. These examples show that individuals are rewarded for their proven abilities and past achievements, the reward being access to better jobs, groups, organizations, or schools. If you cooperate, you are included with other cooperators and separated from defectors. But if you defect, you are excluded from access to future cooperative endeavors. From the point of view of cooperators, ideally, the mere threat of exclusion motivates defectors to cooperate. But exclusion can benefit a defector in different ways. When defectors interact with each other, they do not face social punishments for free-riding. Free-loaders may find their footing, perhaps unrelated to the cooperation problem all face.

It may even be that defectors provide each other with social support that interactions with cooperators cannot provide (for a theoretical example showing how defectors provide each other with the support cooperators refuse to give, see Flache & Macy, 1996). For example, a defector not investing time and effort in a school assignment may be considered "cool" by fellow defectors and be rewarded status benefits accordingly. Following the classical work by Merton, this type of behavior can be considered "innovation" or "rebellion." Both behavioral strategies build on the notion that behavior deviates from what is considered the norm. Defectors displaying support for similar others may be aware of the norm to cooperate but find ways to abstain from norm-abiding behavior. In such instances, one can think of defecting also as a "counter-culture" (e.g., Yinger, 1982) in which behavior is intended to deviate from the norm. For example, adolescents may revolt against cooperative norms set by their parents and teachers and express this by not doing homework or drinking.

Yet, in scientific research on cooperation, the power of (the threat of) exclusion – though based on real-world phenomena – is not assessed in real-life situations in which many, simultaneously operating mechanisms are present. This is not necessarily a problem, as it relates to experimental control over the mechanism of interest that is expected to promote cooperation. However, if this control is achieved at the expense of excluding other relevant explanatory mechanisms, it does become a problem. Notably, the effect of exclusionary mechanisms on cooperation may be distorted by other processes known to affect people's behavior and relational choices. I highlight two.

First, research shows that decision-making is not always rational due to factors such as limited information and bounded rationality (Macy & Flache, 1995; Simon, 1982; Udehn, 2001; Wittek et al., 2013). Thus, people do not always behave rationally; they make errors and do not always choose the best course of action, even by their lights. This can also impact decisions about whether to cooperate or defect in a given situation. For instance, individuals may not always cooperate even when it is in their best interest to do so because they do not have complete information about the situation or cannot accurately assess the consequences of their actions. To illustrate this, imagine a group project in school where each student is assigned a task and is expected to contribute equally to the final product. One student may not have all the information they need to make an informed decision (e.g., lack of clarity about the project's requirements), leading them to make a contribution that is not optimal for the group. Another student may have limited cognitive resources to deal with the choice at hand and struggle to make a decision, leading them to follow the crowd rather than make an independent choice. Also, some need to learn that cooperation is in their best interest over time and may thus need time to cooperate consistently.

Especially the interaction between decision-making and exclusion renders the control for the decision to cooperate or defect – accounting for irrationality – necessary. Building on the example of students working together in a group, someone may be excluded from the group for not cooperating even though one was, for example, unaware of the requirements, simply "tired," or did not have time to alter one's initial defection. But it is problematic if the process leading up to exclusion is based on a decision affected by limited information. It can be that someone has a wait-and-see approach, do what their friends do, or need more time to learn to cooperate. This example shows that exclusion based on limited information in decision-making can be inefficient and lead to suboptimal outcomes. Ideally, exclusion builds on clear information in which all involved know the potential exclusionary consequences of defection. Yet, that is not the case in reality. I include the limits of decision-making when studying the effectiveness of exclusion.

Second, experiments and theoretical models generally do not include many social processes. The reasons for not doing this are clear: It allows for analytical tractability so that the research can explicate clearly how, in this case, exclusion promotes (or not) cooperation. But this is problematic for cooperation research if one aims to further empirical relatedness in which various mechanisms are present simultaneously. Studying whether exclusion still promotes cooperation if other features simultaneously affect the outcome of interest requires the step I take here.

Examples of social processes that interact with the mechanism of exclusion include the tendencies of individuals in networks to form reciprocal ("you are my friend, and I am your friend") or transitive ("a friend of a friend will most likely become my friend") relations (Kadushin, 2012). For example, relational reciprocity may undercut exclusion if a defector befriends a cooperator. Relational reciprocity may thus make for tolerating defectors in a group, rendering the exclusionary sword blunt. Similarly, the empirically documented tendency to pursue transitivity in social relations may spur network relations between cooperators and defectors who are friends of the cooperator's friend, again countering exclusion. A potential consequence of including reciprocity and transitivity considerations into the mixture may be that exclusion is not an ideal mechanism to promote cooperation. At the same time, if relational reciprocity further strengthens the main aim of exclusion – i.e., promoting cooperation – then including such "other" considerations is crucial. Thus, multiple reasons exist to form or sever relations in real life that interact with the exclusion mechanism in determining cooperation decisions.

1.1.1.4 Formulating the main research question

I thus identify important behavioral and social mechanisms that drive group and network formation that are theoretically expected to interfere with the effectiveness of the exclusion mechanism in cooperation settings but that have not been part of existing studies into this effectiveness due to experimental or theoretical control. I formulate the following research question to decrease this abstraction and incorporate the empirical context in which exclusion occurs.

> **Main research question**: *Can exclusionary mechanisms promote cooperation and if so, how, and under which conditions*?

To answer the main research question, we need to scrutinize the role of *information* actors facing a cooperation choice have about potential partners. Selection or avoidance are decisions based on available information. If one has no information concerning the others' previous dealings (either via direct contact or hearsay) or "inner workings," then exclusion may be unwarranted and perhaps should be avoided. In such instances, the threat of exclusion may send a signal that deviant behavior is unwanted. But for exclusion to work as intended along the way, individuals need information. For this reason, this research focuses on *two tiers* related to the importance of information from the fields of psychology and sociology. I also believe that research into cooperation benefits, in general, from an interdisciplinary approach. **Tier 1** stresses the importance of information on stable personality characteristics related to cooperation for separating cooperators and defectors (*psychology*); **Tier 2** points to the importance of information about observed behavior and the context in which exclusion operates (*sociology*). I argue that information from both tiers must be utilized to make inferences about how cooperative someone might be.

1.1.2 The five sub-research questions

Following the argument of Tier 1, a stable measure of prosociality may indicate whether purported behavior is also relatively stable. Researchers should be aware of whether the measure they rely on captures a stable trait, especially if the stability of prosociality is crucial in studying exclusion. If the measure is not, research on exclusionary mechanisms that rely on stable prosociality tendencies may fail to promote cooperation effectively. An unstable cooperative orientation may indicate whether one more easily flip-flops in behavior as a response to, for example, social influence. To be clear, "flip-flopping" means changing from cooperation to defection and vice versa. Using information about stable personality characteristics related to cooperation to decide whom to exclude, increases the likelihood that purported cooperators are not mismatched with purported defectors. That leads me to the following sub-question that sets the goalposts in Chapter 2:

> **Research question 1**: *In what way are personality traits – on which someone could base an exclusion choice – of potential cooperation partners stable over time*? Whether an individual is more prosocial or egoistic may become apparent via telltale behavioral signs or a measure capturing this. I dive into this question in Chapter 2, using representative panel data of six repeated measurements.

Tier 2 sets the importance of context and behavior forth when engaging in exclusion. Information on past behavior comprises what someone did in the past, how they behaved in prior encounters with others, and where and with whom they interacted previously. Exclusion mechanisms studied in this research include decisions about the inclusion of potential new and exclusion of potential current interaction partners. Ideally, cooperators exclude defectors and include similar cooperators. But if such information on individual behavior is unavailable, unreliable, or imperfect, it is difficult to decide about exclusion properly. For example, not knowing whether the other is a cooperator or defector makes it impossible to differentiate between types, and a mismatch in needs and interests may occur. I formulate the following research question that is addressed in Chapter 3.

> **Research question 2**: *In what way does imperfect information about others' behavior impact the effectiveness of exclusionary mechanisms in groups? Does exclusion then still work as a solution for the problem of cooperation?* I start answering this question with a theoretical investigation in Chapter 3, where I develop and apply a computational theoretical model of the effects of exclusion under imperfect information conditions.

Chapter 3 points to exclusion in networks as a solution for the problem of cooperation and highlights the detrimental consequences group contexts can have on the effectiveness of exclusion in groups. Yet, the answers given in Chapter 3 are based on a theoretical modeling study that excludes (no pun intended) many complexities that may interfere with exclusionary mechanisms in cooperation problems as they occur in real-life situations.

In the real world, there are many other possible individual and social characteristics on which individuals could select potential cooperation partners. Research question 3 focuses on this. For instance, are friends more likely to form a cooperative relationship? Does one prefer to cooperate with the same-gender, or does an individual form a cooperation relationship with others with a similar personality trait? If there are preferences to form cooperation relations with certain others, the consequences are that dissimilar others are excluded. We explore the impact of various features on forming relations with cooperative others empirically in Chapter 4.

> **Research question 3**: *Do we observe exclusion based on cooperative considerations in social networks "in the wild" when also other network mechanisms play a role*? I study the role of exclusion empirically in Chapter 4, using a higher education setting where students can benefit from their academic progress by forming cooperative relations amongst each other. The data comprises information on friendship and cooperation network relations that students utilize to work together and individual characteristics such as gender, grades, and personality traits affecting those with whom one works together.

Chapter 4 primarily investigates empirically whether students' cooperation relationships are affected by exclusion dynamics. These dynamics include students' rejection of certain peers for working together while preferring to collaborate with others.

However, social influence may interfere with exclusionary mechanisms in the real world, distorting the effect of exclusion on promoting cooperation. The process of social influence assumes that individuals assimilate to their peers' behavior. For example, peers actively influence individuals via social learning, imitation, or tit-for-tat dynamics. Exclusion may halt the impact of social influence by isolating the defectors from "outside" influences aimed at promoting cooperation. That is why we extend the research reported in Chapter 4 with an empirically-calibrated model in Chapter 5 to investigate whether and to what degree social influence interferes with exclusionary processes in networks. An empirically-calibrated model maintains some relatedness to the empirical situation studied in Chapter 4 but allows me to explore "what if" questions that apply to situations beyond the scope of Chapter 4. An example of a "what if" question is whether exclusion still works if social influence is strongly present (or mildly present). I formulate the following sub-research question:

> **Research question 4**: *In what way would the level of cooperation in an empirically realistic setting be affected if actors' relational choices are more affected by exclusion*? *Or is actors' cooperation more affected by influence from peers*? In Chapter 5, I take a step back from the observed empirical context in Chapter 4 and rely on empirically-calibrated simulations as a tool and data to answer RQ 4.

So far, most researchers studied empirically only situations where cooperation partners are selected dyadically (Chapters 4 and 5; or theoretically in the network part in Chapter 3). Yet, we know that individuals must cooperate in groups or teams, as studied theoretically in Chapter 3. Especially the composition of these teams is an important predictor of their success in achieving a valuable collective outcome. The problem is that few studied self-organized groups or teams, and little is known about whether individuals want to form groups based on how cooperative potential team mates are.

Chapter 6 tests the assumption that students aim to be part of teams with friends, familiar others, and similarly cooperative others. Again, I build on insights into the higher education setting and incorporate multiple features, such as gender, grades, and knowing one another via prior teamwork that may similarly affect ending up in the same team. Insights comprise the tendency to form teams with same-gender others or seek teammates with similar grades. The approach I set out in Chapter 6 applies to contexts in which teams are formed by the constituents themselves.

> **Research question 5**: *In what way does cooperativeness affect preferences for team compositions among students in a setting where real-life features such as friendships, gender, grades, and familiarity also affect team formation*? Chapter 6

uses empirical data to investigate if the exclusionary mechanisms postulated theoretically in Chapter 3 occur empirically in such a setting.

1.1.3 Moving forward

To answer the research questions, I included insights from psychology and sociology, various data, and tools to provide a more comprehensive picture of exclusion and cooperation in the "real world" than current research solely relying on either experiments or simulations. My approach stresses the importance of information (*i*) on stable personality characteristics related to cooperation for separating cooperators and defectors and (*ii*) about observed behavior and the context in which exclusion operates. Exclusionary mechanisms utilize changing, dynamic circumstances we observe in real life as they explicitly build on them; that is, dynamics – i.e., changing network and group relations – are required for exclusion to work as intended. I show that exclusion is a solution for the problem of cooperation in some conditions, mainly if reliable information is available. But even if "good" information is available, the many behavioral and social mechanisms active in real life can distort whether exclusion or the many other processes that affect people's behavior and relational choices promote cooperation. I show that an integrative approach – combining insights from personality and sociology research – provides more information on the effectiveness of exclusion in promoting cooperation.

Main research question: *Can exclusionary mechanisms promote cooperation and if so, how, and under which conditions?*

This dissertation indicates that contextual features remain crucial for cooperation to be sustained in the long haul since individually stable traits related to cooperation alone are not enough for cooperation to thrive and sustain. The exclusion of defectors works on the collective level, increasing overall cooperation levels. But exclusion can backfire. As a consequence of exclusionary mechanisms, non-cooperative actors have fewer encounters with more cooperative ones, diminishing their chances to learn to cooperate through those encounters. The multiplexity (overlap) of networks and groups and the spillover of information from one context to another creates segregation patterns in which cooperators and defectors interact with similar others in the network and group. These segregation patterns can spillover from groups to networks and networks to groups via exclusionary mechanisms.

1.1.4 Data and methodology

I use a triangulation of methods and data to answer the research questions. Chapters 2, 4, and 6 rely on empirical data to answer the RQs, whereas Chapters 3 and 5 rely on simulation models as a tool to answer the research questions. All scripts to analyze the data and simulation code are available online on the Open Science Framework (OSF). The URLs to the OSF folders are listed in each chapter.

First, we need a repeated measures design to answer research question 1 (Chapter 2) because only a longitudinal data design allows us to assess whether individuals on multiple occasions are stable in their personality traits related to cooperation. Chapter 2 builds on a 6-month repeated measures design in which a representative sample of the Dutch population is used to investigate the stability of personality traits related to cooperation. The methods I employ are well-suited to analyze nested data with multiple observations nested within individuals.

Second, I utilize a longitudinal field study of students in the context of a higher education institution (Chapters 4–6). This data and methodology allow me to answer research questions 3, 4, and 5. I use this context to assess cooperation in an empirically realistic environment in which cooperation is needed. The data entails 95 first-year university students (Brouwer et al., 2018), which allows me to test whether network and group dynamics affect cooperation among innately different students. The data contains information on friendship and cooperation relations, personality traits, gender, grades, and project team memberships. Each academic year is divided into two semesters in which students are enrolled in multiple compulsory courses and projects. Data was collected from three waves, one at the start of the year, one in semester 1, and the third in semester 2.

Third, I rely on stochastic actor-orientated models (SAOMs) to analyze longitudinal network data in Chapters 4 and 5 (Snijders, 2017; Brouwer & de Matos Fernandes, 2023 Snijders et al. 2010, 2013). SAOMs represent an essential methodological breakthrough in modeling the co-evolution of networks and behavior (such as cooperation). SAOMs are utilized to answer research questions 3 and 4 in Chapters 4 and 5, respectively. The SAOM framework builds on agent-based computational modeling (ABCM; e.g., Snijders and Steglich 2015) and resembles a multinomial logit model. SAOMs are necessary because they allow analyzing what is commonly known as network autocorrelation or the homogeneity bias ("actors who are connected tend to be more similar on certain dimensions than actors who are not so well connected"). On the one hand, exclusion and inclusion of others (selection in SAOM terms) build on network relations as the dynamic component, meaning that the dependent variable is the change (or not) in network connections. On the other hand, social influence (the same definition in SAOMs) builds on behavior as the dynamic component, meaning that behavior is the dependent variable in the model. Network autocorrelation can be the result of both processes. One needs a tool to control for network autocorrelation when assessing the contribution of different processes in explaining observed co-evolution dynamics. SAOMs is such a tool. SAOMs can simultaneously assess exclusion and influence effects and account for both mechanisms' inherent interdependency. Therefore, I need to use SAOMs to answer research questions 3 and 4, accounting for other real-life regularities and individual characteristics that affect behavioral and relational choices.

Fourth, for Chapter 6, I analyze empirical project team data using exponential random partitions modeling (ERPM) (Hoffman et al., 2023)—a recently developed method. Here, I define a partition as a set of teams. I study project teams, a fixed set of students who must form an exclusive team to work together on a project. Current network models are insufficient to analyze team data because team relations are not dyadic but with a whole team. Team relations do not overlap as students are part of one team and not multiple. Therefore, I need to use ERPMs to answer research question 5.

Fifth, ABCMs are required as a tool in Chapters 3 (RQ 2) and 5 (RQ 4) because they allow for the simulation of complex systems and emergent properties beyond the scope of the empirical situation in Chapter 4. ABCM allows researchers to analyze macro consequences based on assumptions about micro-level interactions between agents (Bianchi & Squazzoni, 2015; Flache & de Matos Fernandes, 2021). In ABCM, the modeler translates the conceptual idea into a formal model and then computer code, explicating how agents behave according to the researcher's theory (e.g., their cognitive abilities), formalizing the context in which agents are embedded (e.g., their group or network), and how behavior and the context influence each other. The empirical data I have has constraints; it does not allow for manipulating social structural conditions, such as the network structure, group compositions, prevalence of exclusionary mechanisms, strength of social influence, and the number of waves beyond empirically measured. In Chapter 3, I implement conditions in which agents either have perfect information (or not). Using the ABCM, I answer research question 2 and analyze the conditions under which imperfect information about individual cooperation jeopardizes the effectiveness of exclusionary mechanisms. In Chapter 5, I move one step further in the use of ABCM and combine empirical data obtained in the field studies of Chapter 4 with a theoretical ABCM. I model behavior based on behavioral and relational choices measured empirically. I apply theoretically different conditions – a stronger or less strong tendency for exclusion and social influence – unable to estimate empirically. This approach allows us to answer research question 4.

1.1.5 Roadmap for the rest of this chapter

In what follows, I first map the current state of the cooperation literature in section 1.2. In sections 1.3 and 1.4, I highlight the two-tier approach taken in this dissertation. In section 1.5, I briefly discuss the theoretical framework used. I give a brief overview in section 1.6 on what will come on a chapter-by-chapter basis. In sections 1.7 and 1.8, I provide some concluding remarks, discuss advances in the literature, and ponder where we should go from here.

1.2 THE STORY OF COOPERATION SO FAR

Contemporary social life is interwoven with situations in which students, community members, activists, employees, athletes, and scholars must effectively work together to realize benefits they cannot achieve by themselves. Examples of such situations include combatting climate change, organizing informal care, engaging in business agreements, mobilizing movements, realizing productive workplace collaborations, working with fellow students, or realizing local renewable energy initiatives. In all examples, producing collective benefits is an enterprise where constituents share responsibilities for success. A successful outcome is uncertain if not all pull their weight and cooperate (Hardin, 1968; Komorita & Parks, 1996; Olson, 1965). Cooperation is thus of fundamental importance for a thriving collective. The clear joint advantage of cooperation would suggest that there is no other way for people than to cooperate. Or is there?

No, we must consider the problem that all face a social dilemma: the optimal individual action (defection) is not the collectively desirable outcome (cooperate) (Dawes, 1980). There is thus a mismatch between individual and collective interests: i.e., “social dilemmas are situations in which individual rationality leads to collective irrationality” (Kollock, 1998, p. 183). If students work together on a project, it is costly to cooperate. Costs comprise time and effort, precluding investments in other personal or social endeavors. It is individually beneficial if others bear this cost: e.g., “I do not have to write an introduction if others do so.” The problem is that many hope that others do the work. The collective fares best when all bring their cooperative A-game to the table. However, when many have a wait-and-see approach, alas, the collective founders (Kollock, 1998; Ledyard, 1995). Cooperation is thus not self-evident since the battle between individual and collective interests may hinder the success of collectives. Therefore, many scientists have confronted the problem of cooperation in many disciplines (Apicella & Silk, 2019; Bianchi & Squazzoni, 2015; Nowak, 2006; Olson, 1965; Ostrom, 1990; Perc et al., 2017; Rand & Nowak, 2013; Van Lange et al., 2013), questioning under which conditions collectives succeed in getting cooperation going.

Fortunately, the story of cooperation is not that bleak. In the movie *The Dark Knight*, Batman (Christian Bale) is on-and-off fighting with the Joker (the late Heath Ledger) in a 20-story building. Batman is unaware that the Joker rigged two fully-crowded ferries with bombs. For simplicity, let me label them ferry A and B. The detonators are swapped, meaning the people on ferry A can detonate the bombs on ferry B and vice versa. To defect (cooperate) is (not) detonating the explosives on the other ferry. To survive, people have two options: (1) Press the button or (2) not, hoping that the other ferry also does not press the button. The risk is that the other ferry presses the button first. The Joker intends to show Batman that people are innately self-interested and always defect. Ultimately, people on both ferries cooperate. Although each individual on ferry A and B has a rational incentive to defect to survive, they choose the collectively most beneficial action of cooperation. The Joker thus fails to show that individual rationality trumps collective rationality. This example is somewhat extreme but attests to individuals cooperating in social dilemmas.

Individuals overcome social dilemmas in both everyday mundane and life-changing settings (Apicella & Silk, 2019; Attari et al., 2014; Baldassari, 2015; Kraft-Todd et al., 2015; Van Lange & Rand, 2022). For example, we combat climate change—a social dilemma in which each country and individual faces the tension between individual and collective interests (Ostrom, 2010), demonstrations to signal societal issues are frequent phenomena (Greijdanus et al., 2020; Opp & Gern, 1993), local renewable energy initiatives in communities are often realized (Goedkoop et al., 2021), and high vaccination rates are reached (Korn et al., 2020). As mentioned earlier, multiple mechanisms are fundamental for cooperation to arise and keep going.

I study exclusion as a mechanism to promote cooperation because present relationships are prerequisites for many mechanisms to work. In doing so, I build on two tiers. This approach is nicely introduced in an *Annual Review of Sociology* piece by Simpson and Willer (2015, p. 44):

> "*Whereas these literatures [psychology, personality, behavioral economics, and evolutionary biology] typically locate the sources of cooperation and prosocial behavior within individuals—in personality, emotions, motivations, and preferences—sociological work views cooperation and prosocial behavior as heavily impacted by factors outside individuals*."

Tier 1 contains factors *within* individuals, while Tier 2 builds on factors *outside* individuals. Past decades have seen a proliferation of research on the understanding features promoting cooperation. Within-individual features comprise personality traits, social preferences, and social value orientations, showing that some are innately more likely to cooperate than others (Thielmann et al., 2020; Van Lange et al., 2014). Exclusionary mechanisms refer to social features that lie outside individuals, allowing them to separate defectors from cooperators (Guido et al., 2019; Gunnthorsdottir et al., 2007). In what follows, I provide more in-depth information on the two-tier approach I take in this dissertation.

1.3 TIER 1: STABLE INDIVIDUAL DIFFERENCES IN COOPERATIVENESS

Although most traditional game-theoretically inspired experimental and classical rational choice research treats individuals as inherently selfish and rational (Camerer, 2003; Elster, 1986; von Neumann & Morgenstern, 1944), we know from decades of research that individuals are not alike in their cooperativeness. Previous research shows which individuals generally, albeit conditionally, cooperate; that is, prosocial types (Balliet et al., 2009; Fehr & Gintis, 2007; Fischbacher et al., 2001; Pletzer et al., 2018; Van Lange et al., 2014; Wittek & Bekkers, 2015). The first step is retrieving information from individuals showing whether one is prosocial or not. The second step is investigating whether one is stable in their prosociality. Both steps are discussed next.

1.3.1 Finding out if one is genuinely cooperative

There are multiple ways to infer whether an individual is innately cooperative. One can think of resorting to information on prior behavior or reputations (Rossetti et al., 2022; Simpson & Willer, 2015; Takács et al., 2021), but this is then the product of the context in which behavior or reputation is realized. If the context is largely defecting, we still do not know whether one is truly a defector or simply adhering to the norm set by the environment. This argument holds for actors involved in the cooperation problem and researchers studying cooperation. Contrarily, personality research attempts to explain why some behave differently than others in multiple situations. Various scales to measure personality are available in the literature (DeYoung, 2015;

McCrae & John, 1992; Murphy & Ackermann, 2014; Thielmann et al., 2020; Van Lange et al., 2014).[1] The all-encompassing conclusion of research into personality and cooperation is that personality matters: People differ in their inclination to cooperate.

I study a personality construct widely used in the cooperation literature: social value orientation (SVO). SVO theory assumes that prosocial types are more prone to cooperate as they generally assign more value to collectively beneficial outcomes than proselfs (Bakker & Dijkstra, 2021; Balliet et al., 2009; Murphy et al., 2011). This tool aids researchers in capturing innate features related explicitly to cooperation. SVO is purportedly a trait. However, SVO must be stable over time to qualify as a trait. The empirical evidence for the stability of SVO is currently scant. Therefore, I assess whether measures of SVO capture a stable trait in Chapter 2, potentially serving as a reliable indicator of the "individual" during exclusionary practices in groups and networks.

Research question 1: *In what way are personality traits – on which someone could base an exclusion choice – of potential cooperation partners stable over time?*

Using several statistical indicators and visualizations, the findings in Chapter 2 suggest that most are stable in their SVO over 6 months. The results also show that there are stable differences between cooperative and defecting individuals: Some are just more prosocial than others.

The answer to research question 1, based on the results in Chapter 2, is yes: The personality traits of potential cooperation partners are relatively stable over time.

1.3.2 Stability in prosociality is not enough for sustainable cooperation

Imagine a group of eight individuals with stable individual differences in prosociality. This may be a group of students studying together for an exam, scholars drafting a paper, or employees working on a project. A person is either prosocial or proself-oriented and stable in their SVO. Now, what will happen? A simple thought experiment leads to the expectation that more cooperation is achieved when more prosocials are in the group. This intuition builds on findings from decades of research: Groups with more cooperative types achieve more easily collective success than groups with more defecting-oriented types (Chaudhuri, 2011; Ledyard, 1995). Yet, solely focusing on individuals and their inner workings is not enough for cooperation to sustain in the long haul because cooperation deteriorates over time. After all, people tend to condition their behavior based on what others do (Chaudhuri, 2011; Fischbacher et al., 2001; Ledyard, 1995; Kollock, 1998). If one or a few start to freeride, even cooperators follow suit. Thus, cooperation or defection does not have to be a reliable indicator of one's personality.

1 The Five-Factor Model (FFM), also known as the Big Five, is a widely used measure and captures a taxonomy of five personality traits (McCrae & John, 1992). Especially *agreeableness* is important because a higher score on agreeableness relates to a higher propensity to cooperate (Thielmann et al., 2020). In the absence of SVO information, I utilize the FFM inventory in Chapter 4 as indicator of personality.

1.4 TIER 2: THE EXCLUSIONARY POWER OF HOMOPHILY

Robinson Crusoe, at one point, lived in a social vacuum. Yet, we are all embedded in multiple social or institutional contexts in which more than one other individual resides. To smoothen cooperation, one needs to be able to separate cooperators from defectors. We need an exclusionary mechanism for this.

Exclusion builds on *homophily* (Blau, 1977; Lazarsfeld & Merton, 1954; Kossinets & Watts, 2009; McPherson & Smith-Lovin, 1987; McPherson et al., 2001). Homophily alludes to the tendency of individuals to bond with similar others. Research shows that these similarities can be, among others, gender, age, religion, ethnicity, socioeconomic status, behavior, attitudes, and beliefs (McPherson et al., 2001). I stress that individuals may prioritize selecting cooperative partners in an environment where cooperation is essential. The expectation that individuals do preferentially select others similar in cooperation is supported by experimental and model-based work (Apicella & Silk, 2019; Fehl et al., 2011; Helbing et al., 2011; Melamed et al., 2018; Perc et al., 2017; Rand & Nowak, 2013; Simpson & Willer, 2015) but not by empirical work (Ehlert et al., 2020; Melamed et al., 2020; Simpson et al., 2014). For example, experiments with university students indicated that university affiliations superseded cooperative reputations for the formation of homophilous relations in a stylized network (Melamed et al., 2020).

There are two reasons why it remains unclear to what extent homophily by cooperation occurs in empirical networks. The first reason is the static nature of measured networks in fieldwork (Apicella et al., 2012; Girard et al., 2015; Simpson et al., 2014). The mere observation of homophily as a static feature does not allow us to conclude whether homophily is stable, persistent, or dominant, nor does it reveal the process through which homophily as an exclusionary process came about. Including network dynamics mirrors what is done in theoretical and experimental work where researchers observe the importance of cooperation homophily for exclusion and inclusion in the long haul.

The second reason is that cooperation matters for homophily in different ways. Behaving and valuing things similarly – such as intentions to pursue the well-being of others, for family, friends, or study mates – leads to a higher likelihood to preferentially form a connection (McPherson et al. 2001). Cooperation homophily, as such, captures the alignment of individuals' interests both socially and in other domains. For cooperators, it may be easier to successfully work together in joint endeavors like studying together if one is matched with a similarly cooperative other. For defectors, homophily may be beneficial because one is not forced to behave in a personally unwanted way, and it may provide support or social approval that a cooperation partner cannot provide a defector (Bianchi et al., 2020; Flache & Macy, 1996). Furthermore, why actors support or punish each other (or not) lies outside the scope of the cooperation problem – and the solution to the problem (exclusion) – I aim to understand. I leave this for now for future research. Even so, whether cooperation homophily affects relational choices has significant implications for the design of exclusionary mechanisms. On the one hand, if individuals tend to select similar peers and group mates, exclusionary mechanisms that target non-cooperative individuals may effectively promote cooperation. On the other hand, if individuals do not prioritize cooperative behavior in selecting partners, alternative mechanisms may be necessary.

Here, I focus on formal groups and informal social networks. We are all embedded in formal (work, school, neighborhood) and informal (friends, acquittances, social media) contexts. This duality is also reflected in cooperation research. Many researchers study in what way groups can produce collective benefits via cooperation (Chaudhuri, 2011), whereas others focus on dynamic networks affecting cooperation levels (Rand & Nowak, 2013; Fehl et al., 2011). Yet, many studies focusing on groups treat these groups as fixed. I argue that both should be treated as dynamic, and both are affected by exclusion.

1.4.1 Homophily in changing groups

No one permanently resides in the same group, and no group configuration is fixed for eternity; we move around and change groups, friends, and jobs. For example, football players move now and then to a different club, employees switch jobs or departments, and project teams break up when their task is finished. The group's composition is dynamic; some leave a team and join others. I use teams and groups interchangeably in what follows.

Meritocratic matching solves the cooperation problem by bolstering cooperators' assortativity (Duca et al., 2018; Gunnthorsdottir et al., 2007; Nax et al., 2018). I analyze meritocratic matching as an exclusionary mechanism for group dynamics in Chapter 3. Ideally, prosocial types can signal to others that they are willing to cooperate, separating them from proselfs who are more prone to defect. Empirically, an implementation of meritocratic matching could be that a study program formally matches cooperative students with similar others, for example, based on evaluations of cooperative behavior in earlier group projects or courses. Yet, following Young's (1957) dystopian classic *The Rise of Meritocracy*, meritocratic systems are criticized since they can perpetuate inequality by shifting it to merit-based inequality (Mijs, 2016; Sandel, 2020). Inequality based on exclusion in a meritocratic system is then viewed as just. However, that may not be the case because, for example, family status, access to education, and wealth play a crucial role in whether one advances on the societal ladder. Chapter 3 considers only cooperation as input for meritocratic matching. This theoretical work contributes to the literature about meritocratic matching by showing under which conditions the success of matching is threatened.

Research question 2: *In what way does imperfect information about others' behavior impact the effectiveness of exclusionary mechanisms in groups? Does exclusion then still work as a solution for the problem of cooperation?*

Using meritocratic matching as exclusionary mechanism for groups in Chapter 3, I extend prior literature and make more realistic assumptions about information and cognition. I uncover a limitation of meritocratic matching: Incomplete information makes prosocials – who we show in Chapter 2 are stable in their orientation – end up not fully exploiting their cooperative potential, hindering cooperation in general. Yet, the spillover of information from networks – in which agents also interact and face cooperation problems – creates conditions under which meritocratic matching can function more as intended in the group context. Pooling information on cooperative behavior from informal social networks and groups advances the functionality of meritocratic matching.

Answering research question 2, imperfect information leads to a exclusionary mechanism that is less able to promote cooperation in the group context, since cooperative potential is lost. Even so, exclusion – albeit in a situation with imperfect information – still promotes cooperation if information from network interactions is included when forming groups.

Whether individuals form groups based on cooperative considerations is investigated using empirical data in Chapter 6, extending the theoretical approach in Chapter 3. Research on how and why specific team configurations emerge is scant (cf., Bailey & Skvoretz, 2017; Kaven et al., 2021). I study the self-selection of students in project teams. Given there are so many individual and social features that could also matter for group formation, does matching based on cooperative consideration happen in empirical settings where there is a clear incentive for cooperators to match with other cooperative types? I include three non-cooperation features in Chapter 6: homophily by grades and gender, familiarity due to prior encounters, and friendships.

Research question 5: *In what way does cooperativeness affect preferences for team compositions among students in a setting where real-life features such as friendships, gender, grades, and familiarity also affect team formation?*

To answer this question, I rely on a novel statistical tool in Chapter 6 using insights from statistical network modeling to model team formation dynamics. Exclusion in Chapter 6 comprises joining one team instead of another, separating students apart based on multiple indicators. Using data from 70 students in higher education, the results indicate that friends as well as same-gender students are more likely to end up in the same team. Cooperative reputations are not a major antecedent to choosing one team over another in this context. Our results suggest that gender homophily and friendships lead to multiplexity across networks and teams, meaning that same-gender students tend to interact with friends within the same team.

Formally answering research question 5: Cooperative considerations do not play a decisive role for students to join one project team over another in the first year in their studies. Non-academic preferences – joining teams with friends and same-gender others – primarily affect team formation.

1.4.2 Separating "friends from foes" in networks is challenging

The second context in which homophily operates as an exclusionary mechanism is informal social networks, primarily friendship relationships. Friendships provide safety, social support, trust, social capital, and linkage across communities (Coleman, 1990; Kadushin, 2002), affecting individual happiness, well-being, life satisfaction, and mental health (Dunbar, 2018). In the workplace, networks – or the lack of network relations – crucially affect what information people have and whether they can create organizational change (Burt, 1992; Granovetter, 1985). In cooperation research, networks allow cooperators to exclude defectors and select similar others. Network relations themselves are then utilized as a homophilous exclusionary mechanism. For example, one can sever the network tie if a network partner behaves differently from you. At the same time, you can scan in your social vicinity for a similar other to form a relationship with and then select the other as a partner. This feature is commonly described as network selection, network reciprocity, or dynamic networks (Fehl et al., 2011; Nowak, 2006; Rand & Nowak, 2013).

Research question 3: *Do we observe exclusion based on cooperative considerations in social networks "in the wild" when also other network mechanisms play a role?*

The "exclusion part" in Chapter 4 is primarily focused on with whom students form cooperation relations. Who remains behind (and are thus excluded) and who not is of interest to me in this study. Understanding the formation of friendship and cooperation relations – in which information and resources are shared – is essential for first-year students who need to adapt to a new environment. Stochastic actor-oriented model analyses of 95 students in higher education show that gender homophily is pervasive in multiplex networks. Multiplex networks capture different types of ties (i.e., friendship and cooperation relations) among the same set of students. A byproduct of being friends is that they are more likely to become cooperation partners and vice versa.

Answering research question 3, we infer that exclusion in terms of having cooperation relations tends to occur as a byproduct of other social mechanisms and relations. Not being friends may lead to not having a cooperation relation and same-gender students are more likely to form cooperation relations, excluding dissimilar others. Thus, the answer to RQ 3 is "yes", we do observe exclusion "in the wild", albeit in terms of having a cooperation relation and not based on cooperative behavior.

Information about individuals' prior cooperation is key for forming network relationships. However, it is difficult to comprehend what others did in the past, especially those with whom one did not interact personally. A possibility to infer someone else's type is reputational information (Nowak & Sigmund, 1998). Cooperative reputations are individually ascribed, flow easily through the network, and convey to a reasonable degree whether one is cooperative or not (Takács et al., 2021). More importantly, reputations solve cooperation problems via so-called chatter or gossip in which information is exchanged about each other (Giardini & Wittek, 2019;

Raub & Weesie, 1990; Wu et al., 2016). You can think of chatter like “that is a good one,” “we worked together in the past, and it went well,” “it is not wise to work together with him or her,” or “you should avoid working with that person.” Generally, cooperating (defecting) leads to a “good” (“bad”) reputation. Past behavior wrapped up in a reputational construct is thus a potential source to separate friends (“good” ones) from foes (“bad” ones). Therefore, reputational information conveys in an imperfect world – in which you cannot interact with everyone and directly experience what type of person the other is – whether someone is a cooperator or defector. In Chapters 3 to 6, I rely on cooperative reputations and relations, allowing the separation between cooperators and defectors in various ways.

1.4.2.1 The impact of social influence

Social influence affects homophily. Social influence means that individuals learn from, conform to, adapt to, or are influenced by friends in terms of how to conduct themselves (Bandura, 1977; Bicchieri, 1990; Cialdini et al., 1991; Coleman, 1990; Ehlert et al., 2020; Friedkin, 1998). Cooperators can influence people around them to follow suit, whereas defectors can do the same. Again, the cooperative potential of some is lost if prosocials and defectors are forced to defect due to social influence. Notably, homophily and social influence give rise to the behavioral similarity among friends (for an overview, see Steglich et al., 2010). Homophily builds on individuals preferentially selecting similar others as network partners, while social influence builds on network relations that are already present. If a network relation is present, social influence potentially counters homophily as an exclusionary process. As such, social influence is included as a mechanism of interest. Chapter 3 considers theoretically the role of agents adjusting their behavior to others in their group. Chapter 5 includes in an empirically-calibrated model the impact of social influence by network partners on cooperative behavior.

Research question 4: *In what way would the level of cooperation in an empirically realistic setting be affected if actors’ relational choices are more affected by exclusion? Or is actors’ cooperation more affected by influence from peers?*

In Chapter 5, I assess the extent to which cooperation homophily (the exclusionary mechanism of interest) and social influence lead to cooperation to “spread”, “segregate” or “die out”. Chapter 3 studied this question theoretically. I take the next step in this chapter and pair homophily and social influence with other behavioral and social mechanisms found empirically. I use data from 95 students in higher education as input for the simulations. The results in Chapter 5 reveal that cooperation benefits the most when homophily by cooperation and social influence are strongly present. Strongly present means in this context that I amplified the tendency to exclude dissimilar others and being socially influenced by peers. An explanation for this finding is that cooperators form local dense clusters, influencing their peers to maintain their cooperation. The model in Chapter 5 – corroborating the findings in Chapter 3 – highlights a downside for defectors: Some remain socially excluded and are not stimulated to cooperate, diminishing their chances to escape from defecting.

To formally answer research question 4, cooperation levels segregated if homophily by cooperation and social influence are strongly present.

1.5 A SONG OF MICRO AND MACRO

The approach set out in this dissertation heavily builds on the macro-micro-macro model in combination with the social-mechanism-perspective (Coleman, 1990; Hedström & Ylikoski, 2010; Schelling, 1978; Lindenberg, 1990; Udehn, 2001). Especially James Coleman (1990) explained using his famous diagram – colloquially known as *Coleman's boat* – how macro-level outcomes depend on micro-level mechanisms. That is, the diagram explains the relationship between social events, facts, and phenomena (macro) via individual decision-making as the unit of analysis (micro) (Figure 1.1).

I explain the macro-micro-macro model succinctly. The initial macro condition imposes constraints and possibilities for individuals (arrow 1 in Figure 1.1). The micro-level playing field of individuals – desires, beliefs, opportunities (Hedström, 2005) – is influenced by groups or networks in which individuals are embedded. SVOs are an example of individual beliefs that influence behavior. Then, the diagram explicates how individual actions are influenced by said constraints and possibilities (arrow 2 in Figure 1.1). In this dissertation, actions, or micro outcomes, comprise cooperation or defection, staying in or leaving a group, and dropping or maintaining a network tie. All these actions are influenced by the desires, beliefs, and opportunities individuals strive for. Arrow 3 in Figure 1.1 eventually captures how the choices of multiple interdependent individuals generate a particular macro-level outcome. For example, emergent outcomes are updated network or group configurations. A network may either be more integrated or segregated, and overall levels of cooperation may have increased or decreased over time. Via modeling arrows 1 to 3, I explain the relationship between macro-level conditions, as visualized via arrow 4 in Figure 1.1.

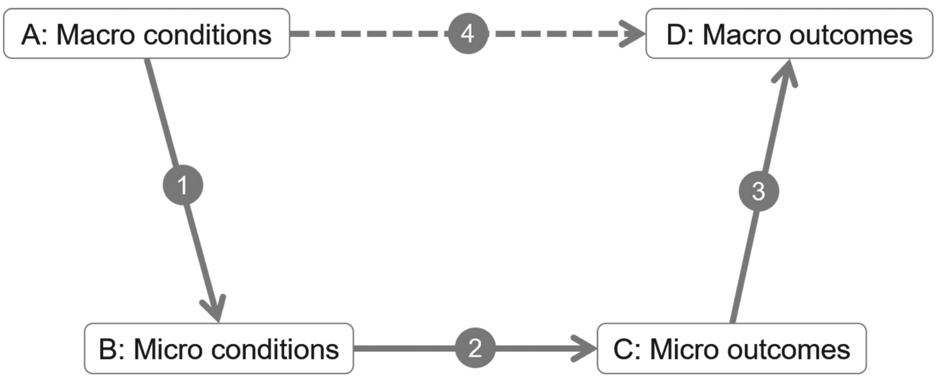


Figure 1.1: The macro-micro-macro topology that is commonly used in sociology to explain macro-level phenomena via the micro-level.

The macro-micro-macro model can be interpreted as a *representative actor story*. Namely, it allows me first to figure out what an actor does in a given social situation, which is, in turn, shaped by macro-level conditions. But there is a problem with the representative actor approach. We know that more individuals are involved, and what others do affects what an actor can do. An example of this process is social influence. Social complexity tells us that the interplay of interactions of many actors may produce outcomes that cannot be readily "aggregated up" from their actions (Flache & de Matos Fernandes, 2021). Macro-level outcomes do thus not simply arise via arrow 3 following a representative actor story but more as a consequence of interdependent actors interacting and influencing each other to behave in a certain way. The macro-level outcome (node D) is based on the actor and context interaction. A quote by Schelling is particularly telling about the importance of understanding this complex micro-macro link.

> "*To understand what kinds of segregation or integration may result from individual choice, we have to look at the processes by which various mixtures and separations are brought about. We have to look at the incentives and the behavior that the incentives motivate, and particularly the way that different individuals comprising the society impinge on each other's choices and react to each other's presence.*" (Schelling, 1978, p. 142)

I discuss two examples showing the need to account for the complexity underlying the macro-micro-macro linkage and representative actor approach. First, Schelling (1971, 1978) showed how residential segregation (macro) arises due to unintended consequences of relocation decisions (micro). The link is as follows. When a resident selects to move somewhere else, that resident influences how satisfied members are in their old and new neighborhoods with the adjusted ethnic composition. Consequently, unhappy residents decide to relocate. Thus, a single relocation decision may lead to a relocation cascade giving rise to residential segregation, even when residents have non-xenophobic preferences. Second, Stadtfeld (2018) exemplified how the prevalence of gender segregation (macro) depends on various interdependent micro-level mechanisms. Stadtfeld considers actors with homophily preferences (*h*) and non-homophilous actors (*nh*). *h* actors form ties with similar others, refusing to connect to dissimilar ones. *nh* actors strive for reciprocated and transitive ties. Given that the *h*'s assort together, *nh* actors seek connections with other available *nh* actors. The macro-level consequence of parallel operating micro-level mechanisms for homophilous and non-homophilous actors is gender segregation. That is why researchers generally need a detailed analysis of the micro-processes and how they relate to macro outcomes.

The content of this dissertation utilizes a combination of the macro-micro-macro approach and complexity perspective as a guiding principle. In Chapter 2, I study the stability of SVO as a micro condition and precursor for cooperation (node B in Figure 1.1). In Chapters 3 to 6, I follow parts of the full macro-micro-macro model envisaged in Figure 1.1. Chapter 3, for example, covers all steps in the model. Based on different initial network and group conditions (node A), I consider the following micro conditions: Social influence in groups and networks, preference for joining cooperative groups, and homophily in networks. The macro-level outcome comprises whether groups and networks are more or less segregated and whether cooperation levels in groups and networks increased or decreased. Chapter 4 takes an empirical network as an initial macro-level condition (node A) and focuses on preferences for homophily as a micro condition and other relational preferences affecting with whom students form network relations. The outcome of interest is how homophily and other features affect relational choices (node C), giving rise to the final network configuration (node D). Chapter 5 takes a slightly different approach as I examine artificially different micro conditions (node B) affecting cooperation and relational choices (node C). Different micro conditions comprise absent or present homophily and social influence (and combinations thereof). Chapter 5 thus allows me to test how micro and macro-level outcomes depend on the relative presence of two intertwined micro-level processes. Chapter 6 explores in what way prior group and network configuration (node A) preferences for joining certain groups (node B) and acting upon those preferences (node C). The outcome of interest is the team composition of project teams (node D).

Table 1.1: I list several unique selling points per chapter.

Number	Chapter	Unique selling point
1.	Chapter 2	The results support classifying SVO as a personality trait as it captures a stable pattern of preferences related to cooperation.
2.	Chapter 2	One should refrain from categorizing SVO scores since imposing boundaries on a continuous scale heavily affects the stability of SVO.
3.	Chapter 3	Information about others' cooperative behavior is vital for promoting cooperation if meritocratic matching is used to form groups.
4.	Chapter 3	Chapter 3 points to the downside of imperfect information. I show theoretically that when prosocial agents escape from uncooperative groups, proselfs have fewer encounters with prosocials, diminishing their chances to learn to cooperate through those encounters.
5.	Chapter 4	Following Chapter 3 showing how information network relations help foster cooperation, I show that students informally mix friendship and cooperation ties: Friends become cooperative partners, and cooperative partners become friends.
6.	Chapter 4	The results indicate that more creative students are more active and popular as network partners than their less creative counterparts.
7.	Chapter 5	Whether cooperation "spreads," "segregates," or "dies out" depends on cooperation homophily and social influence as well as other behavioral and social mechanisms.
8.	Chapter 5	Cooperation flourishes the most – relative to the setting of students in higher education on which I calibrated the model – when cooperation homophily and social influence targeting cooperation strongly affect network formation and behavioral change.
9.	Chapter 6	Friends and same-gender students are likelier to end up in the same team.
10.	Chapter 6	Gender homophily spills over from friendship networks to teams.

1.6 SUMMARY PER CHAPTER

Chapters 2 to 6 consider one or more features pivotal to answer the main research question: *Can exclusionary mechanisms promote cooperation, and if so, how and under which conditions*? I list several key results per chapter in Table 1.1. After that, I briefly summarize the chapter's key components and how they answer the RQs.

1.6.1 Chapter 2: The stability of social value orientations

In what way are personality traits – on which someone could base an exclusion choice – of potential cooperation partners stable over time? In Chapter 2, I study the social value orientation (SVO) slider measure as a key explanatory measure for stability in prosocial preferences. I find that – using panel data with repeated measurements – one's prior SVO is highly indicative of future SVO. The analyses in Chapter 2 validate the slider measure as a reliable scale. SVOs are potentially a valid basis for exclusion.

1.6.2 Chapter 3: A bad barrel spoils a good apple

In what way does imperfect information about others' behavior impact the effectiveness of exclusionary mechanisms in groups? Would it be enough that agents are stable in their prosociality? Does exclusion then still work as a solution for the problem of cooperation? I answer this theoretically in Chapter 3 by studying whether meritocratic matching is less effective when estimations are based on imperfect information. I distinguish between individual-level and group-level information. Prosocials in uncooperative groups are then indistinguishable from proselfs from the same groups, preventing them from accessing cooperative groups. A solution studied in Chapter 3 is pooling information from groups and social networks. Homophilous networks create conditions under which matching can function as intended. There is a downside to this: Chapter 3 shows that when prosocial agents escape from uncooperative groups, proselfs have fewer encounters with prosocials, diminishing their chances to learn to cooperate through those encounters.

1.6.3 Chapter 4: Studying the multiplexity of social life

Do we observe exclusion based on cooperative considerations in social networks "in the wild" when also other network mechanisms play a role? The simulation results in Chapter 3 indicate that exclusion in groups works better if the information on cooperation from networks is utilized during the process. The question is whether individuals incorporate cooperative considerations in forming network relations in an empirical setting (students in higher education). Chapter 4 investigates whether and in what way students form friendship and preference-for-collaboration networks. This time, the exclusion pertains to whether students who are friends also tend to have cooperation ties, excluding non-friends. The network relations themselves are then utilized as an exclusionary mechanism. The empirical results show that friends become cooperation partners (and vice versa). What is more, especially popular cooperators are more likely to become friends. The findings in Chapter 4 indicate that networks are exclusionary for cooperation relations: Becoming a cooperation partner is easier for known others (in this case, friends).

1.6.4 Chapter 5: Cooperation in an artificial world

In what way would the level of cooperation in an empirically realistic setting be affected if actors' relational choices are more affected by exclusion? Or is actors' cooperation more affected by influence from peers? Using an empirically-calibrated model as a follow-up to Chapters 3 and 4, I explore in Chapter 5 "what if" situations in which I vary the strength of homophily by cooperation (exclusion) and social influence. The model in Chapter 5 builds on empirical data as input (friendship network and cooperation data from students), tailoring the model to accommodate contextual relatedness. The empirically-calibrated simulations reveal that cooperation benefits the most when cooperation homophily and social influence are strongly present. Echoing the theoretical findings in Chapter 3, the model in Chapter 5 again highlights a persistent downside for defectors: Some remain socially excluded and not stimulated to cooperate, diminishing their chances to escape from defection.

1.6.5 Chapter 6: There is no I in TEAM, but there is a M-E in there

In what way does cooperativeness affect preferences for team compositions among students in a setting where real-life features such as friendships, gender, grades, and familiarity also affect team formation? Chapter 6 extends the theoretical analysis in Chapter 3 by empirically studying whether students form teams based on cooperative considerations. The exclusionary mechanism studied in Chapter 6 is whether students prefer to join teams with similar others, excluding dissimilar cooperative others. I control for other features – friends, familiar students, preferences for same-gender others, and preferences for students with similar grades – affecting exclusion during team formation. The results indicate that friends and same-gender students tend to join the same project team. Cooperative considerations do not affect joining one team over another. Our sample has no empirical evidence for an explicit exclusion or separation pattern regarding cooperation.

1.7 LESSONS LEARNED FOR COOPERATION RESEARCHERS: WHAT (NOT) TO DO?

The most important contribution of this dissertation is providing more insight into the conditions and processes through which exclusionary mechanisms can be a solution (or not) to cooperation problems. I worked with an integrative framework in which insights from psychology (individual differences in cooperativeness) and sociology (exclusionary mechanisms) are combined to understand conditions under which cooperation breaks down or thrives when exclusion operates as a social mechanism.

I point to the following six takeaways from which I believe future cooperation research could and – perhaps would – benefit from (*i*) uncovering prosociality tendencies, (*ii*) combining insights from sociology and personality research, (*iii*) studying the effectiveness of mechanisms to promote cooperation among defectors in the long haul, (*iv*) including a dynamic approach in which traits function as a reactive trait, (*v*) studying the sustainability of cooperative behavior, and (*vi*) using novel statistical methods that account for the interaction between the individual and the context. Each paragraph starts with a statement showing how I view future cooperation research based on this dissertation's results.

First, a key takeaway is that observed behavior is not always a reliable indicator of innate cooperativeness because an individual's social context also influences behavior. Thus, knowing what type of person one is provides information on whether a (mis)match between behavior and personality and between context and behavior occurs. Adjusting behavior to contextual influences does not convey whether one is truly a cooperator or defector or simply adhering to the norm set by the social environment. For example, individuals can be cajoled into cooperation or reciprocate with defection if others defect. If one is unaware that a prosocial type defects due to social influence, one may wrongly attribute the label "proself" to the defecting prosocial. Thus, it is essential for future research to gather information from the individual to, firstly, understand what is driving behavior and, secondly, what potentially explains shown behavior.

Second, combining insights from psychology and sociology provides a framework for deepening our understanding of cooperation. An exclusively sociological approach – studying how individuals affect their social surroundings and how the context influences behavior – provides a comprehensive taxonomy of why individuals behave in a certain way and why a social context is what it is. Yet, I also show that studies benefit from an interdisciplinary approach in which insights from psychology and personality research are included. For instance, incorporating under which conditions individual features, such as SVO, affect cooperative behavior provides more insights into why some cooperate and why some defect. Some may thus react differently to social pressure than others. A sole sociological view would overlook this issue. As an illustration, promoting cooperation via punishment may spark retaliatory feelings and thus lead to less cooperation (Nikiforakis et al., 2012), even though implementing punishments ideally enforces cooperation (Fehr & Gächter, 2002; Flache et al., 2017a).

Third, connections, interactions, and social motivation are needed for defectors – who may reluctantly maintain to defect – to learn that cooperation is more beneficial in the long haul. Most social mechanisms are designed to incentivize and motivate innately proself types to cooperate (Simpson & Willer, 2015). To make cooperation sustainable, one hopes that people do not need constant external motivation to cooperate but that they are, at some point, intrinsically motivated to do so. Yet, the question is whether defectors alter their innate prosociality in response to social cues. I show that an interdisciplinary approach to this is considering contextual influences and the role of personality in behavior.

Fourth, a dynamic, interactional account that includes personality insights is valuable for cooperation research, and it also mirrors the complexity real-life encompasses (Van Lange et al., 2013). I treat context and individual behavior here as a dynamic process with malleable configurations. The same reactivity – behavior to context and context to behavior – may be well applicable to prosociality. Persons may have a reactive trait (Ackermann et al., 2016). For example, a prosocial but vengeful person may react to social pressure from friends by withdrawing cooperative efforts. Thus, the social context influences cooperative behavior and provokes reactions that find their roots in personality. Incorporating a dynamic, interactional approach and accounting for inter and intra-individual features provides a comprehensive taxonomy of what drives cooperation.

Fifth, solutions promoting cooperation need to be sustainable. For example, think of the example of organizing an event together. If the solution is to offer monetary compensation to defectors, then cooperation comes at the expense of the welfare of the cooperator. In the long haul, cooperation is unstable as the solution of monetary compensation for cooperative efforts is imperfect and not endlessly applicable. The social and group context profoundly affects the sustainability of cooperation. I go further than standard cooperation research showing how cooperation decays over time in fixed groups (Chaudhuri, 2011; Ledyard, 1995). I show theoretically that the mere possibility of an "out" – i.e., leaving a group or severing a friendship relationship – does not have to impede sustainable cooperation. Implementing the option of exclusion can thus be a stable solution to promote cooperation. Yet, one first has to investigate this in an empirical context. I must note that defectors can draw the short straw on multiple occasions. They may fall through the network *and* the group if they do not change their ways. For them, exclusion is a double-edged sword.

Sixth, future research can benefit by studying a single topic via multiple lenses. I use a multitude of methods and data. Namely, I link longitudinal data to the appropriate data analysis methods. The data encompasses students—a suitably diverse population where students need and have opportunities to cooperate in a concrete empirical context. These data are input for the empirically calibrated simulations (Stadtfeld, 2018; Steglich, 2018; Steglich & Snijders, 2022). The main benefit of this approach is that it allows studying the impact of social influence and/or homophily – and varying their presence – on cooperation levels and network segregation based on an empirically realistic representation of the network context under consideration.

1.8 NEXT STEPS AND A BRIEF OUTLOOK

1.8.1 Practical consequences of this dissertation

This dissertation points to practical problems related to group and network formation and contributes to science and society. The "science" part is already discussed. I offer three applications for society but note that the content in this dissertation applies to more situations in which individuals are linked in more than one way and need to work together to realize benefits unable to realize alone.

First, the results of this dissertation provide insights into policies applicable in various contexts where students need to work together. For instance, small group teaching (i.e., learning communities) has become increasingly prevalent in higher education (Brouwer et al., 2018). Yet, it is still uncertain how such communities are best organized. Specifically, Brouwer et al. (2018) indicated that higher-achieving students segregate and benefit more from small-group teaching than lower-achieving students. The structural relations in learning communities are thus critical for students to succeed in higher education. In this context, this dissertation elucidates that mixing students based on SVO in groups can improve the overall cooperation of students, but only if, at the start, enough cooperate; otherwise, the group is less able to produce collectively valuable outcomes. Also, switching project partners might promote cooperation in learning communities, but the question of who a cooperator is, is difficult to answer if the information is lacking. The default option – relying on who you know – may lead to segregation-like effects in multiple contexts.

Consider, for example, friends selecting cooperation partners and vice versa – i.e., a spillover from one context to another. On the one hand, having friendships in your studies or work environment may be beneficial. It may be more convenient to work with friends, to have low costs in working together because you already know one another, and perhaps to experience more fun from working with friends. On the other hand, the key features of a friendship relationship (informal roles) and a study or work relation (formal roles) may be troublesome. For example, friends may shield each other from social repercussions; if not contributing to a team project is punished – e.g., by being assigned extra work – a friend may offer support or help, possibly diminishing the punishment's impact and goal. As such, friends may stand by each other, and the formal and informal roles may become blurrier. A clear example is when a manager is reluctant to make a hard decision about an employee because they are friends. A manager perhaps does not want to offend a friend or chooses to protect their friendship instead of making an organizational decision. Hence, pursuing relational and study or work goals may thus be at odds with one another. This multiplexity needs to be managed in some way. One way is that a manager ensures that employees do not always self-organize their teams (as friends tend to stick together in teams; see Chapter 6). Another way is to foster interactions via lunches or meetings with others from different departments or study groups. This may lead to cross-network interaction in the workplace or studies.

Second, this dissertation provides organizations that incorporate teams and groups with information on how interventions based on group formation (and thus exclusion) might work and how interventions related to social influence might be sufficient. For example, if research indicates that individuals who defect can "learn" to cooperate through being excluded, it would be appropriate to implement exclusion mechanisms that allow for the reintegration of "reformed defectors." However, the simulation results in Chapter 3 suggest that defection is not easily deterred via exclusion, and harsher exclusion mechanisms may thus backfire.

Building on this, another policy perhaps is to work with nudges instead (Thaler & Sunstein, 2008; Tversky & Kahneman, 1974); first to motivate individuals to cooperate by averting exclusion; and second to reintegrate defectors swiftly by providing them easy ways back in. Nudges are small, subtle hints that can influence behavior. Nudges can avert unwanted behavior by making the desired behavior more attractive or easier to perform and making the undesired behavior less attractive or more difficult to perform. For example, a nudge to encourage healthy eating in a cafeteria could involve placing fresh fruits and vegetables at eye level and in prominent locations while placing less healthy options in less visible or less accessible locations. Combining the knowledge of nudging and exclusion, a policy could be "graduated exclusion." Such a policy would start with a 'nudge' indicating that one gets an extra reward for cooperating instead of defection. Then, if defection persists, a mild form of exclusion could be to exclude someone from a meeting and not indefinitely from the group. If this 'nudge' does not work, one could increase in exclusion severity (excluded for a week) if the individual continues to defect. For example, students who do not contribute to a group project receive a lower grade than their cooperating peers. If they continue to defect, they could be excluded from participating in future projects with the group or fail the course. Bridging the gap between theoretical mechanisms and policies is difficult (Truijens, 2022), but it can provide a framework for understanding complex issues and the potential consequences of different policy options.

Third, stakeholders, such as study directors and managers, can implement inclusive policies to deter segregation-like effects where cooperating actors benefit most. An integrative approach – understanding the individual and the relations among individuals in multiple contexts – is key to avoiding cooperators primarily interacting and working together with similar others, leaving defectors without the option to learn to cooperate. For instance, it is found that individuals tend to select cooperation partners based on factors such as friendship and social background (e.g., socio-economic status, gender, grades) similarity rather than cooperativeness. A policy could be to encourage the formation of cross-cultural or cross-socioeconomic groups to foster cooperation between individuals from different backgrounds so that they can learn to cooperate from possibly dissimilar others. This could be achieved through mentoring, team-building activities, or community service projects, encouraging individuals from different backgrounds to interact and work together. Said policy aims to create an environment fostering a cascade of cooperation among individuals who – without the policy – would not interact.

1.8.2 And now, the end is not near: Pondering where we should go next

What is next? Future research may want to include a similar method-of-decreasing-abstraction approach in which many interrelated social mechanisms are monitored. For example, incentives are often used to motivate individuals to cooperate, and building trust is often used to foster stable relationships. Concretely, communicating before exclusion may ensure that, for example, being excluded as a defector does not lead to a loss of trust or even be averted if a defector is allowed to ensure that they will cooperate in the future. It may even be that a "healed" defector receives rewards for not maintaining the defect.

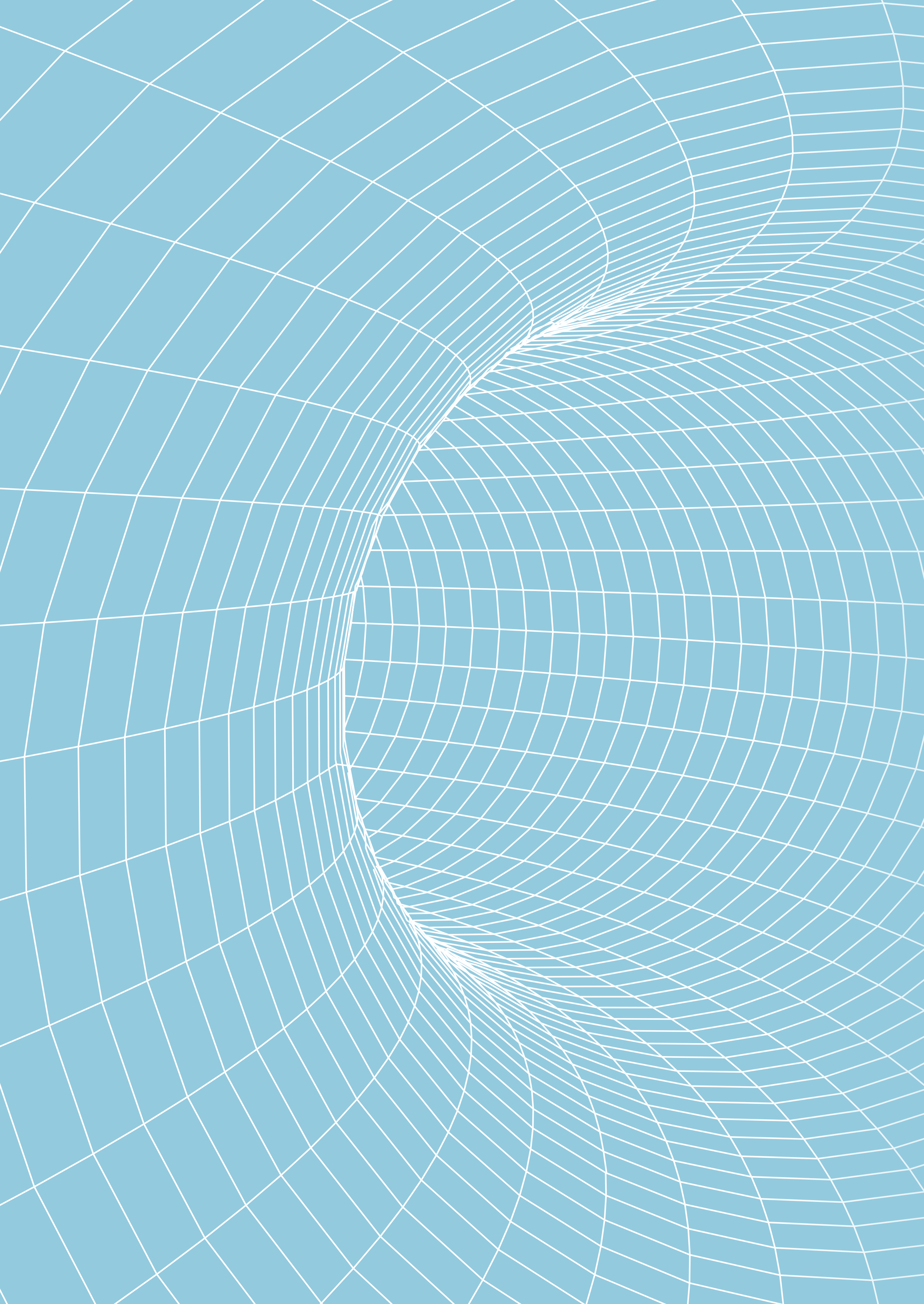
Furthermore, ABCM and computational social science (CSS) are two influential fields closely intertwined in studying cooperation (Flache, Mäs, & Keijzer, 2022). ABCM is a computer simulation technique that allows researchers to create virtual worlds populated by agents interacting to simulate real-world scenarios. This allows researchers to gain insights into the dynamics of cooperation in various social contexts (see Chapter 3 for an application). On the other hand, CSS is the application of computer science and data science to study social phenomena (Edelmann et al., 2020). This field is used to analyze large datasets to understand better the dynamics of cooperation between agents in complex systems. By combining the two fields, researchers can gain a more comprehensive understanding of the dynamics of cooperation and use this knowledge to build better models and simulations of real-world scenarios (see Chapter 5 for an application of empirically-calibrated simulations; for an example, see Manzo et al., 2018).

Although the future is difficult to predict (de Matos Fernandes & Keijzer, 2020), cooperation research can help study potential scenarios by using data on past interactions and trends to identify potential future patterns and trends (Hofman et al., 2017). Yet, Watts (2017) stresses that social science should be more solution-oriented: Rather than focusing solely on understanding the causes of complex social issues, social science should also focus on developing evidence-based solutions. This involves considering the context in which the issue occurs. By doing so, researchers can develop interventions tailored to specific contexts and effectively address underlying issues. A recent example is how behavioral and social science research can aid in dealing with the social consequences of the COVID-19 pandemic (Van Bavel et al., 2020). Research can help inform public health policies by better understanding how people respond to social distancing measures, the impact of economic downturns, and the psychological effects of prolonged isolation. By using evidence-based research, policymakers can develop effective strategies for mitigating the social consequences of the pandemic (Ruggeri et al., 2022).

1.8.3 Outro

Cooperation is part of everyday interactions. Prior research identified multiple mechanisms for cooperation to arise, spread, and thrive. I focused on the power of exclusionary mechanisms to solve the problem of cooperation but also point to the paradoxical nature of the mechanism. I explicate in this dissertation that exclusionary mechanisms (Tier 2) matter in the group and network context. Above all, for collectives to achieve the optimal outcome, we need practical solutions suitable for the context under consideration. I show that an integrative approach – combining insight from psychology and sociology – is a step in the right direction.

Finally, it must be noted that exclusionary mechanisms have a downside: If a defector is on the “wrong side of things,” then learning to cooperate is difficult. Being embedded into cooperative groups and social structures is thus key for cooperation to arise. A defector is, in such instances, more motivated to follow suit. However, insulated from social pressures to cooperate, defectors continue defecting when others do so as well in their social environment. The situation becomes even more stringent if defectors are isolated, receiving no backlash from defecting and no social pressure to change their ways. The cooperative collective needs to keep trawling in its social vicinity to allow everyone to receive benefits from cooperation; otherwise, segregation persists in groups or social networks. Simply providing an “out” via exclusion thus needs to be accompanied by an “in.” Not only do defectors need to work to be included in cooperative endeavors, but also cooperators need to provide conditions in which defectors can “find cooperation.”



Chapter 2

The stability of social value orientations[1]

Life is like a box of chocolates.
You never know what you're gonna get.

—A quote by Tom Hanks as Forrest Gump
in the movie *Forrest Gump* (1994)

[1] This chapter is the result of joint work with **Dieko Bakker** and **Jacob Dijkstra**, which appeared in *Judgment and Decision Making* in 2022 under the title "*Assessing the test-retest reliability of the social value orientation slider measure.*" Small modifications are made in comparison to the journal version to stay in line with the central tenet in this dissertation.

ABSTRACT

After establishing the content of Tiers 1 and 2 in Chapter 1, the present dissertation proceeds to investigate the stability of social value orientations (SVOs) in Chapter 2. Specifically, Chapter 2 aims to determine whether the SVO slider measure, a tool designed to assess cooperative intentions, is consistent over time. This inquiry is motivated by the potential use of SVOs as a criterion for exclusionary decision-making. To ensure that purported cooperators are not inadvertently excluded alongside defectors, it is imperative to establish the stability of SVOs as a trait. Thus, Chapter 2 serves as a precursory step towards making informed exclusion choices by examining the temporal stability of SVOs among potential cooperation partners.

2.1 INTRODUCTION

"Personality traits are probabilistic descriptions of relatively stable patterns of emotion, motivation, cognition, and behavior" (DeYoung, 2015, p. 64). Social value orientation (SVO) is purportedly such a personality trait and is frequently invoked to explain individual variation in cooperative behavior (Van Lange et al., 2014). However, SVO must be stable over time to qualify as a personality trait. In particular, empirical measures of SVO should exhibit high degrees of test-retest reliability. The empirical evidence for this is currently scant. Therefore, we test whether the SVO slider measure (SVOSM), a popular SVO measure frequently used after its introduction in 2011 (Bakker & Dijkstra, 2021; Murphy et al., 2011), captures a stable personality trait. Stability is an important psychometric property required of any measure claiming to translate to an internally valid, consistent, and reliable assessment of the studied trait. Measuring SVO reliably has long been a scientific goal (Au & Kwong, 2004; Balliet et al., 2009; Murphy & Ackermann, 2014). We contribute to reaching this goal by analyzing SVOSM panel data ($N = 495$) from six-monthly repeated measures and assessing test-retest reliability.

Even though several measures exist (Thielmann et al., 2020), we focus on the SVOSM because this measure is specifically designed to assess a trait related to cooperation, namely: SVO is defined by the weight individuals assign to their own and others' outcomes in situations of interdependence (Messick & McClintock, 1968). Primary reasons for researchers to rely on the SVOSM instead of other measures include the fact that SVOSM is not very burdensome for participants (consisting of just six items), has clear consistency checks, purportedly has high test-retest reliability, and yields a continuous score (Murphy & Ackermann, 2014). Categorical classifications may fail to capture minor individual differences in SVO (Bakker & Dijkstra, 2021), and the SVOSM allows researchers to utilize continuous scores. Even though most researchers utilize the slider measure to capture SVO as a categorical construct, the designers of the SVOSM recognized that SVO is "best represented as a continuous scale" (Murphy et al., 2011, p. 772). Fleeson (2001) conducted a similar approach, studying the Big Five as a distributional continuous measure rather than a discrete categorical one. Yet, we go further than Fleeson and inspect whether distributions of SVO continuous scores are alike over time. We move in Chapter 2 beyond treating SVO as a category and rely on SVO as a continuous construct.

SVO is generally considered a stable construct (Bogaert et al., 2008; Van Lange et al., 2014). If this is true, repeated measurements of the SVOSM should show stable and strong associations between continuous SVO scores over more extended periods. Yet, the little research into the test-retest reliability of the SVOSM there is used measurements just one week apart. We remedy this situation with our panel design of 6 measurements one month apart.

Assessing test-retest reliability is of the highest relevance both empirically and methodologically, as indicated in Chapter 1. Individuals with high SVO scores are shown to cooperate more than individuals low on SVO in observational (e.g., volunteering; Manesi et al., 2019), experimental (Balliet et al., 2009), and computational (de Matos Fernandes et al., 2022b) studies.

Establishing the test-retest reliability of SVO strengthens its case as a reliable predictor of cooperation (and other behaviors). Apart from employing a design with longer time intervals between measurements, we also advance the field by relying on a non-student sample. Van Lange et al. (2014, p. 148) posit that we know surprisingly little about SVO in non-student samples. We rise to the occasion and use a representative sample of the Dutch population.

In the remainder of Chapter 2, we first discuss previous research. We then describe the data collection process and our sample, followed by a presentation of our findings. We end Chapter 2 with prospects for future research and some concluding remarks.

2.2 WHAT WE KNOW THUS FAR

Previous assessments of SVOSM's test-retest reliability are encouraging. With approximately one hour between two waves (*N* = 124), Ackermann & Murphy (2019) report a correlation of 0.72 between SVOSM scores. Most other studies use a two measurements design one week apart. One study with *N* = 872 reports a correlation of 0.79 between continuous SVOSM scores (Höglinger & Wehrli, 2017). The developers of the SVOSM report a correlation of 0.92 with a sample of 46 students (Murphy et al., 2011). Another study reports a correlation of 0.75 in a non-monetary condition (only show-up fee; *N* = 155 students) and a correlation of 0.35 in an incentivized monetary condition with *N* = 62 (Reyna et al., 2018). Hence, the SVOSM seems relatively stable, but some measurement-to-measurement variation is present. To our knowledge, only one study investigates the test-retest reliability of the SVOSM in a much longer time frame. Bakker & Dijkstra (2021) report a correlation of 0.60 between continuous SVOSM scores, relying on *N* = 86 students and two measurements three months apart. On the one hand, temporal instability may result from random measurement errors. On the other, it may result from SVO being systematically affected by, for example, personal experiences. Overall, prior research on the temporal stability of the SVOSM suffers from two defects: (*i*) studies either use very short time frames or have low sample sizes when using longer time frames, and (*ii*) studies rely exclusively on student samples. We remedy both shortcomings.

2.3 METHOD

2.3.1 Social value orientation slider measure

The SVOSM has six items. Each item contains several alternative resource allocations, with the ranges of own and others' payoff changing across items. Per item, respondents are asked to decide how they wish to allocate units of some hypothetically valuable good between themselves and a random other person. Respondents were informed about the hypothetical nature of the questions and did not earn extra money in addition to their participation fee. We chose this non-incentivized design because most existing SVO studies do not use monetary incentives. In light of the findings of Reyna et al. (2018) mentioned above, investigating test-retest reliability across longer time frames in non-student samples in incentivized designs also seems valuable. We leave this question for future research. Finally, to calculate SVOs, we need to compute each respondent's SVO degree. We first calculated the mean payoff allocated to themselves and the other for all items and then measure a single SVO degree score based on the mean-self to the mean-other ratio (see Murphy et al., 2011, or Murphy & Ackermann, 2015, for more information on the measure and how to compute continuous SVO scores).

2.3.2 Data collection and our sample

We used a 6-month repeated measures design where respondents filled in the SVOSM each month in a non-experimental context. The first wave of data collection occurred in January 2021, with subsequent waves administered in February (wave 2), March (wave 3), April (wave 4), May (wave 5), and June (wave 6). Questionnaires started with an introduction, followed by an example SVO question to get acquainted with the type and format of SVO questions. Then respondents answered six allocation questions. Data were collected by the Longitudinal Internet studies for the Social Sciences (LISS) panel administered by CentERdata (Tilburg University, the Netherlands). The LISS panel is a representative sample of Dutch individuals. The panel is based on a true probability sample of households drawn from the population register and consists of 4500 households, comprising 7000 individuals. Our sample ($N = 495$) is a random subset of the panel.

2.3.3 Consistency check

A vital property of the SVOSM is its consistency check, allowing researchers to exclude respondents who are inconsistent in their allocation preferences. This may indicate random answers or a lack of understanding. Murphy et al. (2011) suggested excluding respondents whose answers result in intransitive preferences over SVOs. As an alternative, Bakker & Dijkstra (2021) suggest excluding respondents whose answers were so inconsistent that their resulting vector is too short (i.e., whether distance, D, is smaller than some cutoff value). The more consistently a respondent chooses allocations corresponding to a particular SVO, the longer their D will be. Vectors shorter than 35 are considered inconsistent and are excluded from the sample.[2] For more information on computing vector lengths and the mathematical function, we refer the reader to the supplementary file attached to Bakker & Dijkstra (2021). We find that 9 percent of answer profiles are intransitive across all waves while 7 percent fail the vector length criterion. A total of 302 (approximately 14%) out of 2176 responses were excluded because they failed to meet the transitivity criterion, the vector length criterion, or both.

2 The choice for 35 as the criterion is based on 39.99 (mean vector length) − 2 ⋆ 2.47 (standard deviation). Additional analyses using 37.5 (applied by Bakker & Dijkstra, 2021) or 40 show that relying on stricter vector length criteria leads to more stringent filtering of then considered inconsistent answer profiles: excluding 12.5 and 36.4% of responses respectively. The intra-class correlation coefficient (see section 4.2) goes up from 0.78 ($D = 35$, $N = 230$) to 0.81 ($D = 37.5$, $N = 214$) and 0.90 ($D = 40$, $N = 103$). A more conservative vector criterion leads, as expected, to fewer inconsistencies in answer profiles and higher test-retest reliability.

2.4 FINDINGS[3]

2.4.1 Distribution of SVO continuous scores in our sample

In Table 2.1, we provide descriptive statistics for all six items in our questionnaire. Generally, we find that the mean scores do not vary that much. The standard deviations across payoffs allocated to themselves and the others, however, do show some variance. Especially items 1-other, 6-other, 4, and 5 show differences in allocation choices. Murphy et al. (2011) denote that, next to the self-other dimension, SVO items capture differences in preferences for maximizing own and others' outcomes and (in)equality. If respondents favor maximizing their payoff, they tend to select a self-payoff of 85 (other-payoff = 15) in item 4.

Table 2.1: Inspecting the six SVO items separately, average payoffs to self and the other, and average SVO scores. Payoff ranges of items 1 to 6 are reported in a note below.

	Wave 1	**Wave 2**	**Wave 3**	**Wave 4**	**Wave 5**	**Wave 6**
Item	***M (SD)***	***M (SD)***	***M (SD)***	***M (SD)***	***M (SD)***	***M (SD)***
1-self	85 (0)	85 (0)	85 (0)	85 (0)	85 (0)	85 (0)
1-other	79.1 (14.3)	81.3 (11.5)	82.2 (10.0)	82.1 (10.6)	82.0 (10.4)	82.1 (10.8)
2-self	99.3 (2.2)	99.5 (1.6)	99.5 (1.5)	99.5 (1.7)	99.5 (1.6)	99.7 (0.9)
2-other	48.3 (5.1)	48.8 (3.7)	48.9 (3.6)	48.9 (4.0)	48.9 (3.8)	49.2 (2.1)
3-self	82.8 (6.5)	83.4 (5.4)	83.0 (6.5)	82.8 (6.8)	83.1 (6.1)	83.4 (5.3)
3-other	85.9 (2.8)	85.7 (2.3)	85.9 (2.8)	85.9 (2.9)	85.8 (2.6)	85.7 (2.3)
4-self	68.7 (10.1)	69.2 (10.0)	68.6 (9.8)	68.6 (9.6)	68.6 (9.8)	69.5 (10.2)
4-other	54.6 (24.6)	53.4 (24.2)	54.8 (23.9)	54.9 (23.3)	54.9 (23.7)	52.7 (24.7)
5-self	83.8 (11.5)	83.3 (11.4)	83.0 (11.6)	81.7 (11.4)	82.3 (11.6)	83.4 (11.5)
5-other	66.2 (11.5)	66.7 (11.4)	67.0 (11.6)	68.3 (11.4)	67.7 (11.6)	66.6 (11.5)
6-self	89.7 (6.2)	89.1 (5.9)	89.1 (6.1)	88.9 (5.8)	89.0 (6.0)	89.3 (6.2)
6-other	74.0 (14.5)	75.4 (13.8)	75.4 (14.1)	57.9 (13.6)	75.6 (14.0)	75.1 (14.4)
Self	84.9 (4.4)	84.9 (4.2)	84.7 (4.4)	84.4 (4.3)	84.6 (4.4)	85.0 (4.4)
Other	68.0 (8.7)	68.5 (8.2)	69.0 (8.0)	69.3 (7.9)	69.1 (8.0)	68.6 (8.3)
SVO	27.3 (13.7)	28.0 (12.9)	28.9 (12.8)	29.4 (12.5)	29.1 (12.7)	28.0 (13.0)

Note. *M* = mean; *SD* = standard deviation; Ranges of the SVOSM items in the questionnaire comprise from left to right: 1-self = 85, 1-other = 85 to 15, 2-self = 85 to 100, 2-other = 15 to 50, 3-self = 50 to 85, 3-other = 100 to 85, 4-self = 50 to 85, 4-other = 100 to 15, 5-self = 100 to 50, 5-other = 50 to 100, 6-self = 100 to 85, and 6-other = 50 to 85.

3 The data used in this Chapter and R-script to analyze the data and plot the figures are freely available at the Open Science Framework, via https://doi.org/10.17605/OSF.IO/TW8DQ.

Similarly, if respondents prefer equality in outcomes, then they would choose an allocation in, for example, item 5 that leads to an equal distribution. Yet, respondents favoring inequality in outcomes choose either a higher payoff for themselves or the other. The standard deviations in said items show variation in payoff allocations across waves, attesting to the need to assess the test-retest reliability of SVO via distributions and not solely based on mean scores or discrete categories. Finally, the observed average payoff allocated to themselves of all six items combined ranges from 67 to 93, with a mean of 84.7 (*SD* = 4.3). Conversely, the average payoff allocated to the other ranges from 38 to 87, with a mean of 68.8 (*SD* = 8.2). The mean payoff scores to self and the other vary little across waves.

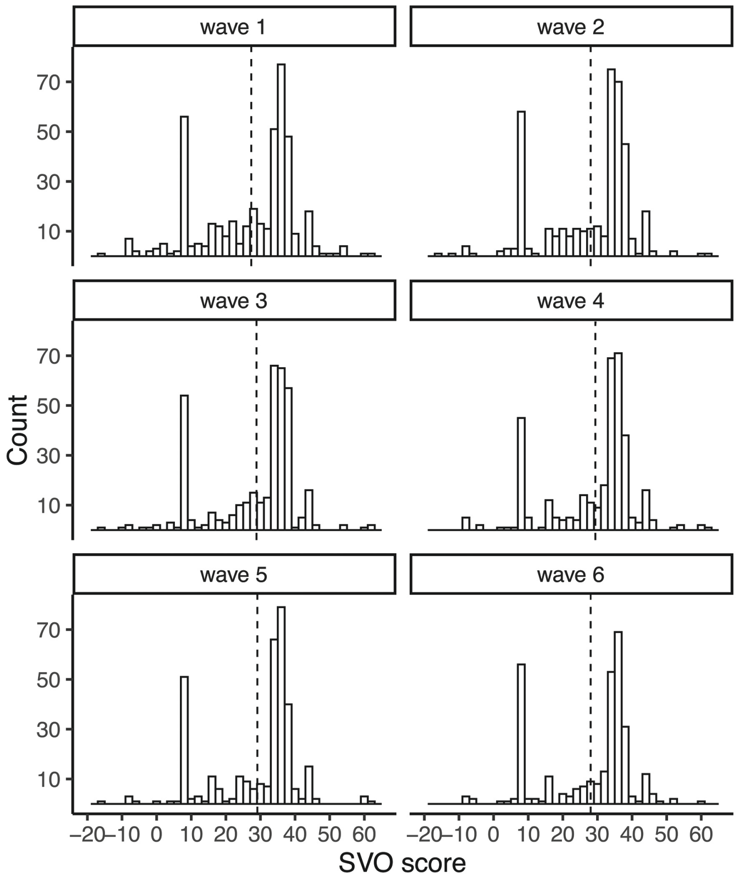


Figure 2.1: Visualizing SVO scores per wave. The mean is shown via a dashed line.

We now turn to our sample's distribution of SVO degrees (the continuous scores) (Table 2.1 and Figure 2.1). Individual SVO degree scores are based on answer profiles on all six items. They are summarized as the mean payoffs allocated to themselves and the other (as shown in Table 2.1) to provide a single index score per wave. Observed SVO degrees range from −16.1 to 61.4, with a mean across all respondents and waves of 28.4 (*SD* = 13.0). Lower scores on the scale indicate a more proself orientation, while higher scores indicate that the respondents orient more prosocially. Table 2.1 shows that mean SVO scores vary marginally across waves, but the high standard deviations point to substantial variance in SVO. Figure 2.1 provides us with a visual inspection of SVO distributions across waves. In particular, we see two major spikes, one approximately at score 8 and one near score 35. These represent respondents who consistently select either prosocial (score 35) or individualistic (score 8) allocations (Bakker & Dijkstra, 2021; Murphy et al., 2011). The descriptive analyses in Table 2.1 and Figure 2.1 point to the presence of variation, showing the need to explore intra-individual differences in SVO rather than wave-by-wave comparisons of mean scores.

Our sample suffered from attrition. Almost 33 percent of respondents dropped out from waves 1 to 6. In brief, attrition did not significantly affect the distribution of SVO in our sample. For example, comparing the wave 1 distributions of SVO scores between respondents who had and had not dropped out by waves 2 to 6, using a Kolmogorov–Smirnov test (which quantifies whether two distributions differ significantly from one another), shows no significant differences. For more information on the impact of attrition on the distribution of SVO, we refer to section 2.4.5.

2.4.2 Descriptive analysis of SVO as a categorical construct

Although a major perk of the SVOSM is its potential to rely on continuous scores, it remains a largely standard practice in SVO research to compute either four or two SVO categories based on continuous SVO scores (Balliet et al., 2009; Bakker & Dijkstra, 2021; Murphy et al., 2011). We provide an overview of the distribution of SVO categories and investigate the test-retest reliability of treating SVO as a categorical construct.

We classify the observed SVO scores into categories: prosocials assign more weight to others' outcomes than individualistic types. In contrast, competitive (altruistic) types want to maximize the positive difference in outcomes between themselves (others) and others (themselves). Altruists have a score greater than 57.15. Prosocial scores lie between 22.45 and 57.15. Individualists have a score between −12.04 and 22.45. Respondents with a score less than −12.04 are classified as being competitively oriented. Most studies lump altruistic and prosocial types into prosocial categories and competitive and individualistic types into proself categories since altruistic and competitive types are rare.

Table 2.2 shows the count and percentage per SVO category and per wave in our sample. Most respondents are prosocially oriented, while many have an individualistic orientation. Our sample contains hardly any competitive or altruistic respondents. Note that the *N* per column in Table 2.2 varies due to the post-hoc removal of intransitive and small vector length responses separately per wave (row excluded). Ignoring the missing values due to sample attrition, we find that the percentage of prosocials (altruistic and prosocial types) is relatively constant, floating within the bandwidth of 67 to 76 percent. At lower percentages, the same holds for proself types with a consistent presence of around 24 to 33 percent (competitive and individualistic types). Consistent with these findings, most work usually reports that roughly two-thirds of their sample classifies as prosocial while approximately one-third are proself (Bakker & Dijkstra, 2021; Höglinger & Wehrli, 2017). A visualization of the proself-prosocial distribution per wave is provided in Figure 2.4.

Table 2.2: Count and percentage of respondents per SVO category per wave.

SVO Type	Wave 1	Wave 2	Wave 3	Wave 4	Wave 5	Wave 6
competitive	1 (0.2%)	1 (0.2%)	1 (0.2%)	0 (0%)	1 (0.2%)	0 (0%)
individualistic	135 (32.5%)	111 (25.6%)	91 (20.5%)	85 (18.9%)	82 (17.6%)	88 (19.2%)
prosocial	277 (66.8%)	275 (63.5%)	275 (62.1%)	267 (59.1%)	253 (54.2%)	219 (47.8%)
altruistic	2 (0.5%)	2 (0.5%)	3 (0.7%)	3 (0.7%)	4 (0.9%)	1 (0.2%)
missing (NA)	0 (0%)	44 (10.2%)	73 (16.5%)	97 (21.5%)	127 (27.2%)	150 (32.8%)
excluded	80	62	52	43	28	37

Note. Excluded refers to intransitive and too short vector length cases; N per wave (without missing and excluded): wave 1 = 415, wave 2 = 389, wave 3 = 370, wave 4 = 355, wave 5 = 340, and wave 6 = 308.

We assess the extent to which such SVO categorization leads to the loss of explained variance in SVO continuous scores. We estimate a multilevel linear regression model for the nested data structure. We take SVO continuous scores as the dependent variable and the four or two SVO categories as the independent variable. Findings indicate that the four (R^2 = 0.81) and two (R^2 = 0.77) category implementations have a high and roughly similar degree of explanatory power.[4] This statistical finding supports the prosocial-proself dichotomy usually employed by researchers using the SVOSM. Still, some variance in SVO remains unexplained due to categorization.

2.4.3 Test-retest reliability of continuous SVO scores

We use Pearson correlations, k-sample tests, and the intra-class correlation coefficient to indicate the test-retest reliability of SVO distributions. First, SVO scores correlate positively and significantly across waves (Table 2.3), meaning respondents' SVO scores tend to be similar across all wave comparisons. Next, the similarity in SVO score distributions is confirmed by the Anderson-Darling (AD) k-sample test: the *p*-value of 0.57 indicates that we cannot reject the equality of SVO score distributions across waves. This result aligns with the distributions in Figure 2.1, the minor differences in mean continuous SVO scores per wave, and the strong positive correlations reported in Table 2.3.[5] The distribution of SVO scores is reasonably constant on the whole.

4 Assessing explained variance in multilevel models can be done via multiple R^2 measures (LaHuis et al., 2014). We rely on the Snijders & Bosker (1994) R^2 measure because it captures variance in two-level models, which we have.

5 AD k-sample test *p*-value of the slider measure items comprise: 1-self = not applicable, 1-other = 0.25, 2-self = 0.08, 2-other = 0.08, 3-self = 0.38, 3-other = 0.38, 4-self = 0.95, 4-other = 0.95, 5-self = 0.07, 5-other = 0.07, 6-self = 0.94, and 6-other = 0.94. The *p*-values above 0.05 indicate that we cannot reject equality of distributions. Items are thus similarly distributed over time.

Furthermore, we utilize the intra-class correlation coefficient (ICC) coefficient to inspect intra-individual consistency in SVO continuous scores. We find an ICC score for SVO scores of 0.78 (95% CI = [0.74, 0.82]) among respondents who participated in all six waves (N = 230). The high ICC score indicates that SVO continuous scores have high test-retest reliability.

Table 2.3: Pearson correlations of SVO scores across waves.

	Wave 1	**Wave 2**	**Wave 3**	**Wave 4**	**Wave 5**	**Wave 6**
Wave 1	—					
Wave 2	0.75	—				
Wave 3	0.72	0.78	—			
Wave 4	0.71	0.75	0.82	—		
Wave 5	0.72	0.78	0.83	0.84	—	
Wave 6	0.66	0.70	0.81	0.83	0.84	—

Note. All p-values are < 0.0001.

Moreover, we employ a multilevel linear regression to inspect whether prior SVO continuous scores are predictive of later SVO scores. Using a multilevel model, we control for the nested structure of our data in which SVO measures are nested within individuals. We include a lagged variable of SVO continuous scores, representing one's SVO score at wave minus 1 ($t - 1$). The results are reported in Table 2.4. Notably, the SVO score in the previous wave is significantly and highly indicative of later SVO scores (estimate = 0.79, SE = 0.01, $p < 0.001$). Note that the wave coefficients represent the difference between the respective waves and the intercept coefficient (wave 2). Hence, the wave 3 coefficient combines the intercept and wave 3 parameters, i.e., the estimate is 6.91 (wave 2 plus wave 3 estimates).

Table 2.4: Results of the multilevel linear regression for estimating predictors of SVO scores.

Parameter	**estimate**	**SE**	**p-value**
Intercept (wave 2)	6.11	0.58	<0.001
Wave 3	0.80	0.58	0.169
Wave 4	0.56	0.58	0.334
Wave 5	0.12	0.59	0.842
Wave 6	−0.95	0.60	0.115
SVO score $t - 1$	0.79	0.01	<0.001

Note. N = 426 with 1700 decisions; SE = standard error.

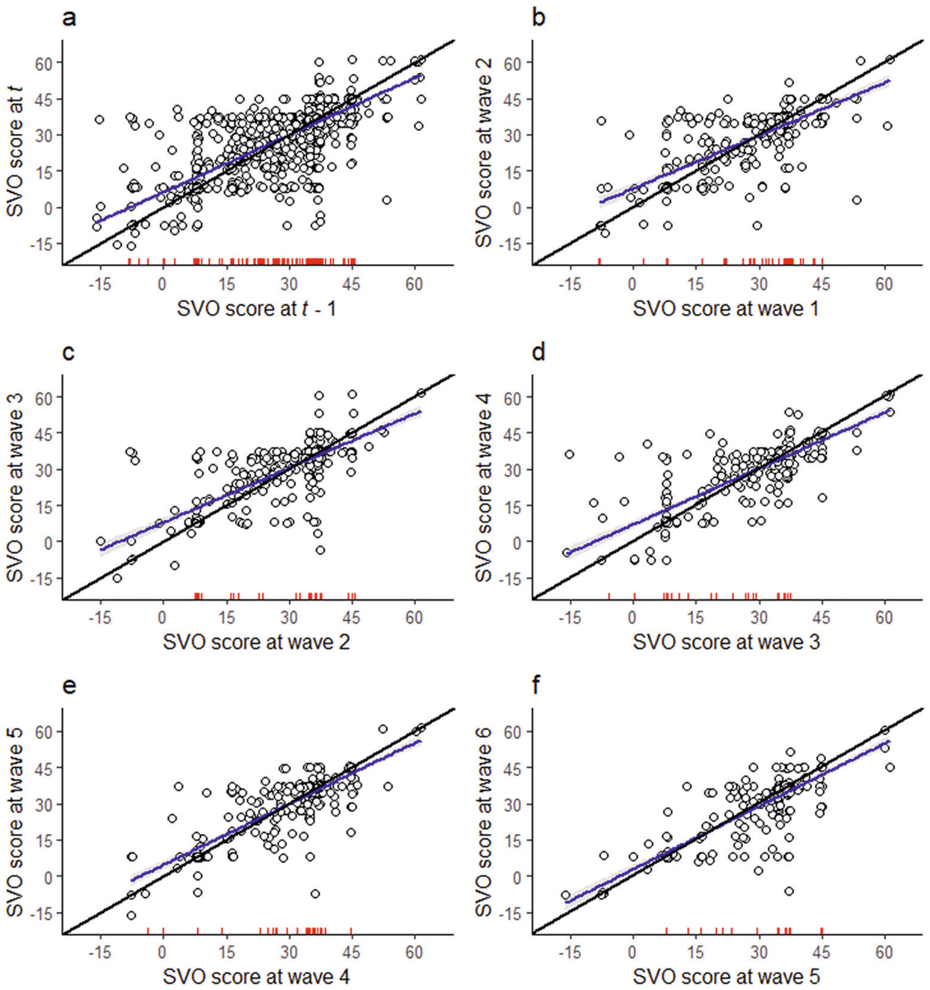


Figure 2.2: Test-retest reliability scatter plots. The diagonal black line represents perfect test-retest reliability. The blue line shows a linear regression with prior SVO (t − 1) as the independent variable and the current SVO score as the dependent variable (t). We show the marginal distribution of dropouts (no SVO score at t) in red. Panel a shows all waves combined, while panels b to f provide a wave-to-wave comparison of test-retest reliability.

We visualize test-retest reliability based on the results of Table 2.4 in Figure 2.2. The diagonal black unity line in the six plots represents perfect test-retest reliability in prior and consecutive SVO scoring. The x-axis shows a respondent's SVO score in the prior wave (referred to in Figure 2.2a as *t* − 1), showing waves 1 to 5. The y-axis shows the SVO score in wave *t*, ranging from wave *t* = 2 to 6. In Figures 2.2b-f, we show a pairwise wave-to-wave comparison. Figure 2.2 also shows the result of a linear regression—blue line—with the prior SVO score as the independent variable and the current SVO score as the dependent variable. Each data point is a paired observation, showing the SVO score at *t* − 1 and *t*. We include a marginal distribution of the SVO score of dropouts in wave *t* − 1. Further analyses in section 2.4.5 show that prior SVO scores are not key predictors of dropping out. In brief, Figure 2.2 shows a strong tendency to score similarly in SVO across waves.

Results in Tables 2.1, 2.2, 2.3, and 2.4 and Figures 2.1 and 2.2 show that the majority are primarily similar in their SVO over time, but some variation persists. We quantify to which extent differences in SVO scores occur.[6] We calculate differences in SVO by subtracting the absolute value of a respondent's SVO score at $t - 1$ from the absolute value of the SVO score at t for respondents who participated in all six waves ($N = 230$ respondents with a total of 1380 scores). The mean difference between SVO at $t - 1$ and t is 3.36 ($SD = 5.9$).[7] The mean difference of 3.36 shows that respondents, on the whole, tend to marginally differ in SVO over time. In what follows, we provide aggregated percentages of absolute differences in SVO—and not separated per wave. Almost 55 percent of scores comparing SVO between $t - 1$ and t, a total of 752 scores, is smaller than 1 (628 cross-wave comparisons differ more than 1 unit in SVO). We see an increase to 71% when we take 3 as an unit, instead of 1, as the dichotomous cutoff value in comparing differences in absolute SVO scores between $t - 1$ and t. Next, we use the standard deviation of 6.8 and two times the SD as cutoff values. Almost 79% and 93% report a difference lower in SVO across waves for 6.8 and 13.6, respectively. In sum, the majority of respondents report minor gradual differences in SVO scores over time, once again attesting to sufficiently high test-retest reliability.

2.4.4 SVO categorical test-retest reliability

Previous assessments of the test-retest reliability regarding SVOSM as categories point to a stable construct over two measurements one week apart. Höglinger & Wehrli (2017) show that 86 percent of SVO categorical classifications remained similar ($N = 872$). The developers of the SVOSM report an 89 percent consistency score (Murphy et al., 2011), while Bakker & Dijkstra (2021) report a 78 percent categorical type consistency score with two measurements three months apart. The categorical test-retest reliability decreased to 67 percent over a period of 1.5 years (Bakker & Dijkstra, 2021), but with only $N = 27$. In the current study, the mean overall instability across all waves is 0.12 (percentage stability is 88%), indicating that, on average, about 12% of respondents change categories from one wave to the next. The measurement-to-measurement variation in categorical instability is as follows: wave 1 → wave 2 = 0.17, wave 2 → wave 3 = 0.14, wave 3 → wave 4 = 0.10, wave 4 → wave 5 = 0.10, and wave 5 → wave 6 = 0.09. The trend appears to show an increasingly stable classification.

We use the following two statistics to inspect test-retest reliability in SVO categories formally: Cohen's (1960) and Fleiss' (1971) Kappa (κ). First, Cohen's κ allows us to check whether respondents stick to their categories in consecutive waves. We find that respondents are consistent

6 Multiple visualizations of individual trajectories of SVO are provided in our open access OSF folder: https://doi.org/10.17605/OSF.IO/TW8DQ.

7 Mean difference in SVO per wave is: wave 1 → wave 2 = 4.99 ($SD = 8.7$), wave 2 → wave 3 = 4.30 ($SD = 8.2$), wave 3 → wave 4 = 3.75 ($SD = 7.3$), wave 4 → wave 5 = 3.41 ($SD = 6.5$), and wave 5 → wave 6 = 3.68 ($SD = 6.3$).

in their SVO: wave 1 → wave 2 = 0.65 (N = 343), wave 2 → wave 3 = 0.70 (N = 342), wave 3 → wave 4 = 0.77 (N = 328), wave 4 → wave 5 = 0.76 (N = 317), and wave 5 → wave 6 = 0.79 (N = 299). Second, Fleiss' κ is an adaptation of Cohen's κ and allows us to simultaneously assess the consistency across all waves. We find a high Fleiss' κ of 0.70 for respondents who participated in all six waves (N = 230).

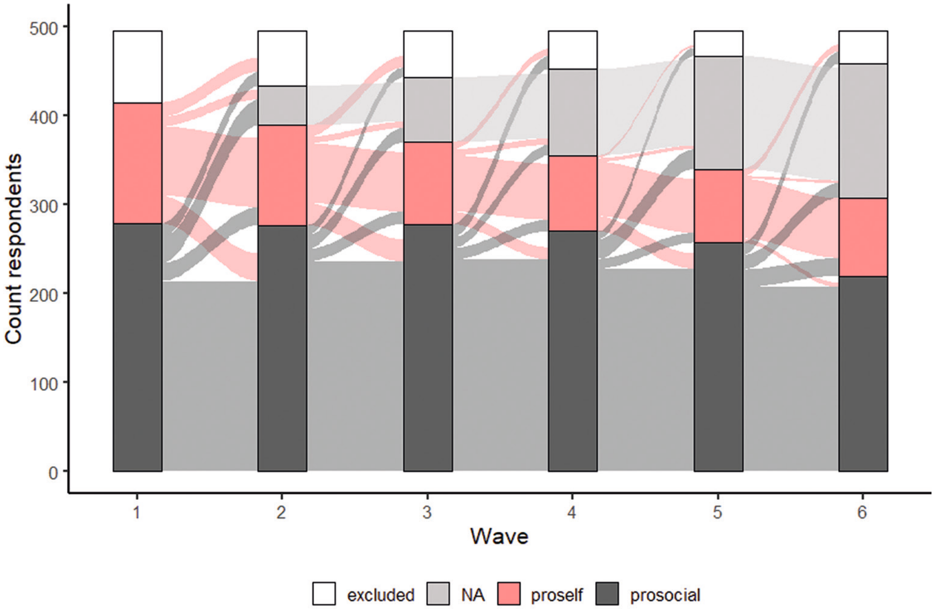


Figure 2.3: An alluvial diagram visualizing changes in SVO. Respondents with similar prior and current SVO are bundled together. Respondents with intransitive and too short vector length answer profiles are visualized in white, whereas dropouts (NA) are in light grey.

Figure 2.3 visualizes variation in SVO. The dark grey block indicates prosocials, while the light red block represents proselfs (light grey is NA, and white is intransitive and too short vector length answer profiles). Figure 2.3 shows how the prosocial and proself categories exchange members over time, while a stable flow of respondents drops out at every transition. The pool of dropouts consists mainly of prosocials, which is unsurprising given that they comprise about two-thirds of our sample.

Prior research indicates that respondents scoring near the classification boundaries are more likely to switch SVO category classification (Bakker & Dijkstra, 2021), attesting to the importance of using continuous scores. We estimate a multilevel logistic regression model to investigate whether this holds in our sample. The dependent variable is whether individuals changed regarding the SVO category from one wave to the following (1 = change and 0 = no change). To calculate proximity to the prosocial-proself boundary, we first subtracted 22.45 (the category boundary in degrees) from the continuous SVO scores and converted scores to absolute values. We then reversed the variable by subtracting the calculated distance from the maximum possible distance so that higher scores indicate proximity to the boundary. Results show that respondents scoring near the boundary are more likely to change SVO categories than respondents farther from the boundary (estimate = 0.05, $p = 0.002$). This finding is in sync with prior research (Bakker & Dijkstra, 2021). Minor gradual changes in SVO categories may thus lead to major consequences in SVO stability in the long haul. Also, the negative wave effects (e.g., wave 1 → 2 estimate = −3.12, $p < 0.001$) indicate that changing SVO categories is not very likely from the outset and becomes even less likely in later waves (e.g., wave 5 → 6 estimate = −4.07, $p < 0.001$), which confirms the low instability percentages discussed earlier. Furthermore, prosocial-oriented respondents are less likely to change their SVO category than proselfs (estimate = −0.93, $p < 0.001$).

The results are fairly the same as with the continuous SVO scores: we find that respondents tend to orient similarly over time, while some measurement-to-measurement variation persists.

2.4.5 Attrition in our sample

Data collection started with $N = 495$ in wave 1 and ended up with $N = 345$ in wave 6 (see Table 2.2). A total of 44 respondents dropped out in wave 2, 73 in wave 3, 97 in wave 4, 127 in wave 5, and 150 in wave 6. The total attrition rate is 33% when comparing the sample size from waves 1 to 6. Figures 2.4a and b visualize the distribution of prosocial and proself categories with and without the NA, unavailable cohort. The percentages of types remain rather similar over time. This is a first indication that attrition did not substantively affect our sample's distribution of prosocial and proself types.

We formally test the role of attrition on SVO using Fisher's exact test (categorical SVO) and the Kolmogorov-Smirnov test (continuous SVO). Applying Fisher's exact test to the distributions of SVO categories of wave 1 respondents who had and had not dropped out by wave 2 to 6, we find no statistical difference ($p = 1$ for all wave 1 to future wave, 2 to 6, comparisons). Thus, the impact of attrition on the SVO category distribution seems minimal. The same holds for treating SVO as a continuous construct. A Kolmogorov–Smirnov test shows that the SVO continuous score distribution of wave 1 respondents who, again, had and had not dropped out by wave 2 to 6 are equally distributed. To be clear, the Kolmogorov–Smirnov test p-value per wave is as follows: wave 1 → wave 2 = 0.24, wave 1 → wave 3 = 0.45, wave 1 → wave 4 = 0.28, wave 1 → wave 5 = 0.36, and wave 1 → wave 6 = 0.73. Attrition thus did not lead to significant differences in SVO distributions.

Next, we investigate whether changing SVO categories is a prerequisite for dropping out in later waves. We conducted supplementary logistic regression analyses with dropping out measured at waves 3, 4, 5, and 6 as dependent variables (1 = dropping out, 0 = maintaining participation). To be clear, we tested in four separate logistic models whether, for example, changing in SVO from wave 1 to 2 increases the likelihood of dropping out in wave 3. Subsequent models include changing SVO from waves 2 to 3, 3 to 4, and 4 to 5 as independent variables and dropping out in respectively waves 4, 5, and 6 as dependent variables. We included changing in SVO categories from wave 1 to 2 (estimate = 0.14, *SE* = 0.57, *p* = 0.81), 2 to 3 (estimate = −0.97, *SE* = 1.04, *p* = 0.35), 3 to 4 (estimate = −0.27, *SE* = 0.76, *p* = 0.73), and 4 to 5 (estimate = −0.58, *SE* = 1.05, *p* = 0.58) as independent variables. Our analyses reveal that instability in SVO in the past is not a significant predictor of dropping out in future waves.

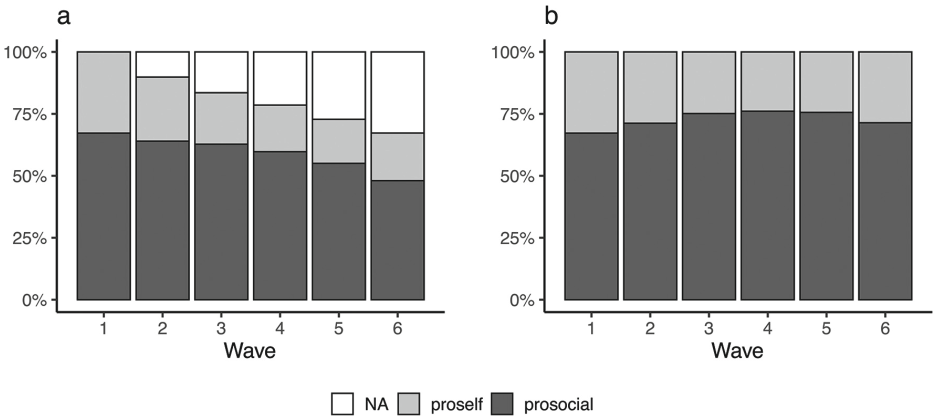


Figure 2.4: Visualizing percentages prosocial and proself types with (a) and without (b) NA's.

We explore whether a proself vs. prosocial orientation and orientating high or low on the continuous SVO scale predict dropping out in our sample. The dependent variable is again dropping out (1) or not (0). We include SVO at $t-1$ as a dichotomous or continuous independent variable. The continuous SVO scores in $t-1$ are generally not predictive of dropping out in wave 2 (estimate = 0.01, *SE* = 0.01, *p* = 0.46), wave 3 (estimate = −0.00, SE = 0.01, *p* = 0.82), wave 5 (estimate = 0.00, *SE* = 0.02, *p* = 0.96), or wave 6 (estimate = 0.01, *SE* = 0.02, *p* = 0.69). The sole exception is dropping out in wave 4 (estimate = −0.03, *SE* = 0.01, *p* = 0.04). Respondents with higher SVO scores are less likely to drop out than respondents lower on SVO. Twenty-four respondents (6%) dropped out in wave 4, and 355 (94%) maintained to participate in our study. The mean SVO score of those 24 respondents is 22.7 (*SD* = 13.3) vs. 28.6 (*SD* = 13.2) of the stayers. An additional Kolmogorov–Smirnov test shows that the SVO continuous score distribution from wave 3 ($t-1$) among dropouts and stayers in wave 4 does not significantly differ according to $p < 0.05$ standards: *p*-value = 0.06. The analysis of SVO as a category does not confirm the higher chances of proself types to drop out more readily than their prosocial counterparts: wave 4 estimate = −0.71, *SE* = 0.43, *p* = 0.10). The non-effect of prosocial and proself categorization is confirmed in other waves: wave 2 (estimate = 0.30, *SE* = 0.36, *p* = 0.39), wave 3 (estimate = −0.13, *SE* = 0.42, *p* = 0.76), wave 5 (estimate = 0.54, *SE* = 0.51, *p* = 0.28), or wave 6 (estimate = −0.13, *SE* = 0.49, *p* = 0.80).

Finally, we study whether having an extreme SVO — for example, preferring to maximize payoffs to themselves or the other — is a predictor of dropping out. Extremeness in SVO is calculated as follows: as a benchmark, we take the diagonal line in the distribution between payoffs to self and the other (as visualized in Murphy et al., 2011, p. 773, Figure 2). The diagonal line represents a

SVO score of 45.[8] Then, we calculate the absolute distance to 45 in SVO in $t-1$ and use that indicator as an explanatory variable for the logistic regression. The dependent variable is dropping out (1) or not (0). Distance to the 45 benchmark SVO score in $t-1$ is generally not predictive of dropping out in wave 2 (estimate = −0.01, $SE = 0.01$, $p = 0.34$), wave 3 (estimate = 0.00, $SE = 0.02$, $p = 0.86$), wave 5 (estimate = −0.00, $SE = 0.02$, $p = 0.83$), or wave 6 (estimate = −0.01, $SE = 0.02$, $p = 0.59$). The sole exception is again the impact of distance to the benchmark in dropping out in wave 4 (estimate = 0.03, $SE = 0.01$, $p = 0.04$). Respondents with more extreme SVO preferences — preferring to maximize differences to the other, either beneficial for themselves or the other — are generally more likely to drop out in wave 4. In the previous paragraph, we already stressed that especially respondents with low SVO scores drop out in wave 4, suggesting that these respondents generally favor higher payoffs allocated to themselves than the other. In brief, respondents with a particular SVO score do not disproportionately drop out in our study—cf., respondents with lower SVO degrees in wave 4.

2.5 CONCLUDING REMARKS

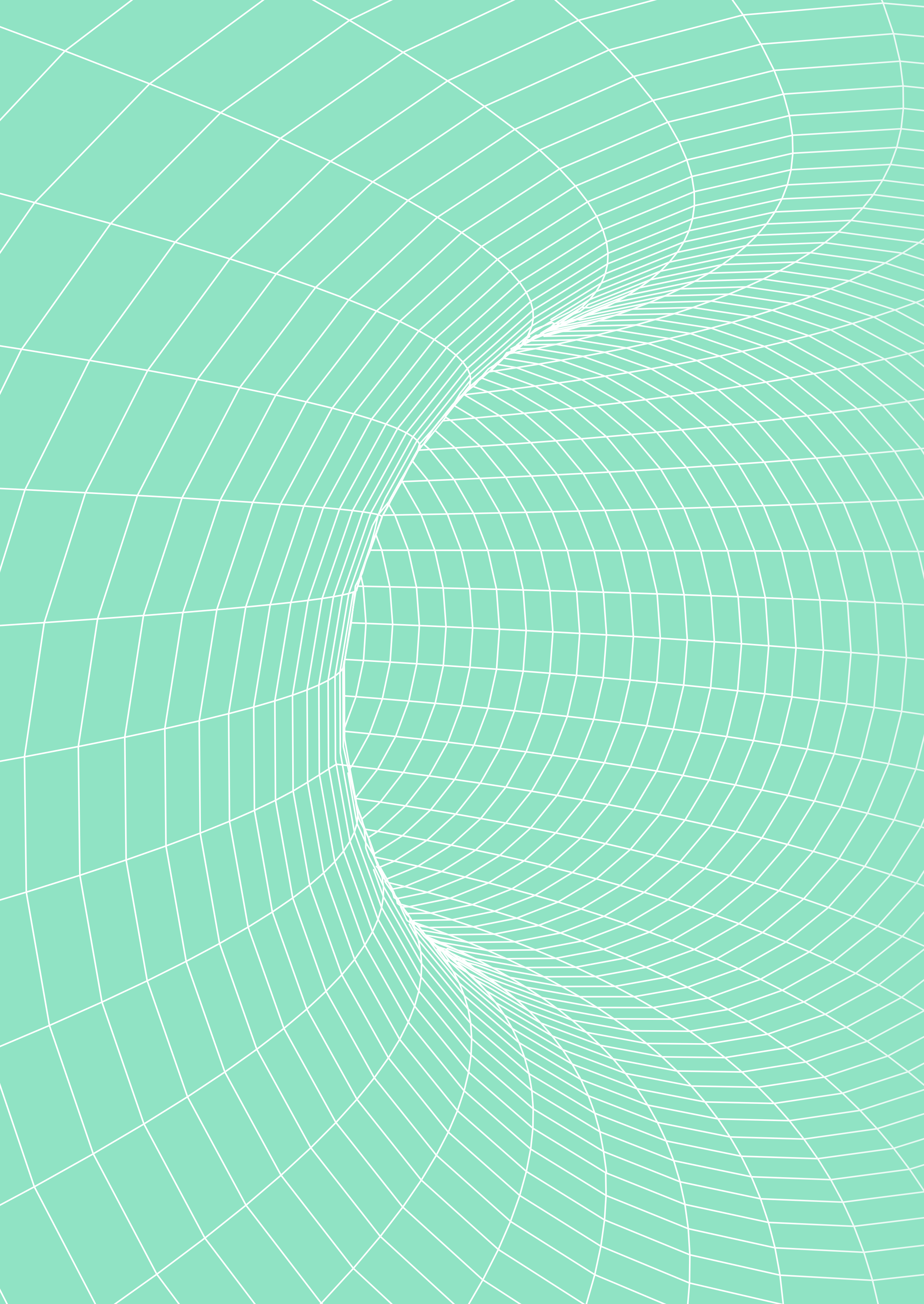
The social value orientation slider measure (SVOSM) is favored over other SVO measures due to its easy implementation, a low burden on respondents, clear consistency checks, high test-retest reliability, and usage of continuous SVO scores (Bakker & Dijkstra, 2021; Murphy et al., 2011; Murphy & Ackermann, 2014). Open questions were whether the SVOSM is reliable in non-student samples and over longer periods than one-week test-retest schemes. Our results show that this is indeed the case. Moreover, additional analyses allow us to recommend refraining from categorizing continuous SVO scores since imposing boundaries on a continuous SVO scale heavily affects the stability of SVO. The attrition analysis shows that, while this study suffered sample attrition, dropouts do not differ significantly in SVO from respondents who did participate in later waves. Moreover, even with attrition, we had a sizeable sample of respondents.

Future work should investigate whether the stability of SVO also translates into stable predictions of cooperative behavior over time. Although major differences in SVO de- pending on monetary or non-monetary incentives are generally not expected (Balliet et al., 2009), findings from Reyna et al. (2018) suggest otherwise. Thus, future research should study the extent to which incentives affect the stability and predictive power of the SVOSM. Specifically, knowing whether monetary or non-monetary incentivized SVO measurements better predict cooperative behavior would be valuable. Future research may also want to consider how individual characteristics and personal or social events — such as changes in income or occupation, experiences with voluntary work, ego depletion of guilt which shows to reduce prosocial behavior (Baumeister et al., 1994) or social integration — influence the stability of one's SVO.

8 Mathematically, the benchmark represents $\pi/4$, a perfectly straight diagonal line in a plot. For more information on describing SVO as a degree angle score, we refer to Murphy et al. (2011).

The high test-retest reliability found in Chapter 2 resembles the stability observed for other personality measures related to cooperativeness. Van Lange (1999) reported a 59 percent consistency score over 19 months with repeated measures using the SVO triple-dominance and ring measure. Bakker & Dijkstra (2021) found consistency percentages of 78, 71, and 67 for the slider, ring, and triple-dominance SVO measure, respectively, over a three-month period. Still, Van Lange and Bakker and Dijkstra utilized said SVO measures as categorical SVOs even though the ring and slider measure allows us to assess SVO as a continuous construct. Akin to our results, some long-term variation is found in the test-retest reliability among categorical SVOs. Moreover, similar accounts of high test-retest reliability are reported for the HEXACO (Dunlop et al., 2021), NEO (McCrae et al., 2011), and Big Five (Henry & Mõttus, 2020) personality inventories in which prosociality related to cooperation is assessed. We show that the slider measure can be added to the list.

The prime contribution of Chapter 2 lies in answering the empirical question of whether SVOs, as measured by the slider measure, are relatively stable over time in non-student samples. Our results support classifying SVO as a personality trait. Still, we find stable differences in prosociality: some are innately more prosocial than others. The question of whether SVOs are a reliable input for exclusionary mechanisms is studied in Chapter 3.



Chapter 3

A bad barrel spoils a good apple[1]

Avengers,
assemble.

—A quote by Chris Evans as Captain America
in the movie *Avengers: Endgame* (2019)

[1] This chapter is the result of joint work with **Andreas Flache**, **Dieko Bakker** and **Jacob Dijkstra**, which appeared in the *Journal of Artificial Societies and Social Simulation* in 2022 under the title "*A bad barrel spoils a good apple: How uncertainty and networks affect whether matching rules can foster cooperation*." Small modifications are made in comparison to the journal version to stay in line with the central tenet in this dissertation.

ABSTRACT

Building on the theoretical foundations established in Chapter 2, this chapter investigates the interplay between contextual and behavioral factors when engaging in exclusionary practices within groups and networks. Specifically, Chapter 3 aims to determine whether exclusionary mechanisms within groups effectively achieve their intended goal of separating cooperators from defectors, motivating defector to engage in cooperative behavior in situations where information regarding SVOs and behavior is unavailable, unreliable, or imperfect. Moreover, by accounting for the multiplexity of social network relationships, Chapter 3 examines whether network ties that cut across group boundaries enhance the effectiveness of exclusionary practices within groups. Chapter 3 offers insights into the potential spillover effects that can occur when exclusionary practices are implemented in one context (networks) and affect social behavior in another (groups).

3.1 INTRODUCTION

Cooperation is central to human life and difficult to achieve. For example, students, community members, activists, employees, or scholars must join forces with their peers to realize benefits they could never generate alone. Yet, individuals are also tempted to free-ride on others' efforts. This jeopardizes the successful cooperation they would like to benefit from in the first place (Chapter 1; Heckathorn, 1996; Olson, 1965; Simpson & Willer, 2015). Among a range of viable solutions to this "social dilemma" (Dawes, 1980; Nowak, 2006), matching mechanisms prevent people who are less inclined to cooperate from entering a group that needs cooperation from its members (Chaudhuri, 2011; Guido et al., 2019). An example would be a student project group whose members only allow peers with high grades to join because they believe that those peers are hard workers. But there are many more real-life examples of matching mechanisms:

> "*[I]n the real world, many mechanisms and institutions exist that are based on the logic of meritocratic matching. Admissions to schools or types of education, for example, are often based on rewards of past school or exam performances which are a function of the work/effort applicants had invested. An important determinant of what makes places that are more competitive to enter `better' is the promise of being matched with others who also performed well in the best*" *(*Nax et al., 2015b, p. 2).

Matching mechanisms exploit stable individual differences in individuals' tendencies to be cooperative, conceptualized, for example, as prosocial value orientation (Balliet et al., 2009; Chapter 2; de Matos Fernandes et al., 2022a) or as personality traits related to altruism or agreeableness (Chapter 4; Thielmann et al., 2020). A successful matching mechanism ensures that only members sufficiently personally disposed to cooperate (hereafter: prosocials) can enter. Matching mechanisms also provide a powerful incentive to behave cooperatively, even for those only motivated by self-interest (hereafter: proselfs). Game theoretical models of so-called *meritocratic matching* show theoretically how matching mechanisms foster cooperation (Gunnthorsdottir et al., 2007; Nax et al., 2017a, 2017b; Nax & Rigos, 2016). Through meritocratic matching, cooperative group members are selected into cooperative and, thus, highly profitable groups. Persistently uncooperative individuals are effectively punished by being left to team up with other persistent defectors in poorly performing groups. Yet, defectors who change their behavior will be rewarded for becoming and remaining cooperative by being allowed to enter productive groups. In other words, a well-functioning and meritocratic matching system under ideal conditions fosters cooperation in a population. The matching system protects genuine cooperators from exploitation by free-riders and incentivizes non-cooperative individuals to act cooperatively.

We contribute to the literature about meritocratic matching in two ways. First, we demonstrate and analyze how imperfect information threatens the success of meritocratic matching. We address the largely overlooked problem that "bad barrels can spoil good apples," referring to situations where imperfect initial matching discourages cooperation among prosocial group members. For example, at the beginning of a course, teachers may match students in groups based on alphabetical order or date of enrollment when other information signaling students' cooperativeness is not available as input for matching. In this case, some prosocial students can end up in project groups comprising many non-cooperative members. This provides a perverse incentive to those prosocials to change their behavior from cooperation to defection to protect themselves against exploitation by their fellow group members. However, "spoiled" prosocials may also have a hard time escaping from unproductive groups because both their (involuntary) non-cooperative behavior and the low performance of the group they reside in make it difficult for other groups to recognize their cooperative intentions.

The "bad barrels" problem occurs to the extent that actors in other groups lack full and accurate information on individuals' "true" cooperative nature. Consider the student groups discussed above. Students may prefer members with high grades to join their project group. Still, these grades can reflect outcomes from earlier group projects in which the final grade was determined on the group-level, not on the level of the individual students. There may be substantial differences in the effort that individual students were willing to invest, but this heterogeneity is not reflected in their grades. It is hard for outside observers to disentangle individual actions from the group context.

To study situations where imperfect information undermines meritocratic matching as a solution to cooperation problems, we developed an agent-based computation model (ABCM) in which cooperation decisions are based on a learning process. Using this model, we analyze the conditions and mechanisms under which imperfect information about individual cooperation jeopardizes the effectiveness of meritocratic matching. Specifically, we compare different information rules on the degree to which they effectively promote cooperation. To be clear, "rules" do not refer to exogenously imposed institutions but reflect different conditions in which agents have (in)complete information due to individual and contextual constraints. The baseline for this comparison is the standard implementation of meritocratic matching based on full information regarding individual merit. We thereby deviate from the conventional full rationality assumption underlying meritocratic matching. To summarize, we investigate whether and under which information conditions meritocratic matching is meritocratic enough in an uncertain world.

Our second contribution to the literature is our investigation of informal social networks as a possible solution to the "bad barrels" problem. Social networks provide an additional source of information agents can use for matching. Dyadic interactions in informal social networks provide signals of individual cooperativeness, which are easier to interpret and more explicit. For example, students matched in project groups often also have academic support relations with peers (Brouwer et al., 2018; see Chapters 4 and 6). In these relations, they can learn more about whether these peers are desirable partners for academic cooperation. Therefore, our agent-based computational model incorporates a mechanism describing how network ties cutting across groups provide additional individual information.

In particular, we use our model to analyze whether homophily in informal social networks, one of the most prominent structural features of social networks, helps to restore the effectiveness of meritocratic matching in a world of imperfect information. Homophily refers to the tendency to preferentially connect to similar others in a network (Lazarsfeld & Merton, 1954; McPherson et al., 2001), driven by shared attributes (gender or educational background) or geographical closeness (neighborhood). Similarity may arise from sharing a status (gender, educational background) or value attribute (attitudes, behavior). We stress in the next paragraph how value homophily in cooperation may arise as a byproduct of status homophily. We show how homophily conditions the effectiveness of informal social networks as an additional source of information for overcoming the bad barrels problem.

As said, a potential source of homophily by cooperation is a byproduct of similarity in other dimensions, often coined as multidimensional homophily (Block & Grund, 2014; Hooijsma et al., 2020). The concept of multidimensional homophily builds on the notion that actors are homophilous on various attributes as a byproduct of a more dominant preferential selection process. For example, similarity in cooperation among network partners may arise as a byproduct of gender or grade homophily. The linking assumption here is that actors sort in the network for reasons unrelated to cooperation per se but tend to trust others who are more similar to themselves (in-group bias or favoritism; Balliet et al., 2014; Fu et al., 2012) and feel more attachment to more similar others (e.g., the attraction paradigm; Byrne, 1971). As a prime example of in-group favoritism or out-group bias, the literature on affective polarization indicates that Democrats and Republicans in the US view the other negatively, affecting the willingness to partner up with opponents (Iyengar et al., 2019). Cooperation in such instances may primarily occur with others similar in the political orientation domain. Our model indirectly includes these considerations and assumes homophily by cooperation as a byproduct network-sorting feature. Relatedly, recent work shows a correlation between sociodemographic attributes and cooperation. For example, economics students are more likely proselfs than other students (Marwell & Ames, 1981). Also, Chapter 4 shows that friends are more likely to work together (Brouwer et al., 2018). In other words, personal predispositions toward certain forms of cooperative behavior are more likely to occur for individuals with similar characteristics or socialized in some similar way.

Combining these differences in cooperativeness with the tendency to preferentially connect to socio-demographically similar others, informal social networks are likely to be homophilous also in terms of cooperativeness. The importance of homophily for cooperation is also reflected by the fact that individuals cooperate more willingly with similar others (Melamed et al., 2020). An informal homophilous network link may help mismatched “good apples” escape from groups where defection prevails by giving them a chance to dyadically show behavior that convinces their network neighbors (and likely members of good groups) of their genuine cooperativeness. We thus argue that homophily improves the chances of spoiled prosocials being identified as potentially valuable candidates for future groups.

But there is also a downside to homophily in social networks. If prosocial actors can quickly escape uncooperative groups by displaying their cooperativeness in informal social networks, proselfs increasingly find themselves stranded in poorly performing uncooperative groups. This undermines the other mechanism through which meritocratic matching works: The provision of incentives for defectors to change their behavior. The more proselfs are concentrated in a group, the more difficult it will become for them to change their ways. Homophily would further exacerbate this problem by restricting their network interactions with other non-cooperative individuals. Thus, our second contribution to the literature is that we use our ABCM to clarify the network conditions under which homophily promotes or jeopardizes the effectiveness of meritocratic matching in a world of imperfect information.

In Section 2, we discuss earlier formal models of meritocratic matching and show how we build on and move beyond this work, formulating three intuitions about the implications of the mechanism the ABCM implements. Section 3 describes the ABCM, and Section 4 presents a detailed analysis of the information conditions and network conditions under which meritocratic matching helps solve the conundrum of cooperation.

3.2 THEORETICAL FOUNDATIONS AND NEW INTUITIONS

Prior affiliations to groups, organizations, firms, or teams can serve as signals of an individual's potential merit under uncertainty (Bacharach & Gambetta, 2001; Gambetta, 2009; Spence, 1973). For hiring committees, for example, the previous affiliation to a reputable firm is a signal of an employee's unobservable "true" individual qualities. Similarly, past affiliation with a fraudulent firm may be interpreted as indicating bad qualities. Signaling theory proposes that the rational use and interpretation of the information conveyed by signals sustain trust and cooperation in an uncertain world (Bacharach & Gambetta, 2001). This type of signaling assumes that individuals rationally display and read such signals so that credible signals (e.g., cooperative behavior that is prohibitively costly for proselfs) differentiate between genuinely prosocial and proself types. In our model, cooperative behavior is the only available behavioral cue. A rationality assumption on which signaling theory rests, related to classical rational choice theory, is that individuals have the unlimited cognitive capacity to process signals and information. This rationality assumption is challenged by decades of research, demonstrating that people are boundedly rational, incompletely informed, and cognitively constrained (Wittek et al., 2013). For the analysis of meritocratic matching under uncertainty, this is a particularly relevant concern, as elucidating individual cooperative signals from behavior in groups requires a lot of cognitive capacity and information processing, and even in dyadic interactions in a network, cooperative behavior cannot be separated from the network context. Led by the critique of rationality assumptions, we rely on simple "low rationality" decision heuristics to explore what happens if agents select new group members, and group members rely on similarly simple heuristics to decide whether to cooperate or not in a given group or dyadic context and then to relate observed behavior to infer cooperative traits.

Our work advances earlier ABCM literature linking cooperation to matching mechanisms. First, Bowles and Gintis (2004) used an ABCM to show how cooperative strategies that ostracize free-riders from groups can thrive and foster cooperation in an evolutionary context. Similar to meritocratic matching, their key mechanism is that ostracized agents are less likely to be accepted into cooperative groups in the future. However, the authors rely on random matching rather than matching based on merit and do not incorporate informal social networks. The work by Bowles and Gintis builds on having the option of playing the game optionally or exiting the game. Suppose agents have the option to leave an interaction altogether. The literature shows that cooperation thrives because cooperators tend to stick together and interact more frequently than defectors, who are more prone to exit (Orbell, Schwartz-Shea, & Simmons, 1984; Schuessler, 1989). Such an exit strategy is not part of the model because we build on the assumption that many cooperation dilemmas do not have outside options where players can work alone or choose not to do anything (e.g., when collaboration is required for coursework or employees are assigned to work on projects).

Second, Duca et al. (2018) studied how meritocratic matching is affected by heterogeneity in endowments. Assuming myopic best response behavior, they show how inequality can strongly hamper the effectiveness of meritocratic matching. Our work introduces inequality likewise, albeit in individual cooperativeness traits rather than endowments. Unlike Duca et al., we add the assumption that information about individual merit is unreliable but that networks can provide additional information. Third, we build on Nax et al. (2015a) who were among the first to show, using simulations of an evolutionary imitation dynamic, how reliance on group merits (*group scoring*) during matching deteriorates cooperation compared to individual merit-based matching. Interestingly, they find that cooperation still arises even when there is only a 1% chance that individual merits are used instead of group merits. However, their model neglects heterogeneity among individuals as well as network solutions and does thus not allow us to highlight conditions for the bad barrels problem we identify. Unlike Nax et al., we focus on conditions under which the bad barrels problem arises.

We follow earlier work modeling bounded rationality with learning theory and evolutionary models to explicate how bad barrels can spoil good apples when it comes to cooperation in groups. To be more precise, simulation studies of evolutionary dynamics in cooperation problems highlighted how cooperation can thrive as a successful strategy, but only when combined with (in)direct reciprocity (Axelrod, 1984; Nowak & Sigmund, 1998). If others defect, then even cooperative recipients are more prone to reciprocate defection. In other words, a good apple can be spoiled by contact with a bad apple. Simulation models based on reinforcement learning lead to a similar conclusion, especially when combined with the assumption that being exposed to others' defections leads initially cooperative agents over time to lower their expectations and be content with low cooperation as an outcome (Macy & Flache, 2002). To be sure, there are differences between explanations of cooperation based on evolutionary dynamics (Nowak, 2006) and explanations based on stochastic learning models (Macy & Flache, 2002). Stochastic learning stresses that changes in cooperation are not driven by payoff-dependent variation in rates of offspring across different strategies or types of agents – as in evolutionary selection – but by variation in the likelihood that agents choose particular cooperation or defection over time. Despite this difference, agents increasingly adopt behavior associated with better outcomes under both approaches. In this chapter, we choose a model based on stochastic learning because we believe that success-driven change of behavior within agents better captures the decision-making of human social actors than the assumption that behavioral strategies are fixed, and all change comes from mutation and selection (for a statement reflecting on this critique on the use of evolutionary algorithms in ABCM, see, e.g., Chattoe-Brown, 1998).

Combining and advancing the perspectives of signaling theory, bounded rationality, reinforcement learning, meritocratic matching under heterogeneity, and uncertainty, we develop, in what follows, a set of theoretical intuitions that serve to guide the design of our ABCM and simulation experiments. Earlier work leads to the intuition that prosocials develop a low level of cooperation through reciprocity if mismatched into groups with many non-cooperative members. Thus, when outsiders are incapable of perfectly inferring the qualities of a group member by observing group outcomes and individual contributions, these "good apples" are spoiled. Various strands of literature support the assumption that human decision-makers tend to (falsely) infer individual qualities from group characteristics, as suggested, for example, by research on fundamental attribution error (Ross, 1977) and statistical discrimination (Fang & Moro, 2011).

> **Intuition 1**. Due to mismatching, prosocial agents cooperate less when matching is based on agents' prior group performance.

In order to escape the negative reputation of a poorly performing group, innately cooperative individuals need some other channel through which they can show their individual qualities. Dyadic interactions in informal networks allow for the development of individual reputations (Buskens & Raub, 2002; Raub & Weesie, 1990). Once cooperative reputations have been established in such dyadic interactions, agents from other groups can use the network information in addition to the information in the group context when determining the matching of agents into new groups. This allows them to eventually achieve higher levels of cooperation.

> **Intuition 2**. The possibility to signal prosociality in dyadic network interactions increases cooperation among prosocials.

In a homophilous network, cooperative types are more likely to cluster together and mainly, but not exclusively, interact with other cooperative types. Homophily thus increases the chances that members of highly cooperative groups interact with mismatched "good apples" from low-performing groups. It thus further improves the possibility of mismatched cooperators being "spotted" as potentially promising recruits.

Intuition 3. Network clustering and information from dyadic network interactions increase cooperation levels of formerly mismatched prosocials in the group context.

3.3 THE AGENT-BASED COMPUTATIONAL MODEL

Agents in our model are either prosocial or proself. Prosocial agents are more cooperative, while proselfs are more egoistic and defection-oriented. Figure 3.1 depicts the basic interaction structure. The matching procedure places each agent into one group. Each group produces its own local collective good (Figure 3.1a). Figure 3.1b shows how the same population of agents is connected by an informal network of dyadic relations that potentially links agents across group boundaries. Both in their group as well as in dyadic interactions, agents are confronted with cooperation problems. These cooperation problems are modeled as iterated *n*-person Prisoner Dilemmas (PD) at the group level and iterated 2-person PDs at the dyadic level, respectively. From time to time, agents can decide to leave their current groups, and groups need to admit new members (rematching). After rematching, a new iterated PD game is started in all groups. Throughout the entire simulation, agents play bilateral PD games with one of their network partners at randomly selected moments. Thus, they sometimes decide whether to contribute to their group's collective good and cooperate in the ongoing private interaction with a particular network partner in the same iteration.

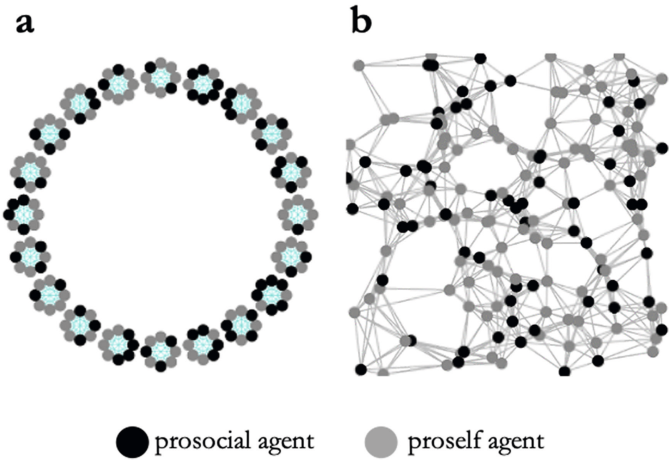


Figure 3.1: Stationary set-up of the model. Agents are embedded in a single group (a) and network (b). Magenta-colored ties show links within the group; grey ties are network ties.

In what follows, we describe the various elements of the model. More precisely, we elaborate on the behavioral and learning algorithms for cooperation, the implementation of prosocial and proself agents, the timing of cooperation decisions in groups and network dyads, the network model, and the different matching rules we compare to assess the effectiveness of meritocratic matching under different conditions. We end with the design of our simulation experiments. The pseudocode of a simulation run is freely available online (de Matos Fernandes et al., 2022b).

3.3.1 Cooperate or not? The decision-making model for cooperation

Cooperation is modeled with a probabilistic threshold model (Macy, 1991b, 1991a; Macy & Evtushenko, 2020; Mäs & Opp, 2016). We apply the model both in the group and in dyadic interactions. Somewhat simplified, agents cooperate if enough others in their group (or network partner) also cooperated in the past; otherwise they defect. How many others are "enough" is defined by an agent-specific threshold. Cooperative types have lower initial thresholds than non-cooperative types. All other things being equal, prosocials are thus more likely to behave cooperatively. Yet, others' behavior also affects agents' propensity to cooperate. This happens through reinforcement learning. Agents become more likely to repeat a behavior associated with a satisfactory outcome and avoid behavior that resulted in an unsatisfactory outcome. Generally, if cooperation (defection) generates a positive outcome, thresholds decline (increase), making cooperation more (less) likely.

We now explain the model for the cooperation decision in the group. Its application to the network game is explained further below. Figure 3.2 provides an overview of the decision and learning sequence. First, each agent compares their current threshold ($\tau_{i,t}$) to the most recent proportion of cooperation by group mates (k_t). Second, all agents decide probabilistically in a random sequence to cooperate or defect (Equation 3.1; $p_{i,t}$). Third, after all decided ($c_{i,t}$), each agent calculates their payoff (Equation 3.2; $s_{i,t}$) and standardized outcome (Equation 3.3; $o_{i,t}$), subsequently applying a learning heuristic to adapt the threshold accordingly for the next iteration (Equation 3.4).

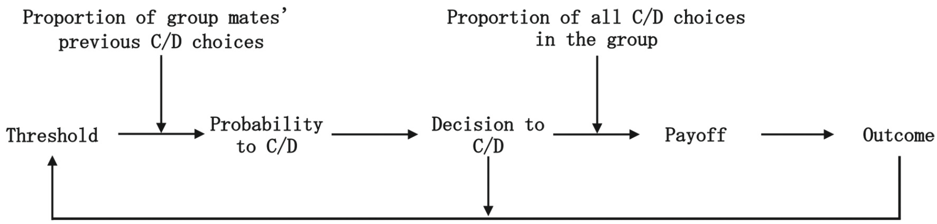


Figure 3.2: A schematic overview of the threshold model. C = cooperation; D = defection.

In the first iteration of a game, others' cooperation is unknown. Initial behavior is governed by agents' innate characteristics given by the initial threshold (τ_i) such that the lower τ_i, the higher the probability of initial cooperation, $p_{i,t}(c_{i,t} = 1) = 1 - \tau_i$ (Mäs & Opp, 2016). After the first decision, thresholds and behavior then change based on past outcomes, reflecting adaptive learning within a group or network relation over time. More precisely, the more the current rate of cooperation in a group (k_t) exceeds an agent's adaptive threshold ($\tau_{i,t}$), the higher the probability of cooperation. Equation 3.1 formalizes the logistic function modeling this link. The slope parameter m controls the degree of randomness in agents' decisions. The higher m, the more the cooperation decision is determined by the difference between the threshold and past group cooperation rate.

$$p_{i,t}\,(c_{i,t} = 1) = \frac{1}{1 + exp[m\,(\tau_{i,t} - k_t)]} \tag{3.1}$$

where $0 \leq \tau_{i,t} \leq 1$, $0 \leq k_t \leq 1$, and $m \geq 1$.

After all agents decided, each agent calculates its payoff, denoted by $S_{i,t}$ in Equation 3.2. The cost of cooperation is 3 (h), while the benefit of cooperation is 4.5 (b). These payoffs constitute a PD in which agents are tempted to defect. For cooperators, we multiply b by the count of cooperative acts in the group ($v_{c,t}$) and divide it by group size (FS) to calculate payoffs minus the cost of cooperation (h). Defectors benefit from cooperating others while not paying the cost of cooperation. However, agents receive -0.5 (d) when all defect. Thus, when all agents defect this is detrimental to both the agent and the group.

$$s_{i,t} = \frac{b\,(v_{c,t})}{FS - h} \tag{3.2}$$

where $\tau_{i,t} = 0$ if $c_{i,t} = 0$, and $s_{i,t} = d$ if $c_{i,t} = 0$ and $v_{c,t} = 0$.

After calculating payoffs, agents compare their payoff of the current iteration to their payoff in the previous iteration, followed by dividing the difference by three times the maximum payoff possible. This yields a standardized outcome $o_{i,t}$ specified in Equation 3.3. Current payoffs weigh more heavily in $o_{i,t}$ than the payoff in the previous iteration. Essentially, the higher the current payoff, the higher the standardized outcome, and the more likely behavior that led to this satisfactory outcome is reinforced. The rate at which thresholds adapt is controlled by the learning rate (l). If $o_{i,t} = 0$, we set $o_{i,t}$ to 0.00001 to ensure that thresholds are updated.

$$o_{i,t} = \frac{l\,(2s_{i,t} - s_{i,t-1})}{3\,|s_{max}|} \tag{3.3}$$

where $0 \leq l \leq 1$.

Finally, agents update their threshold based on $c_{i,t}$ and $o_{i,t}$ (Equation 3.4). If cooperation is associated with $o_{i,t} > 0$, thresholds drop, increasing the chances of future cooperation, while outcomes $o_{i,t} < 0$, increase the threshold following cooperation. The same principle holds following defection. Outcomes o$O_{i,t} > 0$ increases the threshold and, thereby, the probability of future defection, while a negative outcome reduces both.

$$\tau_{i,t}{+}1 = \begin{array}{l} \tau_{i,t} - \{o_{i,t}\,[1 - (1 - \tau_{i,t})^{(1/|o_{i,t}|)}]\; c_{i,t}\} \\ + \{o_{i,t}\,[1 - (1 - \tau_{i,t})^{(1/|o_{i,t}|)}](1 - c_{i,t})\} \end{array} \tag{3.4}$$

3.3.2 Prosocial and proself agents

Agents are randomly selected to be prosocial or proself based on a given proportion of prosocial agents in the population. We assume that prosocial types need less external motivation to cooperate at first, implemented by the assumption that their initial thresholds (τ_i = 0.3) are lower than those of proselfs (τ_i = 0.7). Agents' first decision after a rematching phase is governed by their initial threshold (τ_i), reflecting their innate cooperativeness. Thus, agents reset after matching.

3.3.3 Discrete-time steps per iteration

In an iteration of the group game agents decide, in random sequence, to cooperate or defect. An iteration is divided into discrete time steps. Each agent has the same probability (1/*n*) to be selected per time step. Due to a-synchronicity, agents may have different values for the perceived proportion of cooperation in the group, depending on prior cooperation and defection decisions in previous discrete-time steps. The iteration ends when all agents decided at least once to cooperate or defect, calculating their payoffs, and finally updating their threshold.

A different number of time steps is used for the network 2-person PD, reflecting that interactions with network partners occur in a different context and at a different pace than interactions in the group game. Specifically, in the network context, each dyad has an r chance (r = 0.05) to be selected per iteration; this means that each agent has a 10% likelihood to play the game in any given iteration (and a 90% chance not to play a network 2-PD in the given iteration). Hence, the chances to play the 2-person PD are slimmer than playing the *n*-person PD in each iteration. The value of r = 0.05 is chosen to ensure that cooperation is learned slowly enough in network interactions. In this way, behavior in network interactions is not fully determined by an agent's type but is still a signal of it. Different values for r were explored in a sensitivity analysis, available in Appendix A7.

3.3.4 Social network

3.3.4.1 Random spatial graph algorithm

Following earlier ABCM studies (Grow et al., 2017b; Keijzer et al., 2018), we adopted a spatial random graph algorithm (Wong et al., 2006) to generate the network structure. We rely on a NetLogo algorithm, freely available in the CoMSES computational model library (Grow et al., 2017a). This algorithm can create networks with structural features resembling real-life social networks, such as a high level of clustering and short average geodesic distances. Its core idea is that agents are assigned random coordinates in a two-dimensional space and that then network ties between agents are created such that geographically close agents are more likely to be linked than geographically distant ones.

The network is as simple as possible implemented, meaning statically. The network is generated in two steps. First, agents are randomly dispersed in the cellular world. That is, we independently draw a random x and y coordinate from a uniform random distribution of coordinates, followed by assigning agents to the randomly drawn x-y coordinate if the spot is empty. The network algorithm assumes that agents tend to form ties with those geographically nearby. Second, agents are probabilistically linked to nearby agents. Each agent asks k other non-tied agents to form a tie with. Following previous implementations (Grow et al., 2017a, 2017b; Keijzer et al., 2018), we set $k = 5$. Agents form network ties with at least 5 others. The probability of tie creation is driven by the Euclidean distance in the cellular world (Equation 3.5). Two agents who are geographically closer to each other are more likely to form a tie than two agents who are farther away. Dyadic closeness is denoted as u_{ij} The average geographical distance does not affect tie formation processes when $w = 0$. Contrarily, higher w values increase the importance of distance on tie formation, facilitating clustering levels observed in real-life networks. To ensure a representative level of network clustering, we set $w = 8$ (Grow et al., 2017a, 2017b; Keijzer et al., 2018). Agents form thus more easily ties with geographically nearby agents.

$$f(w, u_{ij}) = e^{(-w[u_{ij}])} \tag{3.5}$$

Here, we concentrate on how we adapt it to induce homophily. Figure 3.3 visualizes how structural homophily between prosocial agents is imposed in the spatial random graphs. In random networks (Figure 3.3a), both the geographic location of prosocial and non-prosocial agents and, thus also, the probability for a network link between two agents are unrelated to their type. In homophilous networks (Figure 3.3b), the geographic allocation is such that prosocial agents are locally clustered so that the algorithm more likely links agents of the same type to each other than would be given by random chance. To assess the resulting degree of type-homophily in the network, we adopt Moody's gross-segregation (*MS*) index (Bianchi et al., 2020; Moody, 2001). Intuitively, the interpretation of the measure is that it is *MS* times as likely for a network link to occur in a dyad of same-type agents than in a dyad of agents of different types. The coefficient allows us to compare the likelihood of ties between the same and other-type agents. In Equation 3.6, details of the implementation of *MS* are given.

$$MS = (AD) \,/\, (BC) \tag{3.6}$$

Formally, the count of same-type ties (prosocial-prosocial and proself-proself, *A*) is multiplied by all possible other-type ties (all proselfs * [all prosocials - 1] and vice versa, *D*). Next, *AD* is divided by the count of all other-type ties (prosocial-proself and proself-prosocial, *B*) and all possible same-type ties (all proselfs * [all proselfs - 1] and vice versa, *C*). *MS* reports an odds ratio (OR). An OR of 1 denotes that the chances for a link between same and other-type agents are equal. An OR < 1 shows that the odds of linking to same-type agents are lower than linking to other-type agents. Contrarily, an OR > 1 indicates that the probability of linking to same-type agents is higher than linking to other-type agents. A feature of the coefficient is that the relative group size does not affect the interpretation of the coefficient.

In the randomly-linked networks, the chances to link to the same and other-type agents are about 50/50 ($MS \approx 1.00$)[2]. But if we allow for structural homophily, the chances to link to same-

2 Estimate is based on 100 replications in BehaviorSpace in a population constituting either 20, 40, 60, and 80 percent prosocials.

type agents need to increase. We introduce a stylized implementation of increasing the chances of linking to same-type agents. With homophily, we assign prosocials to a fixed area in the cellular world and draw random x-y coordinates from that restricted area, promoting the clustering of prosocials due to forming ties with geographically nearby agents. Subsequently, proselfs randomly draw x-y coordinates from the whole grid and move to the drawn coordinate. Prosocials agents are first dispersed in the network's lower-left half when the population consists of 20% prosocials. Whereas prosocials are sorted in the lower half of the cellular world when there are more than 40% prosocial agents.

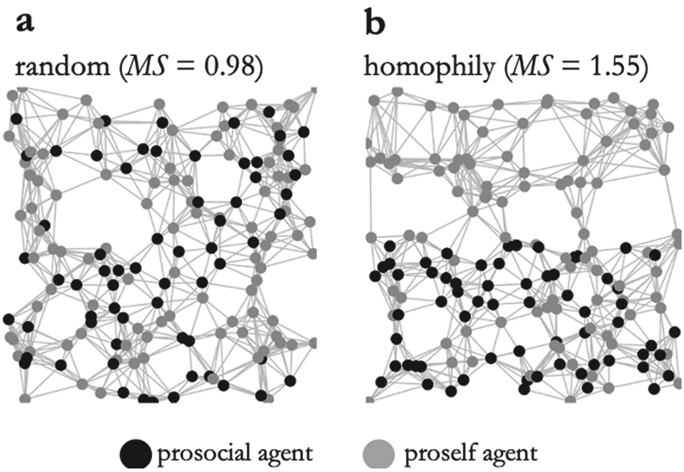


Figure 3.3: Visualization of two spatial random graph single runs with a random (a) and homophilous (b) network. Moody Segregation index (MS) refers to the odds ratio for a link to occur between similar and dissimilar agents.

The stylized set-up of structural homophily ensures a *MS* between 1.5 and 1.8, on average, irrespective of the percentage of prosocials in the population. Estimates are based on 100 replications in a population with 20, 40, 60, and 80 percent prosocials. The stochastic nature of the spatial random graph algorithm inserts minor variance in *MS*. There remains, thus, some minor trembling in forming ties with other-type agents when structural homophily is implemented, meaning that the homophily procedure incorporates some faultiness. This implementation also reflects incomplete information in forming ties with similar others. As Figure 3.3 shows, *MS* is about 1.0 in random networks, while in the networks with homophily same-type agents are linked about 1.5 times as likely than different-type agents.

Note that the *MS* does measure perfect segregated situations. Then, *AD* is divided by 0. We solved this issue in the code by setting $B = 1$ when no cross-type ties were available. This occurred in 33 cases when calculating *MS* in the group context, mostly when learning was slow and complete information was available. We omitted these MS odds ratios because these incur skewness. *MS* was in such situations > 260.

3.3.4.2 Reputation formation through interactions in the network

Whenever a dyad in the network is selected to play their 2-person PD game, both players decide whether to cooperate based on the same decision procedure, learning algorithm, and payoff scheme described above for the group game. We add the subscript sn to indicate network parameters. More specifically, agents make separate decisions per tie, governed by the same threshold for every network partner and taking into account each of their alters' previous cooperation decision. After the interaction, agents update their outcome and adjust their single threshold for all network partners for future interaction with potentially different alters. This implementation also means that an agent can cooperate in an interaction with alter *j* but defect with alter *k*. Alter refers to a directly connected agent. Stable cooperation is more likely to emerge in dyads between prosocial players than proself players. For instance, in prosocial-prosocial interactions in which both previously cooperated, both keep cooperating with a probability of 0.97. After an interaction, thresholds of cooperators tend to lower towards 0. The contrary is true for proself-proself interactions in which both previously defected, then the probability to cooperate is 0.03. In proself dyads, defecting will result in negative outcomes and decrease proself agents' thresholds, making them more likely to cooperate in the near future. However, they are still less likely to cooperate than players in prosocial-prosocial interactions, quickly earning them a worse reputation.

In addition to homophily in the structure of the network, we implement homophily in dyadic interactions. Only similarly-behaving agents will play the 2-person PD. Practically, this implementation facilitates cooperator-cooperator and defector-defector interactions. If agent *i* and *j* cooperated in the previous iterations (or defected) and the dyad is selected, they play the 2-person PD; otherwise, they do not interact. This assumption reflects what in a more detailed elaboration of a backward-looking partner selection process would intuitively be the outcome. Players would be satisfied with mutual cooperation with a network partner and thus repeat that interaction. Players who experience mutual defection or exploitation would abandon their partners and try to find better matches. However, sooner or later defecting proselfs can find only other proselfs to connect with due to the reputation they acquired. Further, assuming that actors prefer mutual defection to not interacting in network relations, we assume that mutual cooperation and mutual defection result in repeated interaction with the same partner. Only after defectors change to cooperation, they are available for interactions with cooperators. In Appendix A8, we show that although plausible, this assumption of 'behavioral homophily' is not crucial for the qualitative results of our analysis. Yet, the differences between conditions in our simulation experiment are highlighted more clearly if behavioral homophily is assumed.

The accumulation of individual cooperation and defection decisions in the network yields a personal reputation score, formalized as $C_{10,sn}$ capturing the most recent 10 network decisions of an individual. Furthermore, we assume that one's reputation is known among alters and alters of their alters. For example, agent *i* plays the 2-person PD with alter *j*, but *i* knows how *j* behaved in all of his last 10 interactions. As we will explain next, personal reputations can be used in the matching phase to assess the cooperative qualities of a potential new group member.

3.3.5 Meritocratic matching

The matching procedure is as follows. Agents decide first whether to leave or stay in their group, followed by matching to new groups.

3.3.5.1 Leave-stay procedure

Agents decide to leave a group when they are not happy with the average level of cooperation from the last 10 iterations (G_{10}) in the group. More precisely, we assume that agents stay if past cooperation exceeds their innate threshold, $\tau_i \leq G_{10}$, and leave otherwise ($\tau_i > G_{10}$). Thus, prosocials accept a lower level of cooperation ($0.3 \leq G_{10}$) than proselfs ($0.7 \leq G_{10}$) reflecting their innate cooperativeness even when others defect. However, agents still condition their decision to leave or stay on what others do so that proselfs also leave a group when cooperation drops too low. Sensitivity analyses were conducted to test the alternative effects of the leave-stay procedure (please see Appendices A5 and A6). Next to $\tau_i > G_{10}$, we test the consequences of leaving if $1 - \tau_i > G_{10}$ and $0.5 > G_{10}$. The leave-stay procedure is activated after 100, 200, and 300 iterations. Agents who decide to leave are put into a pool, followed by matching to a new group. Note that leavers start in the new group with their initial threshold ($\tau_i \rightarrow \tau_{i,t}$), while stayers maintain their current threshold ($\tau_{i,t}$). Resetting is done to model the fact that threshold changes depend on social interactions, and it resembles a reset effect for leavers.

3.3.5.2 Matching rules

After a leave-stay decision, groups are ranked from high to low based on the group-specific G_{10}. We assume that all leavers prefer a higher-ranked group to a lower-ranked one. Agents, in turn, are ranked based on their perceived merit. Generally, the matching procedure assures that groups with higher ranks also receive agents with higher perceived merit. More precisely, the procedure starts by assigning as many agents to the highest-ranked group as there are empty slots, starting with the highest-ranked agents, then takes the remaining highest-ranked agent and assigns them to the empty slots in the next highest-ranked group. Note that it is expected that the best functioning groups do not have empty slots to fill because no agent left the group. This procedure repeats until all agents from the pool are matched. What determines merit depends on the exact matching rule and whether reputational information from the networks is available, as will be explained below.

As we set out in the theory section, the severity of the “bad barrels” problem is determined by the extent to which agents lack information about the individual behavior of others. We, therefore, compare three different matching rules to assess how moving from perfect to imperfect information concerning individual cooperation and the possibility of using a network-based reputation for matching decisions moderate the effectiveness of meritocratic matching. Appendices A1 and A2 show some additional rules we explored as the benchmark for comparison with earlier work. Findings for these additional matching rules are reported in Appendix A1. Figure 3.4 shows three matching rules for agents who left their group.

Rule 1. This rule represents our baseline scenario in which agents have complete information about all prior individual cooperative actions (Figure 3.4a). Agents are initially assigned to groups based on their first cooperation decision, which is itself determined based on the agent's initial threshold. This approach limits initial mismatching. Prosocials initially have a 70 percent chance to cooperate, whereas proself agents have a 30 percent chance to cooperate.

Rule 2. This rule allows us to test intuition 1. Agents are randomly matched to groups and no reputational information from the network is available for assessing their individual merit in rematching decisions (Figure 3.4b). To further model incomplete information of agents in other groups, merit assessments are entirely based on the recent level of group cooperation (G_{10}) of an agent's past group (Duca & Nax, 2018). This rule tests whether mismatched cooperative agents can get away from "bad" groups if agents in other groups know the average level of cooperation in the group.[3]

Rule 3. With this rule, we add the possibility that agents in other groups can use individual reputational information from the network. Thus, agents are now embedded into two contexts (Figure 3.4c). Agents in other groups rely during matching on the combination of social network information and group merit to assess an agent's merit: $GC_{10} = (C_{10,sn} + G_{10}) / 2$, where agents store their last 10 social network decisions in $C_{10,sn}$, while G_{10} represents the average cooperation of the last 10 iterations in their previous group. But there is a caveat: Agents only rely on GC_{10} when local network information is available (yellow pane in Figure 3.4c) and use G_{10} when network information is unknown. Agents thus do *not* have global network information. Local network information may not be available if members of the receiving group do not belong to the social vicinity of an applicant. To be precise, agents know $C_{10,sn}$ only if they are alters or alters' alters of an applicant. Otherwise, they can only use G_{10}. The addition of network information is interesting because GC_{10} allows prosocial agents with low G_{10} to increase their chances of joining better groups when $C_{10,sn}$ is high. The contrary is also true. It may be detrimental for agents with a high G_{10} to incorporate a low $C_{10,sn}$. There are two network implementations under rule 3: a random and homophilous network, allowing us to test intuition 2 and 3, respectively (Figure 3.4c).

3.3.6 Simulation experiment

To check whether our intuitions for the model are correct, we conducted simulation experiments via BehaviorSpace in NetLogo (Wilensky, 1999). Our most important experimental outcome is cooperation levels reached for prosocials, but we also zoom in on proself and collective cooperation levels. We choose a scenario roughly inspired by 2 consecutive academic years, divided into 4 semesters, in which students are grouped for a project and can self-organize new project groups after each semester.

3 Appendix A1 shows that relying on the last 10 individual cooperation decisions in the group context does not alter cooperation levels reached under rule 2.

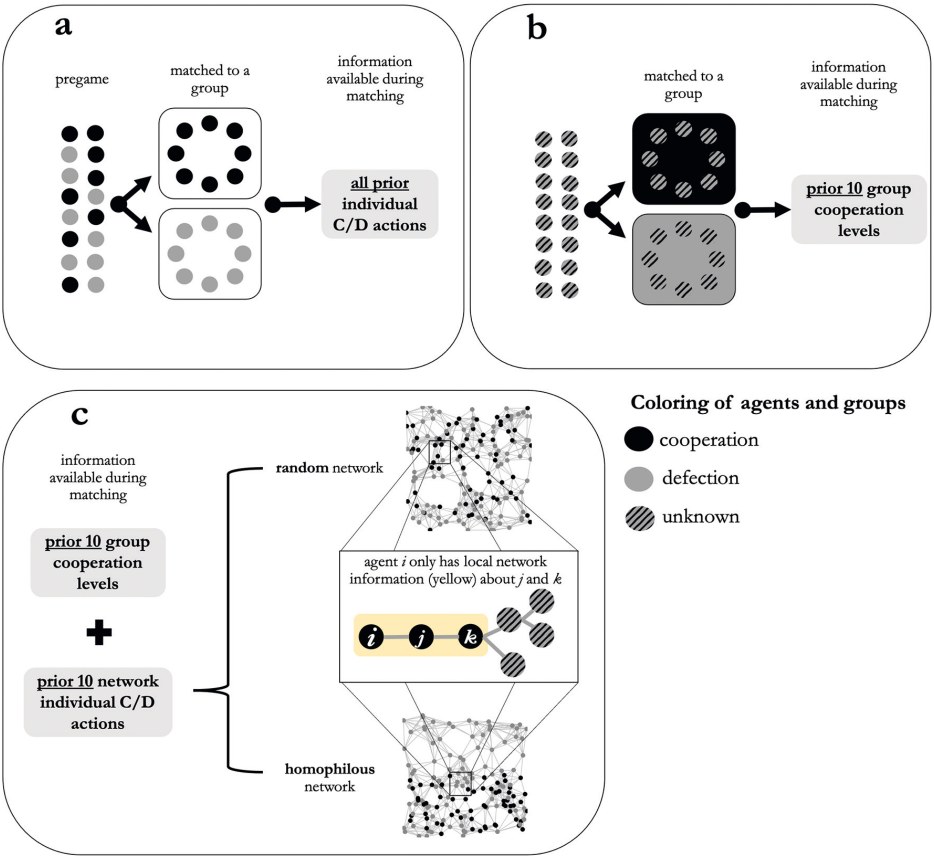


Figure 3.4: A visualization of the three matching rules and information available during matching. Note: C = cooperation; D = defection; all prior = average level of cooperation of all prior iterations; prior 10 = average level of cooperation in prior 10 iterations.

We model a population with $n = 160$ agents placed in $G = 20$ equally sized groups. The population contains a minority of 40% prosocial students ($PA = 0.4$). Agents play an iterated n-person PD for 400 iterations in the groups, with rematching occurring after $X = 100$, 200, and 300 iterations. This assures groups that remain fixed for a sufficiently long period to develop stable cooperation levels. The network is either formed with all dyads being equally likely or based on homophily. The network contains 800 social ties, where each agent has at least 5 network alters.

Regarding the threshold model, we assume a moderate degree of learning ($l = 0.5$) and randomness ($m = 5$), following earlier work (Macy, 1991a). The full parametrization of the model can be found in Table 3.1. In Appendix A, we report the various robustness checks of our findings. The data used in this Chapter and R-script to analyze the data and plot the figures are freely available at the Open Science Framework via https://doi.org/10.17605/OSF.IO/2QPDN (also, see the supplementary material attached to de Matos Fernandes et al., 2022b).

Table 3.1: Summary of parameters and ranges related to the agent-based computational model.

Parameter	Rule	Symbol	Range
Learning rate	1-3	l	{0.1, 0.3, 0.5, 0.7, 0.9}
Slope (noise)	1-3	m	{1, 5, 10}
Adaptive threshold	1-3	$\tau_{i,t}$	(0, 1)
Cost of cooperation	1-3	h	3
Benefit of cooperation	1-3	b	4.5
Payoff when all defect	1-3	d	-0.5
Chance of network dyad selection	1-3	r	{0.01, 0.05, 0.25, 0.5}
Group context features			
Moment of matching	1-3	X	{100, 200, 300}
End iteration of a single run	1-3	E	400
Population size	1-3	$N\{0, \ldots, n\}$	$n = 160$
Fixed group size	1-3	FS	8
Count of groups	1-3	G	20
Proportion prosocial agents	1-3	PA	{0.2, 0.4, 0.6, 0.8}
Fixed (initial) threshold	1-3	τ_i	{0.3, 0.7}*
Leave-stay procedure	1-3	$\tau_i > G_{10}$	{leave, stay}
Noise in leave-stay procedure	1-3		{0, 0.01, 0.05, 0.25}
Matching: Complete information			
Pre-game prob. to cooperate	1	$p_{i,t=0} = 1 - \tau_i$	{0.7, 0.3}*
Individual cooperation	1 & 1 adj.	C_{all}	(0, 1)
Matching: Incomplete information			
Individual cooperation	2 adj.	G_{10}	(0, 1)
Group cooperation	2	G_{10}	(0, 1)
Network cooperation	3	$C_{10,sn}$	(0, 1)
Network and group cooperation	3	GC_{10}	(0, 1)

* The first value refers to prosocials ($\tau_i = 0.3$) and the second to proselfs ($\tau_i = 0.7$).

3.4 FINDINGS

3.4.1 Investigating intuitions 1 – 3

Figure 3.5 reports mean cooperation levels over time, averaged over 100 simulation runs for prosocials, proselfs, and the entire population (collective) per matching rule.

The spikes in Figure 3.5a after matching show prosocial agents initially cooperating with high frequency in accordance with their initial thresholds. However, cooperation soon declines to lower levels than right after the matching moment. The cooperative intentions of prosocials are to no avail in some groups. The loss of cooperative potential after matching points to the presence of mismatched prosocials.

Our model corroborates intuition 1. Comparing matching rules 1 (red) and 2 (blue) in Figure 3.5a indicates that mismatched prosocials are less able to cooperate if meritocratic matching is based on agents' prior group performance.[4] Proself agents are slightly better off when complete individual information is available, but we need to stress that differences between cooperation levels reached under complete and incomplete information rules are marginal (Figure 3.5b). Our model does thus not support what is considered an important strength of meritocratic matching – proselfs do not behave significantly more cooperative over time when their individual merits are visible. Our simulation findings suggest that cooperative agents end up in less-than-ideal groups when matching is based on incomplete information.

Our results also show that incomplete information jeopardizes the collective efficiency of meritocratic matching. Figure 3.5c shows that cooperation rates are highest under complete information, which is collectively optimal under the social dilemma game groups play (Figure 3.5c). What is more, Figure 3.5b suggests that collective cooperation levels are not driven by egoistic agents overcoming their innate inclination to defect. Rather, Figure 3.5a shows that prosocials who cooperate under Rule 1 and Rule 3 (homophily) drive cooperation at the collective level. Consequences of removing meritocratic matching broken down for prosocials, proselfs, and the collective are reported in Appendix A2. Without matching and network information, the model with only a *n*-person PD suggests that the collective fares best when agents in the group, irrespective of group composition, interact for over 10000 iterations without matching to another group (Appendix A2, Figure A2).

We analyzed whether random or homophilous networks help restore the effectiveness of meritocratic matching in a world of imperfect information. On the one hand, intuition 2 proposed that individual information derived from dyadic interactions in a *randomly formed network* mitigates the bad barrels problem (green line in Figure 3.5a). Formerly mismatched cooper-

4 An additional incomplete information rule – see Appendix A1 – corroborates that solely observations of agents' individual behavior in the context of that group also lead to similar lower cooperation levels as under rule 2.

ative agents are better able to signal their prosociality and, therefore, improve their chances of moving into more cooperative groups. However, there is only a marginal increase from the incomplete information rule 2 (blue) and rule 3, in which additional individual network information is available in random networks (green). We cannot confirm intuition 2 for random networks.

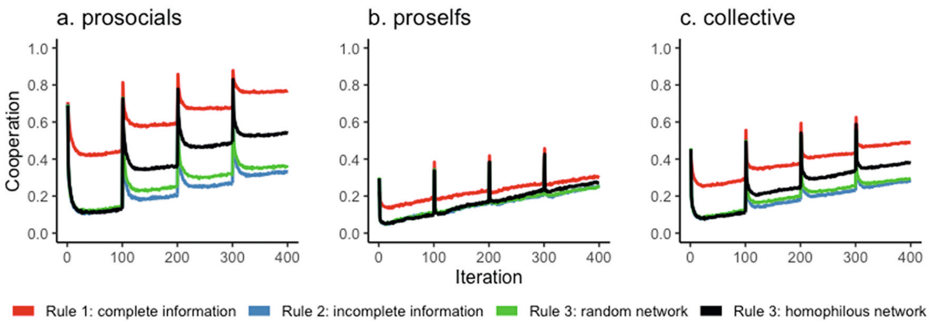


Figure 3.5: Average level of cooperation of 100 independent runs for prosocials (a), proselfs (b) and the collective (c), separated by matching rule. Intuition 1: red vs. blue; Intuition 2: blue vs green; Intuition 3: blue vs. black. Parameter settings: $m = 5$; $l = 0.5$; $PA = 0.4$; $r = 0.05$.

On the other hand, the picture changes radically when homophily is implemented in the network, consistent with intuition 3. The black line in Figure 3.5a shows how cooperative agents increasingly cooperate when information is incomplete, due to the possibility to escape from groups in which defection prevails. In particular, adding homophily increases cooperation rates only for prosocials and not for proselfs (Figure 3.5b). The difference to the random network condition shows the underlying mechanism. Prosocials cooperate more because, due to homophilous networks, they more often succeed in leaving bad barrels and joining groups in which they more readily cooperate. Our findings also suggest that there is still some loss of efficiency due to imperfect information, demonstrated by the large difference between cooperation levels when information is complete or incomplete (red vs. black line in Figure 3.5a).

In the following two sections, we explore underlying reasons why homophily is an important driver for prosocials' cooperation. In a nutshell, we point to prosociality segregation and the impact of homophily on dyadic interactions as underlying reasons for the findings reported in Figure 3.5.

3.4.2 Prosociality segregation

One feature that facilitates the cooperation of prosocial agents is the presence of similar others in the group. Thus, the occurrence or absence of prosociality segregation – i.e., more prosocials in cooperative and proselfs in "bad" groups – may be an important explanans for the reported cooperation levels in Figure 3.5. In Figure 3.6, we use the gross-segregation index (*MS*) to measure the odds of being matched with similar types in the group context (Moody, 2001).

Segregation in the group context is highest when complete information is available (red line in Figure 3.6). Both prosocials and proselfs are three times as likely to be grouped with their own type. Moreover, the *MS* odds ratio value at iteration 0 for complete information shows that initial mismatching is less prevalent than incomplete information conditions. Even if a cooperative agent is spoiled by the mere presence in an uncooperative group, a cooperative effort at the game's early stages still serves as a signal to others when complete individual information is available. This signal, in turn, positively affects prosocials' chances to escape the uncooperative environment and to match to a more cooperative group.

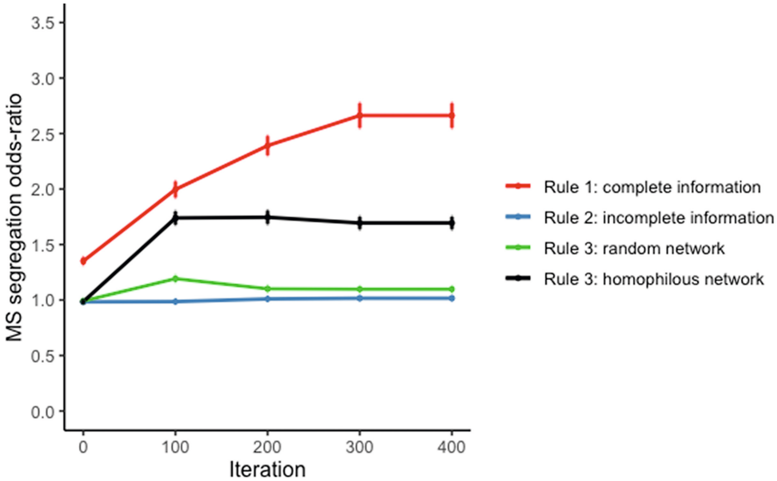


Figure 3.6: Average level of prosociality segregation of 100 independent simulation runs, separated per matching rule. *MS* = Moody gross-segregation odds ratio index. We report 95% confidence intervals at t = 0, 100, 200, 300, and 400. Parameter settings: m = 5; l = 0.5; PA = 0.4; r = 0.05.

Figure 3.6 shows that when agents are embedded in homophilous networks, the odds of joining forces with similar others are around 1.5. Harvesting individual information from a homophilous network allows cooperators to team up, leaving defectors only their own types to be matched with. For prosocials, assorting with similar others promotes more chances to cooperate (Figure 3.5a), while the opposite counts for proselfs. The increases in cooperation in Figure 3.5a appear to be largely driven by mismatched prosocials leaving bad groups and moving to more cooperative groups with many similar others. The contrary is true for incomplete information settings when matching is initially imperfect and remains to be so. "Spoiled" cooperative types may have a hard time escaping from unproductive groups due to the low performance of the group they reside in, making it hard for them to demonstrate their genuine cooperativeness to other groups (green and blue line in Figure 3.6).

3.4.3 Impact of homophily

Here we zoom into how homophily affects cooperative behavior in dyadic interactions and – thereby – the information agents can obtain about potential new group members from their network interactions. In Figure 3.7, we compare a single run of cooperation levels in a random and homophilous network. The full simulation experiments provide a similar picture (Appendix A3). Strikingly, homophily does not increase the cooperation of prosocials (Figure 3.7a), but it reduces the cooperation of proselfs in dyadic interactions (Figure 3.7b). Agents have no choice in random networks but to play the 2-person PD. Such a 2-person interaction scheme in random networks – where there is a 50/50 chance to link to other-type agents – is particularly beneficial for proself agents to learn to cooperate when they have repeated interactions with cooperating others (most likely prosocials). Namely, when interacting with prosocials, proselfs will quickly generate a probability to cooperate of 0.5 $(\frac{1}{1 + exp[5(1 - 1)]})$, in which a random walk from defection to cooperation leads to locking into cooperation. Figure 3.7, green line, shows the tendency toward all-out cooperation in random networks, which, as a result, makes it hard to differentiate between more prosocial and proself agents. While prosocials maintain higher levels of cooperation than proselfs even in random networks (Figures 3.7a and b, green lines), the difference in cooperation rates is small. As a result, dyadic interactions in random networks provide insufficient information to separate prosocials from proselfs. Consequently, dyadic interactions in random networks do not lead to more cooperation in the group context among both prosocials and proselfs (Figures 3.5a and b). As such, information derived from random networks does not serve as an exclusionary mechanism compared to information derived from homophilous networks and therefore does not lead to more cooperation in the group context among prosocials and proselfs.

However, the picture changes when we inspect cooperation levels of proselfs in a homophilous network (Figure 3.7, black line). In such networks, agents mainly have network ties to similar others, and similarly-behaving agents preferentially interact. Cooperators – most likely prosocials – tend to receive cooperative acts in return, while defectors receive mostly defection. The context in homophilous networks has a downside for proselfs as a result of their limited interaction with cooperative others. Proselfs have little opportunity to learn cooperative behavior from interactions with others since the other is most likely a proself type. Meanwhile, homophilous networks enable prosocials to signal that they are cooperative regardless of the group context into which they have been matched. This allows prosocials who find themselves in a bad barrel to nonetheless identify themselves as cooperative. Thus, homophilous networks make prosocials more likely to be identified as good cooperation partners and proselfs less likely to be considered desirable group members. Overall, individual information derived from homophilous networks enables agents to distinguish more readily between prosocial and proself types, consequently allowing prosocials to join forces more easily (Figure 3.6) and cooperate more often (Figure 3.5a). As such, homophily serves as an exclusionary mechanism, clearly differentiating between prosocials and proselfs. Our work, reflecting ideas from earlier research on homophily in networks (McPherson et al., 2001), stresses the importance of homophily as a structural and behavioral process operating in social networks. The pervasiveness of homophily in informal social networks elucidates to what degree cooperative acts and information on others' cooperative behavior flow locally among similar others, contrasting randomly formed networks.

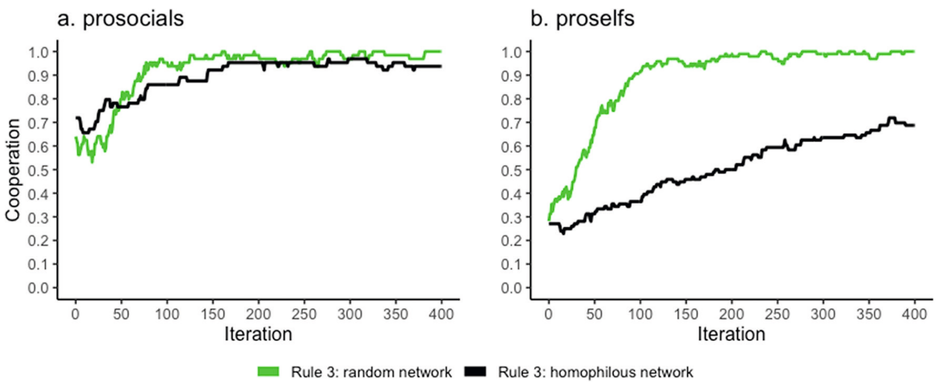


Figure 3.7: The average level of network cooperation in a typical run; one for prosocials (a) and one for proselfs (b) separated by agents' embeddedness in a random (green) or homophilous (black) network. Parameter settings: $m = 5$; $l = 0.5$; $PA = 0.4$; $r = 0.05$.

3.4.4 Sensitivity analysis

We implemented several robustness checks inspecting under which conditions our simulation findings are robust. First, we explored various learning rates (l) since learning dynamics play a pivotal role in solving the problem of cooperation (Macy & Flache, 2002). Second, the presence of more prosocial agents may increase the chances of teaming up with similar others; thus, we inspect the impact of the proportion of prosocials (PA) in the population. Third, noise in the behavioral decision-making model (indicated by m) is bound to play an important role when agents make decisions in threshold models (Macy & Evtushenko, 2020; Mäs & Opp, 2016). We inspect the consequences of more or less noise. Fourth, noise also has a role in the leave-stay procedure in which a proportion of agents who were happy with group performance and therefore stayed in the group will wrongly be put in the leavers pool. On a similar note, we test the consequences of altering input for the leave-stay procedure, either τ_i, $1 - \tau_i$, or 0.5 in relation to G_{10}. Finally, we vary the rate r at which dyadic interactions rather than group interactions occur.

Appendices A4 to A8 provide a comprehensive taxonomy of the various robustness checks with a total of 42600 simulation runs. Our simulation findings regarding cooperation and prosociality segregation turn out to be fairly robust to changes in the proportion of prosocials present in the population (Appendix A4), for learning rates below 0.9 (Appendix A4), and for variations in the leave-stay procedure (Appendices A5 and A6). Notably, a high learning rate ($l = 0.9$) allows agents to learn how to cooperate quickly, providing a different solution than meritocratic matching for cooperation to thrive (Appendix A4). Moreover, we find that the bad barrels problem and homophilous network solution are more pronounced when $m = 5$ and $r < 0.25$ compared to when $m = 1$ or 10 and $r = 0.25$ or 0.5. The sensitivity analyses raise a few points. First, the simulation findings are sensitive to more or less noise in the decision-making model, showing two cooperation equilibria (Appendix A4). More noise ($m = 1$) leads to a self-correcting equilibrium where cooperation levels steadily hover around 0.34, whereas less noise ($m = 10$) leads to a self-reinforcing equilibrium where cooperative agents quickly lock into cooperation (Macy &

Flache, 2002). Second, the importance of complete information rules for prosocials to cooperate is robust to changes in network dyad selection (r), but our incomplete-information-with-network-information solution is not. Figure A10 in Appendix A7 shows when chances for dyad selection increase to values of 0.25 and higher, individual information from homophilous networks does not contribute to prosocials' chances to cooperate more often or to join more cooperative groups with similar others. The reason for model sensitivity to r is found in the inability to differentiate between proself and prosocial agents regarding network cooperation. When $r \geq 0.25$, proselfs more readily learn to cooperate at similar levels as prosocials, even when proselfs are embedded in parts of the social network where initial defection prevails. Finally, we also tested whether model results change qualitatively when we abandon the assumption that homophily is not only affecting the network structure but also who interacts with whom (Appendix A8). While effects become smaller quantitatively, they remain unchanged qualitatively.

3.5 CONCLUDING REMARKS

Our work has uncovered a limitation of meritocratic matching. Information availability strongly affects the matching mechanism's ability to generate cooperative groups. Complete information on individual predispositions provides ideal conditions for meritocratic matching. We introduced several matching rules to analyze the consequences of incomplete information on model outcomes. When only group-level information is available for the matching mechanism, prosocials end up not fully exploiting their cooperative potential, hindering cooperation in general. We also asked whether social network information can solve the bad barrels problem. Our simulations show that if prosocial agents have access to individual information derived from homophilous networks, they can mobilize more of their cooperative potential. Homophilous networks improve the functioning of meritocratic matching systems by allowing cooperators to identify other cooperators. Agents preferentially connect to and interact with similarly behaving others in the network: Cooperators mainly interact with cooperators, while defectors are left to interact with other defectors. This creates ideal conditions for mismatched prosocial agents to display their cooperative tendencies, as they do not have to fear exploitation by uncooperative network partners. Dyadic interactions thus increase *differentiation* between prosocials and proselfs. In addition, homophilous networks create groups of prosocial agents who are aware of each other's behavior. The stronger this *prosociality segregation* is, the better prosocials can cooperate in the group context. The availability of information on prosocial others and the relative effectiveness of behavior in homophilous dyadic network interactions helps prosocials group up, resulting in more cooperation in the group context.

Chapter 3 comes with limitations that suggest avenues for future research. One limitation pertains to comparing the value of network cooperation to group cooperation, which may be context-dependent. Our robustness checks showed that differentiation in the frequency of interactions (r) matters. But the goals of work-related teams may also differ from the social goals of inter-employee friendships. Some may not even want a spillover between the friendship and work domain. This may limit how networks help solve the "bad barrels" problem. The problem is further perpetuated when merit information from the group context surpasses the importance of network-based merit. However, for important contexts, it seems plausible that network information is sufficiently reliable and relevant to improve selection decisions. For example, a scientific department recruiting new staff may want to mobilize informal collaboration networks of employees with many ties to applicants working in the discipline, to collect more individual information about applicants. Especially when work-related information is lacking or unreliable.

An important topic highlighted by our model is the tension between what is individually or collectively optimal in meritocratic matching. Prosocials fare better under meritocratic matching, but proselfs – and thereby the collective – may need more time to follow suit.[5] This tension, i.e., maximizing collective benefits that arise out of cooperation and minimizing individual differences in benefits, finds its roots in the classical societal problem that the meritocratic system aims to attenuate: Inequality. One way to suppress inequality as an outcome is to bolster equality in opportunities—a core tenet of meritocratic matching since it leads to equal opportunities in principle. But meritocratic matching may also perpetuate inequality by shifting it to merit-based inequality. For instance, the ideal situation occurs when proselfs quickly recognize that they need to cooperate in order to advance. However, our model also suggests that proselfs need time to learn to cooperate, and they learn faster in the presence of prosocials (see Appendix A). When meritocratic matching functions optimally, prosocials and proselfs are quickly segregated. One may question whether it is fair to condemn proselfs to collectivities in which many – if not all – defect. The consequence is that meritocratic matching benefits prosocials and harms proselfs. The question is whether such a cleavage between cooperative and defecting groups is collectively optimal. For example, in higher education, our model suggests that it is best for cooperative students to join forces with other cooperative types, leaving non-cooperative students astray. But this risks writing off groups of proselfs who are not, by definition, incorrigible defectors. Thus, meritocratic matching can also have negative externalities for non-cooperative types who initially fell off the cooperative wagon, exacerbating societal inequality.

Also, future work may want to inspect conditions under which the network works as an exclusionary mechanism. For instance, our model shows how random networks operate as an exclusionary mechanism in dividing prosocials and proselfs only under certain conditions. Especially

5 A potential extension of the model could be to introduce the option of exiting the game. Some works show that not having to be part of a 2-person or n-person PD may increase the welfare of the collective because more cooperative oriented agents can quickly interact with like-minded others, avoiding more defecting types (Orbell & Dawes, 1993). The consequences for the network are discussed later by adressing the implications of dynamic instead of static networks. Intuitively, a potential consequence of having an exit option in the group context may be that prosocials find each other more readily because they quickly leave a group where some defection occurs, leaving defectors to squander around. Yet, in incomplete information settings (rules 2 and 3), prosocials who quickly exit a group may still be spoiled by the bad barrel. Still, for spoiled prosocials, not playing the n-person PD long enough may be a second chance if they are not considered spoiled anymore.

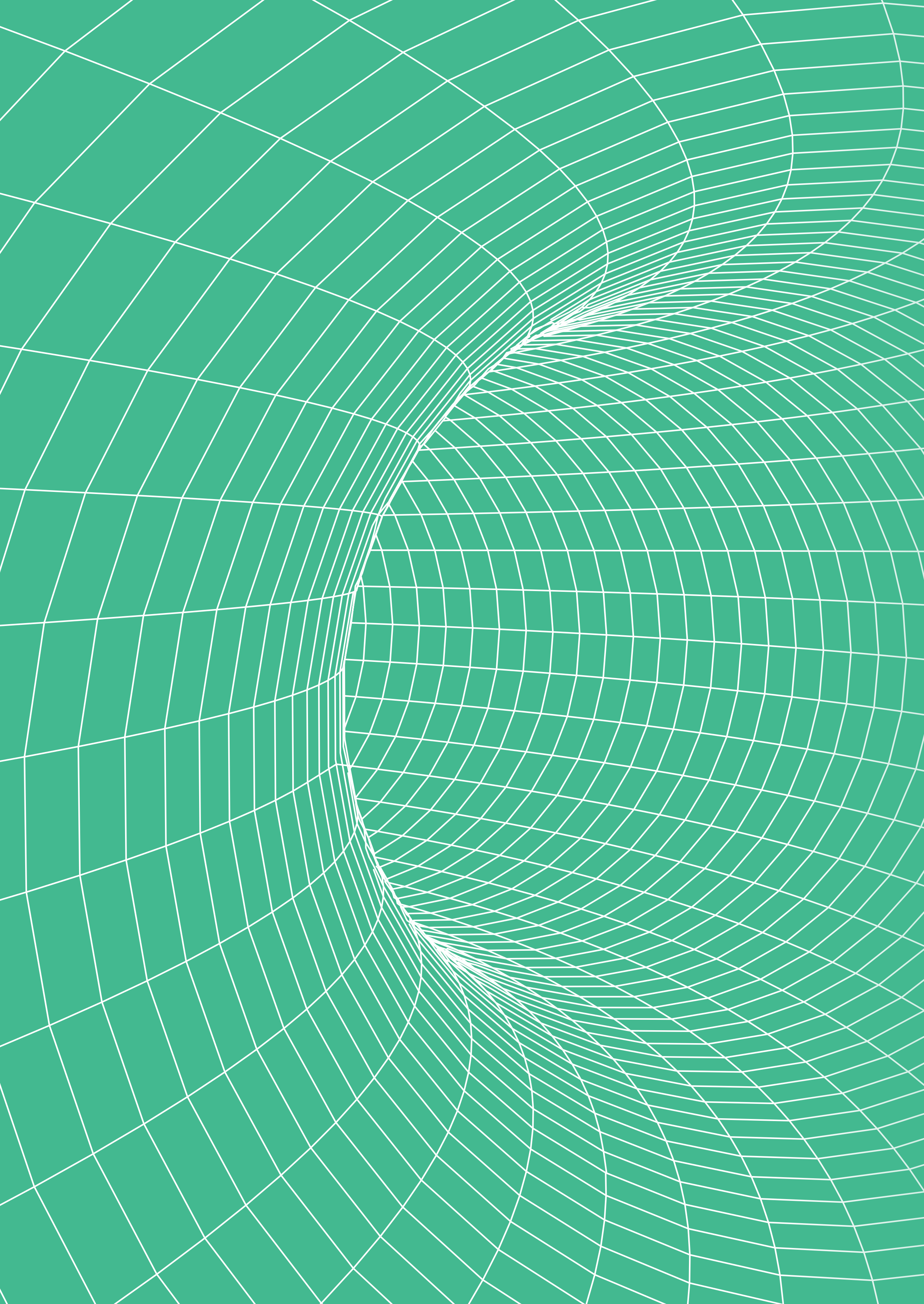
in the early stages of interactions in random networks, proselfs are relatively less attractive than prosocials. Overall, this does not translate to a radical change in cooperation rates in the group context compared to the condition of homophilous networks. A follow-up study may entail exploring network conditions under which the exclusionary mechanism in cooperative relations in homophilous networks also increases the exclusion of non-cooperators in the group context. A further intriguing possibility could be that heterophilous networks – networks in which prosocials are preferentially connected to proselfs – can lead to better chances for proselfs to escape from low-cooperation groups because they can learn more effectively to cooperate in the relational context. Other sources of uncertainty may allow the discernment of proselfs and prosocials more readily in dyadic interactions. One can think of uncertainty concerning with whom one interacts, including end-game behavior via a probability of ending an interaction, or in-group favoritism based on individual dimensions other than cooperation. These potential additional features - when the likelihood of interacting in the future is uncertain or when we expect those interactions to occur in the distant future – may foster differentiating between proselfs and prosocials.

Finally, although we have already introduced some potential model extensions in Appendix A9, our model makes assumptions about how individuals process and respond to the information obtained from group and dyadic interactions. As a first step toward testing the practical implications of our model, it is important to test these behavioral assumptions in laboratory experiments or empirical settings. We envision settings where participants are embedded in group and network contexts and use information concerning merits from one context to inform others in another. Furthermore, the homophily solution in a world where meritocratic matching is based on imperfect information does not particularly exacerbate the problem for proselfs. A reason why proselfs do not experience a backfire effect of homophily may be the static nature of the network in our model. A dynamic network in which agents preferentially form and dissolve ties with (dis)similar cooperative others may eventually result in a cooperative cluster in which prosocials reside, while proselfs are condemned to interact with similar others in a cluster in which defection prevails. Then homophily may be detrimental for the chances of proselfs to cooperate in the group context. Our model already incorporates interaction dynamics (via parameter r), but dynamic networks may introduce another mechanism that separates defectors from cooperators.

The model indicates that dyadic interactions provide noisy signals of which type the other is, especially when all agents tend to cooperate at some point due to repeated interactions. Dynamic networks may also be a feature that affects the distinction between prosocials and proselfs. If prosocials (proselfs) assume that cooperators (defectors) are most likely prosocial (proself), it may alter who is more likely to interact with each other in the network. Since cooperators are likely to sever ties to defectors, a potential outcome is that prosocials tend to cluster because they seek cooperating partners. A consequence is that proselfs – most likely defectors – tend to be isolated or left with no option to interact with similar defecting others. Prosocials and proselfs may be more distinguishable via their shown behavior and interaction partners. Yet, we include stochasticity in the model, and proselfs/defectors always have a chance to cooperate by accident (and thus be part of the pool of cooperators if they continue to cooperate). Another extension worth exploring in future research is a more complex decision-making model for cooperation, incorporating more forward-looking considerations, such as expectations, strategic behavior, and future payoff-driven behavior. Such an extension would make the decision-making model more complex but combines backward (learning) and forward-looking (strategic anticipation) components. Relatedly, we could implement prosocials being disappointed more quickly by proselfs' defection or that defectors avoid groups in which they expect to be punished for their defection. This requires a more complex cognitive assumption we have not currently implemented in the model but may be interesting to explore in future work.

In summary, we showed that meritocratic matching systems in which merit is assessed based on group-level outcomes suffer from what we termed the "bad barrels" problem. Persons with cooperative intentions (the "good apples") end up in uncooperative groups (the "bad barrels"). They cannot change the nature of the group single-handedly and are forced to behave more uncooperatively to avoid exploitation. The good apples are thus spoiled by the bad barrels in which they find themselves. Matching systems that rely on group-level information cannot identify these spoiled good apples, resulting in collectively inefficient outcomes.

As a potential solution, information from informal social networks can be used to improve the functioning of meritocratic matching systems. Informal social networks, particularly when these networks show homophily on traits that relate to cooperativeness, allow individuals to show their merit without being constrained by group-level interdependence. Imagine again the student context discussed earlier. At the start of an academic year, students are randomly grouped to work together in project teams. The course ends at some point, and all groups receive a collective grade. A students' true value as a potential contributor in future project teams may not be reflected by the group grade. Still, social relations with similar others who also generally invest a lot of time and effort into their studies is a way out for students with more to offer. Our findings are in sync with the five rules for cooperation to arise proposed by Nowak (2006); that is, we show that reciprocity within groups in the long haul (Appendix A2), (in)direct reciprocity in network interactions, and network clustering via homophily foster cooperation. However, our work also uncovered a potential downside of homophily: segregation of proselfs limits their possibility to learn cooperative behavior over time in interactions with prosocials. This raises an intriguing possibility for future work: Identifying an optimal degree of meritocratic matching that balances the benefits for prosocials with the benefits for the overall population.



Chapter 4

Studying the multiplexity of social life[1]

When are you going to get it into your head?
We are in this together!

—A quote by Emma Watson as Hermione Granger
in the movie *Harry Potter and the Order of the Phoenix* (2007)

[1] This chapter is based on joint work with **Jasperina Brouwer**, **Dieko Bakker**, and **Andreas Flache**, which is currently submitted to a peer-reviewed journal under the working title *"The role of personality traits in the formation of friendship and preference-for-collaboration networks"*.

ABSTRACT

Chapter 4 aims to determine whether real-world networks exhibit similar tendencies to those predicted by the theoretical model in Chapter 3. In particular, Chapter 4 focuses on the selection of network partners and the potential exclusionary dynamics that arise from such choices. Exclusion is addressed by studying whom students form cooperation relations with. I investigate how individual and social characteristics, such as personality traits and friendships, affect the formation of cooperative relationships within networks: Do friends form cooperative relationships, or do similar individuals in terms of personality traits preferentially form cooperation relationships? This inquiry is motivated by the understanding that the failure to establish a cooperative relationship can result in exclusion from potential benefits that could arise from such a partnership. Using a sample of students, I analyze patterns of rejection and preference in forming cooperative relationships and provides insights into the factors that shape such decisions.

4.1 INTRODUCTION

The first year in higher education is an exciting, uncertain, and challenging time for students in many respects (Christie et al., 2004; Trautwein & Bosse, 2017; van der Zanden et al., 2018; Wilcox et al., 2005). They need to adjust to a new social and academic environment, pursuing academic success in a competitive setting as well as striving to integrate into a new community. Establishing friendships and starting collaborations ease the transition from secondary to higher education. First, friendships are among the most important sources of support, help, or peer feedback to achieve academic success (Brouwer et al., 2018, 2022; Stadtfeld et al., 2019). Second, developing a network of reliable collaboration partners is important since first-year students in higher education face multiple challenges necessitating joint efforts (Blumenfeld et al., 1996; Brouwer et al., 2016; Lin, 1999), such as studying together and sharing study material. For activities like these, friendships and collaboration relationships are, in such instances, crucial to realize valuable study outcomes.

The transition from high school to the university often occurs during emerging adulthood (Arnett, 2007). This phase in life is simultaneously characterized by various life transitions, such as starting at university, finding a new job, and exploring romantic relationships. Not only building a new social network is important but also the development of personality within the social context (Asendorpf & Wilpers, 1998; Parker et al., 2012). Deventer et al. (2019) show that especially personality – and changes therein – are crucial for evolving network relationships (also, see Asendorpf & Wilpers, 1998). For example, insecurity regarding friendships is related to higher levels of neuroticism and lower levels of extraversion. Van Zalk et al. (2020) indicate that friends reinforce each other's extraversion, leading to higher levels of extraversion on both ends of a friendship tie. Also, Mueller et al. (2019) show that especially people with neuroticism – i.e., more anxious or quicker aroused – benefit the most from interacting with friends, fostering their happiness. Yet, neurotic persons generally find it more difficult to form and maintain relationships over time (Robins et al., 2002).

More importantly, personality traits affect the emergence of network relations. For establishing peer relationships, it matters what type of person the other student is – i.e., their personality. Prior research shows that Five-Factor Model (FFM) personality traits impact friendship formation, work ties, help-seeking, and forming romantic relationships (Fang et al., 2015; Harris & Vazire, 2016; Selfhout et al., 2010; Shchebetenko, 2019). Succinctly, FFM traits comprise agreeableness (altruism), extraversion (sociable), neuroticism (emotional instability), openness to new experiences (creativity), and conscientiousness (self-discipline) (DeYoung, 2015; McCrae & John, 1992). For example, findings show that extraversion, agreeableness, and openness are important for forming friendships (Selfhout et al., 2010; Shchebetenko, 2019; Zhu et al., 2013). Conscientiousness appears in some studies to play an important role in forming non-friendships such as work relationships (Fang et al., 2015) but in others as a key indicator of qualitatively valuable friendships (Jensen-Campbell & Malcolm, 2007) and friendship support (Tackman et al., 2017).

However, previous research pays little attention to the possibility that personality is linked in a different way to different types of social relations. Here, friendship relations and academic collaboration are of particular interest, which forms and change simultaneously among first-year students. Moreover, these different types of social relations cannot be seen as independent from each other. Students have multiplex relations, i.e., more than one type of network relationship with each other at the same time (Kadushin, 2012; Kivelä et al., 2014; Stanley & Faust, 1994). Our research addresses the overlooked issue of the impact of personality in multiplex networks: Students simultaneously form friendships and collaboration ties. It is, furthermore, common that friends prefer to collaborate or that preferred collaborators become friends (Brouwer et al., 2018).

Specifically, Chapter 4 focuses on the impact of the five basic dimensions of personality on two types of relationships: friendship and preferences-for-collaboration (PFC) networks. Using multiplex networks allows for a more comprehensive examination of the influence of personality on network formation (extending Chapters 2 and 3). For instance, being sociable (extraversion) or creative (openness to new experiences) may be important for friendship networks but may not be as defining for selecting a collaboration partner. On the other hand, being conscientious (conscientiousness) or altruistic (agreeableness) may be more important for collaboration relationships than friendships. In addition, how both networks influence each other is of interest to us. For example, this multiplexity can be problematic for students who are not well-connected in one network. Students with few friends may miss out on information and resources flowing through collaboration relations, unlike students with many friends. This process may generate *stacked clustering*: If multiplex relations tend to align over time, then networks may increasingly cluster along a shared dimension (Hooijsma et al., 2020). A theoretical example of stacked clustering in groups and networks can be found in Chapter 3. This study focuses on the different effects of personality on friendship and collaboration networks and the interdependence between these networks concerning the effects of personality.

4.1.1 Advancement of Chapter 4

Here, we assess the impact of multiplex ties on network formation across friendship and PFC networks. As such, we question whether, and if so, in what way relations from one network affect relations in another network. Does stacked clustering – friends becoming preferred collaborators and vice versa – occur? Are popular students in one network also popular in another? Failure to account for spillover effects between networks also risks overestimating the role of each network studied in isolation. Moreover, the alignment of network relations in multiple networks may lead to the formation of faultlines in which groups fall apart into subgroups that are homogeneous across multiple network dimensions (Mäs et al., 2013).

To study this, we utilize stochastic actor-oriented models (SAOMs), which allow us to study (*i*) network changes over time, (*ii*) the role of personality traits in the formation of network relations, and (*iii*) the multiplexity of friendship and PFC networks (Brouwer & de Matos Fernandes, 2023; Snijders, 2017; Snijders et al., 2010, 2013; Steglich et al., 2010). A SAOM allows for assessing the co-evolution of personality and network relations as well as the interdependence between network relations by using simulation methods for finding model specifications that yield the best fit between data and theoretical assumptions specified by the researcher. In this way, SAOM solves the problem of the interdependence of observations in evolving social networks that renders conventional statistical methods unsuitable, which rely on the independence of observations. Chapter 4, using SAOMs, will help us understand the impact of personality traits – controlling for grades and gender – in forming friendship and PFC networks among 95 first-year students in higher education.

In the remainder of this chapter, we first dive deeper into the relationship between personality and networks. Then we discuss the data and our analytical strategy. Finally, we provide an overview of the descriptive as well as the SAOM findings and end with a discussion of further implications of our study.

4.2 PERSONALITY AND THE FORMATION OF NETWORK RELATIONS

Personality plays a crucial role in the formation of relationships within networks because it acts as a key determinant of an individual's behavior, network position, and relational choices (Fang et al., 2015; Harris & Vazire, 2016; Selfhout et al., 2010; Zhu et al., 2013). Research using the same SAOM methodology and FFM measure applied here has shown that individuals with similar personalities in terms of agreeableness, extraversion, and openness tend to preferentially form friendships with similar personalities (Selfhout et al., 2010). A plausible explanation is that individuals with similar personalities are likelier to share common values, goals, and interests, which are the foundation for friendship formation.

However, the impact of personality traits on network formation may depend on the content of the network itself. Making friends is essential for adapting to a social environment. For example, friendships are one of the most important sources of support and help to achieve academic success (Brouwer et al., 2018, 2022; Stadtfeld et al., 2019). But developing a network of reliable collaboration partners is also important since students in higher education face multiple challenges necessitating joint efforts (Brouwer et al., 2016; Lin, 1999). Students study together and share study material. For activities like these, collaboration relations are crucial to advance in higher education (Stadtfeld et al., 2019). Whether personality traits differently affect network formation in multiplex networks remains unanswered.

On the one hand, some traits are particularly dominant in friendship networks. Namely, research shows that more extroverted and agreeable individuals tend to have larger friendship networks, potentially due to traits – respectively, being sociable and altruistic – that are perceived as attractive to form a relationship with (Zhu et al., 2013). A potential reason for this finding is that such types are more likely to give timely social support if needed (Barańczuk, 2019). The research of Fang and colleagues (2015) further notes that individuals high on openness may have fewer friends but a more diverse friendship network, serving as a bridge in-between distinct clusters of friends. As an illustration, Zhu et al. (2013) find that students with higher scores on openness and extraversion have more diverse contacts (i.e., not with similar first-year students but with second or third-year students) than students low on openness. Finally, students similar in agreeableness, extraversion, and openness seek each other as friends over time (Selfhout et al., 2010).

On the other hand, certain traits – not affecting friendship formation – are also important in different types of networks. Interestingly, conscientious individuals tend to be more self-organized and perform better at work, potentially because their socially recognized conscientiousness allows them to move into a more central role in work-related networks in which more information and resources are available (Fang et al., 2015). Meta-analytic research shows that conscientious individuals are motivated to work hard and feel responsible for pursuing shared goals (Wilmot & Ones, 2019). These qualities are generally attractive in a more work-related environment where people work together to produce valuable outcomes, such as reports.

The problem is that there is hardly any research – to our knowledge – that assesses whether personality traits affect different network relations students can have in a different way. Earlier research indicated that extraversion and agreeableness affect friendship formation, whereas conscientiousness appears to be the most important trait in forming a more work-related network. Prior research notes that neuroticism tends to affect network formation in the sense that such individuals are avoided. All this suggests that there are different ways in which traits can affect network formation, which we describe next.

We consider three potential ways personality traits affect network formation: activity, popularity, and homophily. First, activity stresses that some students are more active in initiating network relations than others and that this activity depends on individual characteristics such as personality traits (Snijders & Lomi, 2019). For example, extroverted people are more active in their friendship network (Selfhout et al., 2010). Second, popularity captures the opposite of activity: i.e., people with a high value on a trait are more often selected as network partners by others (Snijders & Lomi, 2019). These mechanisms can explain why some students are more popular than others. For example, Fang et al. (2015) note that neurotic individuals are less popular for forming friendship relations with them or to be asked for help in case of need. Also, people high on agreeableness are more popular as friends (Selfhout et al., 2010). Fang et al. (2015) find that conscientious individuals are frequently asked for work-related advice and more information. Third, homophily assumes that students with similar attributes (i.e., similar scores on FFM traits) are attractive as network partners because "they are like me" (McPherson et al., 2001; see Chapter 3). For example, it has been found that students high on openness and agreeableness tend to preferentially form ties with similarly creative and altruistic others to combine their 'creative and altruistic forces,' respectively (Selfhout et al., 2010). In this chapter, we assess how FFM traits are related to differences between students in their activity, popularity, and homophily in both networks.

One personality trait in particular that appears a plausible candidate for increasing students' popularity as friends and collaboration partners is agreeableness. Note that we interpret agreeableness as an operationalization of altruism (see Chapter 2 for a different measure). More altruistic individuals are likely to provide more social support in friendship relations and may be more active in befriending others. However, interestingly, recent empirical work does not identify altruism homophily as an important driver of network evolution among friends (Girard et al., 2015; Melamed et al., 2020; Simpson et al., 2014). For example, experiments with university students indicated that university affiliations are more important than altruism for the formation of homophilous relations in a stylized network (Melamed et al., 2020). A potential reason why altruism homophily is not found in fieldwork may be the impact of previously non-measured traits such as extraversion or openness, or that the trait is important in a different context (i.e., the PFC network). For example, prior research shows that extraversion and openness to new experiences are important reasons for forming network relations. This leads us to inspect the link between personality traits, altruism (as operationalized via agreeableness), and network formation in the present study.

Studying homophily requires using the SAOM tool for accounting for network dynamics. For instance, observing homophily in a static cross-sectional network does not allow us to conclude that homophily is a stable network feature, nor does it reveal the process through which trait homophily came about. Namely, homophily captures whether connected actors are similar on certain dimensions – e.g., attitudes or behavior – than actors who are not so well connected in the network. Yet, homophily as a structural pattern in the network can arise from both selection and influence (to put it in SAOM terms). The selection mechanism stresses that similar actors, for example, select each other as friends. At the same time, the influence mechanism posits that friends tend to become alike in terms of behavior. We build on SAOMs and networks measured on multiple occasions to disentangle the two processes. We, therefore, stress the importance of studying dynamic networks. Our longitudinal network data and method employed are suitable for studying the stability of homophily and the processes by which it develops.

Finally, we control for well-known attributes that have been found to foster network formation among students: gender and grades. For instance, previous work in higher education and the network literature consistently finds gender (Brouwer et al., 2022; Kossinets & Watts, 2009; Weber et al., 2020) and grades (Brouwer et al., 2018; Lomi et al., 2011) homophily. Following the attraction hypothesis (Byrne, 1971), interacting with a same-gender other tends to be more comfortable – as well as leading to more positive attitudes about each other – than with a dissimilar other. Also, grades serve as a signal to others that one is capable to thrive in higher education. High-grade students would prefer to form network relations with similar others, forcing low-grade students to do the same. Also, we may expect that higher-achievers are more popular than lower achievers (Brouwer et al., 2022).

4.3 DATA AND MEASURES

We analyze data from 95 Dutch first-year students. The sample comprises 58 female (61%) and 37 male (39%) students with a mean age of 19.5 years old (SD = 1.6). Participation in this study was voluntary. Students were informed about the aim of the study before giving informed consent to use their data for research purposes. Ethical approval was obtained from the ethics committee of the degree program. Students answered a 20 to 30-minute computer-based questionnaire three times within an academic year. Friendship and preference-for-collaboration (PFC) network data were collected at the end of the first semester ($t = 1$) and the end of the second semester ($t = 2$). Five-Factor Model (FFM) traits are assumed to be stable (DeYoung, 2015) and were only measured at t = 0 (start of the academic year). This data was previously used to assess in what way learning communities and network relations promote academic success (Brouwer et al., 2018, 2022), and whether attitudes affect friendship selection (Brouwer & Engels, 2021). Both studies primarily focused on grades as an attribute of interest. An overview of the variables is provided in Table 4.1.

Table 4.1: An overview of constructs, scales, and waves (t = 0, 1, and 2) in this study. N = 95 students and year of data collection = 2013-2014.

Construct	**Scale**	**Wave *(t)***
Network measures		
1. Friendships relationships	1 = tie; 0 = no tie	1 & 2
2. Preference-for-collaboration (PFC) ties	1 = tie; 0 = no tie	1 & 2
Individual attributes		
3. FFM traits: neuroticism, conscientiousness, extraversion, openness, and agreeableness	1 = very inapplicable to 5 = very applicable	0
Control variables		
4. Grades	1 = lowest to 9 = highest	1
5. Gender	1 = female; 0 = male	0

4.3.1 Friendship network

Students nominated their fellow students as friends on a scale from 1 ("best friends") to 6 ("I don't know who this is"). All students in this cohort were embedded in learning communities – students divided into small groups in the first year who semi-regularly meet to share their, for example, academic experiences. Earlier studies reveal that embeddedness in learning communities does not play a defining role in friendship selection (Brouwer et al., 2018).[2] Nominations followed a two-step procedure in which they were first provided with a complete list of names of members of their learning community. Students were asked to evaluate their friendship rela-

2 For more information on learning communities, we refer to Brouwer et al. (2018).

tion with each learning community member. Subsequently, students could enter the names of other students in the study program on their own accord. This was assisted by a free recall name generator, meaning that complete names popped up even if students only typed a part of a name. To analyze the friendship network with stochastic actor-oriented models, we dichotomize the presence of a friendship relation from the 5-point scale as follows: 1 = "best friends," 2 = "friend," and 3 = "friendly relationships" are coded as 1 (a friendship tie). Options 4 = "neutral, not much in common," 5 = "only known from face or name," and 6 = "I don't know who this is" are coded as 0 (no friendship tie).

Next, we measure the stability of friendship nominations between $t = 1$ and $t = 2$ with the Jaccard similarity index (Snijders et al., 2010). There are, in total, 581 friendship ties at $t = 1$ and 474 at $t = 2$. There are a total of 8930 possible ties. Friendship networks should contain a stable part, owing to the fact that networks generally change slowly. The Jaccard index measures changes in ties between two waves. Too much instability negatively impacts the reliability of the statistical analysis. The network is considered to be too unstable for Jaccard index values below 0.30 (Snijders et al., 2010). The Jaccard index of 0.381 for our network shows sufficient stability in friendship nominations between both waves. Specifically, there are 473 changes in friendship tie presence when comparing $t = 1$ and $t = 2$. From $t = 1$ to $t = 2$, most ties ($n = 8166$) remained absent. 183 ties were formed, whereas 290 ties were severed over time. 291 ties remained present from $t = 1$ to $t = 2$. The density of the network decreases from 0.065 (581 / 8930) to 0.053 (474 / 8930), owing to the decrease in the number of ties. The average degree decreased from 6.12 to 4.99, meaning that students have 5 to 6 friendship ties on average.

4.3.2 Preference-for-collaboration (PFC) network

Indicating with whom one prefers to collaborate followed the same procedure as with friendship nomination. Students indicated at $t = 1$ and $t = 2$ "*I would like to collaborate with* [name]." They could rate each other on a 5-point Likert scale, ranging from 1 ("strongly disagree") to 5 ("strongly agree"), with the option of 6 ("I do not know"). Per dyad, we dichotomized the received nomination. Categories 1, 2, 3, and 6 are re-coded in the matrix as "0" and 4 and 5 as "1". Scores 4 and 5 indicate the sender considers the receiver to be a preferred collaboration partner. All other categories – either unnominated or scores 1, 2, 3, or 6 – are not identified by peers as students with whom they would like to collaborate. We dichotomized the network variable because we are (*i*) interested in differentiating between a present or absent PFC tie and (*ii*) to be able to model changes in the PFC network using SAOMs.

The PFC network is stable over time. The Jaccard similarity index of 0.375 shows enough stability in PFC nominations between $t = 1$ and $t = 2$. A total of 499 changes in PFC relations occurred over time. Out of a total of 8930 theoretically potential relations, we have 8131 PFC ties that remain absent over time. 216 PFC relations were formed from $t = 1$ to $t = 2$. 283 PFC relations were severed, and 300 remained present over time. The average degree is 6.14 and 5.43, for $t = 1$ and $t = 2$, respectively. The total number of ties decreased from 583 ($t = 1$) to 516 ($t = 2$). The density of the PFC network decreases from 0.065 (583 / 8930) to 0.058 (516 / 8930).

4.3.3 Overlap in friendships and PFC

The overlap between the friendship and PFC networks is shown in two ways: (*i*) by comparing the in-degrees and (*ii*) by visualizing the overlap in network ties. First, the correlation between the in-degree – "being popular" – in the PFC and friendship network is significantly positive: $r = 0.87$ ($p < 0.001$) and $r = 0.91$ ($p < 0.001$), for $t = 1$ and $t = 2$, respectively. Students popular in one network tend to be popular in the other. Second, network ties quantitively overlap 78 and 83 percent for $t = 1$ and $t = 2$, respectively. Figure 4.1 reports two Venn diagrams showing the overlap between friendship and PFC ties. The Venn diagrams indicate that friends prefer collaborating with friends and vice versa. One can intuitively argue that students tend to want to collaborate with their friends and be friends with those with whom they want to collaborate in higher education. Figure 4.1 shows essentially the need to account for multiplexity.

4.3.4 Personality traits

The FFM captures a taxonomy of five traits related to everyday behavior: agreeableness, extraversion, neuroticism, openness, and conscientiousness. We use the Ten-Item Personality Inventory to assess the FFM (Gosling et al., 2003). Students answered the question "*To what extent do the following statements relate to you?*" separately for each item. The following 10 items are considered: 1) "*I take time for a talk*," 2) "*I try to avoid conflicts*," 3) "*I work in a structured manner*," 4) "*I am easily enthusiastic*," 5) "*I am open to new experiences*," 6) "*I ignore adversity quickly*," 7) "*I see myself as someone who is generally trusting*," 8) "*I can handle stress well*," 9) "*I am interested in art*," and 10) "*I am self-disciplined*." Students indicated their choice on a 5-point Likert scale, ranging from 1 (very inappropriate) to 5 (very appropriate). Items corresponding to a specific trait (1, 4 = extraversion; 2, 7 = agreeableness; 3, 10 = conscientiousness; 6, 8 = neuroticism; 5, 9 = openness to experiences) were averaged to capture the latent trait accordingly.

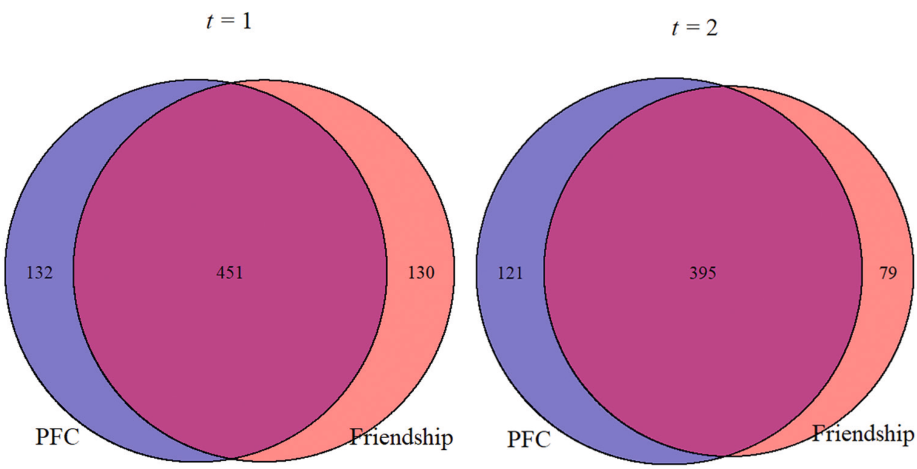


Figure 4.1: Two Venn diagrams of the count of links that overlap (purple) in the preference-for-collaboration (PFC; blue) and friendship (red) network at $t = 1$ (left) and $t = 2$ (right).

The Ten-Item Personality Inventory is not designed to bolster the internal reliability of the measure but rather to provide researchers with a tool to measure the Big-Five personality dimensions in a brief period. Even so, we report various reliability statistics per trait in Table 4.2. A reason for the relatively weak Cronbach's alpha's (α), compared to common standards of α > 0.60, can be the use of only ten items; α generally increases when more items are included. Gosling et al. (Gosling et al., 2003) have a similar issue with relatively low α's per trait. Also, inter-item correlations are indicative of the reliability of few-item scales (Eisinga et al., 2013). The inter-item correlations report relatively strong positive associations, owing to capturing the same content. The α's reported here are comparable to those in the original study by Gosling et al. (2003).

Table 4.2: Summary of the Five-Factor Model personality traits.

FFM personality trait	Mean (SD)	Inter-item correlation	Cronbach's alpha (α)
extraversion	3.84 (0.70)	0.37	0.53
agreeableness	4.14 (0.59)	0.31	0.44
conscientiousness	3.11 (0.95)	0.61	0.76
neuroticism	3.14 (0.78)	0.26	0.41
openness	3.56 (0.77)	0.18	0.26

Note. *SD* = standard deviation.

4.3.5 Control variables

As mentioned in Table 4.1, we include gender and grades as control variables. *Gender* is a dichotomous variable: male (0) and female (1). *Grades* are measured at $t = 1$ as a weighted average grade. The range of the grading system is 1 (min) to 10 (max). We averaged all received grades from the start of the academic year up until $t = 1$. Completing a course (grade ≥ 5.5) results in credit points. We weigh the average grade by the credit points obtained for courses divided by the maximum possible credit points possible and round the grade variable ($M = 6.07$, $SD = 2.02$, $min = 1$, $max = 9$).

4.4 ANALYTICAL TOOL

The interdependence of network and individual attributes is explicitly addressed in stochastic actor-oriented models (SAOMs) (Snijders, 2017; Snijders et al., 2010, 2013). Binary network data is used as input for network selection. The SAOM framework builds on agent-based computational models (Snijders, 2017; Snijders & Steglich, 2015). Namely, the SAOM assesses which theoretically postulated mechanisms can best generate a sequence of mini-steps that explains changes between observed waves. One mini-step can mean forming, dropping, or maintaining a network tie by an actor. How many mini-steps an actor can take is modeled via the rate function, while the objective function determines which mini-step the actor takes. The objective function shows the attractiveness of all new network states that agents can reach by changing a network nomination or a behavioral trait. In SAOM simulations, actors decide myopically and probabilistically which mini-step to take. Mini-steps, which result in more attractive network states, are more likely to be taken.

The SAOM simulation algorithm, combined with statistical methods for model fitting and selection, is used to assess which parameter values for the rate function and the effects constituting the objective function included in a model yield the best match with the empirically observed changes in network relations and individual traits. For example, an attribute of oneself and the (prospective) network partner may drive a tie or suppress the formation of the tie. A positive (negative) parameter for an effect generally indicates the probability of pursuing (avoid) a network situation that exhibits the corresponding feature (all other things being equal). SAOMs are estimated in *R* (R Core Team, 2021) using the package Simulation Investigation for Empirical Network Analysis (*RSiena*; Ripley et al., 2021).

4.4.1 Model specification

We include the following effects to model sources of friendship and PFC selection. The rate estimate indicates how frequently tie changes occur between $t = 1$ and $t = 2$. Following established practice in SAOM modeling (Snijders et al., 2010), seven endogenous network effects are included in the objective function of friendship and PFC selection: out-degree, in-degree activity and popularity, out-degree activity, reciprocity, transitive closure, and an interaction term of reciprocity and transitive closure. The out-degree (density) indicates the general tendency to form ties in the network ($i \rightarrow j$). In-degree activity and popularity parameters show whether popular students – thus students with high in-degree scores – send out more nominations and attract more network nominations, respectively. Similarly, out-degree activity models whether students with high out-degrees tend to send more tie nominations over time. Reciprocity indicates if there is a tendency towards reciprocal relations ($i \leftrightarrow j$). Transitivity measures the likelihood of the following network structure: If $i \rightarrow j$ and $j \rightarrow k$, then this effect calculates the probability for $i \rightarrow k$. The interaction between reciprocity and transitivity indicates how both mechanisms are related. These effects represent basic processes that occur in nearly all social networks.

Per attribute – FFM traits, grades, and gender – we add three effects in the objective function of the selection model: activity, popularity, and similarity. These effects parametrize whether relations are more likely for the individual attribute's lower, higher, or similar values. The activity effect elucidates to what degree higher values on an attribute (indicated via *) determine how active students are in sending out ties ($i^* \rightarrow j$). The popularity effect measures the opposite of activity, indicating whether students with a higher score on the attribute of interest are more likely to be nominated ($j \rightarrow i^*$). Finally, homophily is measured via the similarity effect, capturing tendencies of unrelated students (i^* and j^*) to form ties with similar others or for related students to maintain ties ($i^* \rightarrow j^*$).

We assess three cross-network effects (for more information on modeling multiplex networks with SAOMs, see Snijders et al., 2013). The first effect models a basic cross-network spillover effect: are friends more preferred as collaboration partners (and vice versa)? The second cross-network effect captures whether a reciprocated friendship relation spurs a PFC nomination and whether a reciprocated PFC relation spills over to nominating the other as a friend. This is to capture the expectation that reciprocated relations indicate mutual and thus more stable selections as (respectively) mutual friends or collaboration partners. The third cross-network is the cross-network popularity effect. This effect captures whether two actors are more likely to connect in one network if they are both nominated by the same alter in another network. To be clear, the effect captures whether both i and j being nominated as a friend by k makes i and j more likely to form a PFC relationship. And, whether i and j are nominated as a PFC partner by k makes i and j more likely to form a friendship relationship.

After estimating the SAOMs, we extract each effect's relative importance (RI) in the objective function of network selection using the *sienaRI* function (Indlekofer & Brandes, 2013). The RI shows in percentages in what way changes in the network can be attributed to an effect. Combining all percentages in either friendship or PFC selection cumulatively leads to 100 percent.

We run separate SAOMs for friendship and PFC selection and insert cross-network effects to investigate multiplexity. The seven structural network features, twenty-one attribute effects, and three cross-network effects are included in both SAOMs. We checked the convergence of the *RSiena* algorithm. This is automatically provided in the SAOM output by reporting *t*-ratios for each effect as well as overall model convergence. Convergence statistics for both selection models adhered to the standard criteria for convergence (for more information, see Ripley et al., 2021): all effects have a *t*-ratio below 0.10, and model convergence ratios are below 0.25.

4.5 FINDINGS

4.5.1 Descriptive information

Table 4.3 provides correlations between all continuous variables in our data. First, the correlation between grades and conscientiousness is significant and positive (r = 0.29, p = 0.009). Conscientiousness students tend to score higher grades, while low graders tend to be less conscientious.

Second, we find a significantly positive association between extraversion and openness to new experiences (r = 0.32, p = 0.003). Students higher on extraversion also tend to score higher on openness, meaning that exhibiting enthusiasm and availability for a talk (extraversion) is associated with being open to new, creative experiences (openness). Most FFM personality traits are not significantly correlated with each other, owing to their measurements being constructed for capturing distinct personality traits.

Furthermore, we employed some ANOVAs to inspect the relationship between gender and grades as well as with FFM traits. We find in our sample that females have a significantly higher mean score on conscientiousness than males, $F_{1, 81} = 4.26$, $p = 0.04$. There were no statistically significant gender differences on agreeableness ($F_{1, 81} = 0.47$, $p = 0.50$), neuroticism ($F_{1, 81} = 0.16$, $p = 0.69$), openness ($F_{1, 81} = 0.49$, $p = 0.49$), and extraversion ($F_{1, 81} = 3.34$, $p = 0.07$). Furthermore, there are no significant gender differences in grades, per the ANOVA analysis result of $F_{1, 86} = 1.80$, $p = 0.18$.

Table 4.3: Pearson correlation matrix of Five-Factor Model (FFM) personality traits and grades.

Measure	1	2	3	4	5	6
Grades	—					
Extraversion	0.01	—				
Agreeableness	−0.02	0.07	—			
Conscientiousness	0.29**	0.11	−0.08	—		
Neuroticism	0.05	0.17	−0.07	0.11	—	
Openness	0.02	0.32**	0.01	0.13	0.16	—

* $p < 0.05$; ** $p < 0.01$.

4.5.2 Stochastic actor-oriented model findings

We utilize SAOMs to inspect micro-level drivers of network selection in friendship and PFC networks. Only the final models are discussed because there are only minor differences among the estimated models (Appendix B shows SAOMs with a stepwise inclusion of studied estimates[3]). We start in this section by discussing friendship network selection effects, followed by inspecting the PFC network. We then discuss cross-network features.

4.5.2.1 Evolution of Friendship Selection

The friendship selection SAOM is reported in Table 4.4. The results indicate that most FFM trait effects are not key determinants for friendship selection in our sample. An exception is the trait openness to new experiences. Students higher on openness are significantly more active (estimate = 0.47, *SE* = 0.11) and popular (estimate = 0.38, *SE* = 0.11) than less open students. Relative importance percentages show that FFM traits are among the weakest individual attribute determinants of friendship selection in our sample relative to other SAOM effects.

We found significant structural network and control variable effects on friendship selection. The negative out-degree parameters – the more negative the outdegree, the sparser the network is – indicate that there is a general tendency not to nominate another student as a friend unless

3 The R-script to analyze the data and Supplementary Information are also freely available at the Open Science Framework, via https://doi.org/10.17605/OSF.IO/GSA4E.

the nomination brings with it desirable consequences, such as reciprocating an incoming tie or befriending a desired cooperation partner (estimate = –2.99, *SE* = 0.26). Furthermore, popular students – i.e., high in-degree scores – are not more active in sending out friendship nominations (estimate = –0.26, *SE* = 0.06). The positive reciprocity estimate of 2.77 (*SE* = 0.23) indicates that students tend to reciprocate friendship ties. In addition, we observe that students tend to initiate ties to close transitive triplets in their friendship network (estimate = 0.49, *SE* = 0.07). We further find that the relative importance of structural network features is the largest: A combined total of approximately 59%. Furthermore, we find that students prefer same-gender (estimate = 0.72, *SE* = 0.13) and similarly-achieving (estimate = 1.99, *SE* = 0.65) students as friends. Students with higher grades are more active in sending out friendship nominations (estimate = 0.27, *SE* = 0.14). Also, females tend to be significantly less active (estimate = –0.49, *SE* = 0.14) and less popular (estimate = –0.34, *SE* = 0.14) than males. Gender and grade effects have a combined relative importance of nearly 23%.

Table 4.4: SAOM findings for friendship and PFC selection.

	Friendship selection		**PFC selection**	
	estimate (SE)	**RI**	**estimate (SE)**	**RI**
Rate parameter				
1. Friendship rate $t = 1 \rightarrow t = 2$	10.70** (1.53)		14.86** (1.70)	
Structural network effects				
2. Out-degree (density)	–2.99** (0.26)	17.4%	–2.60** (0.25)	16.6%
3. Out-degree activity	0.05** (0.02)	4.0%	0.03* (0.01)	3.2%
4. In-degree popularity	–0.05 (0.03)	4.5%	–0.05* (0.02)	5.6%
5. In-degree activity	–0.26** (0.06)	9.5%	–0.18** (0.05)	8.3%
6. Reciprocity	2.77** (0.23)	9.5%	2.48** (0.21)	9.5%
7. Transitive triplets	0.49** (0.07)	9.4%	0.45** (0.06)	10.3%
8. Transitive reciprocated triplets	–0.31** (0.09)	3.8%	–0.33** (0.07)	5.3%
Cross-network effects				
9. PFC relation	0.64* (0.31)	2.3%		
10. Reciprocated PFC tie	–0.40 (0.44)	0.7%		
11. PFC popularity	0.15** (0.05)	2.4%		
12. Friends			0.88** (0.19)	3.5%
13. Reciprocated friends			–0.19 (0.33)	0.2%
14. Friendship popularity			0.08 (0.04)	1.4%

[continued on next page]

Table 4.4: *[continued]*

	Friendship selection		PFC selection	
	estimate (SE)	RI	estimate (SE)	RI
Individual attribute effects				
15. Agreeableness activity	−0.26* (0.12)	1.1%	−0.24* (0.12)	1.1%
16. Agreeableness popularity	−0.19 (0.12)	1.2%	−0.14 (0.11)	1.1%
17. Extraversion activity	0.05 (0.11)	0.3%	0.09 (0.09)	0.5%
18. Extraversion popularity	−0.01 (0.10)	<0.0%	−0.11 (0.09)	1.0%
19. Conscientiousness activity	0.06 (0.07)	0.4%	0.01 (0.06)	0.1%
20. Conscientiousness popularity	−0.03 (0.07)	0.3%	0.05 (0.06)	0.7%
21. Neuroticism activity	−0.05 (0.09)	0.3%	−0.03 (0.07)	0.2%
22. Neuroticism popularity	−0.10 (0.09)	0.9%	0.04 (0.07)	0.4%
23. Openness activity	0.47* (0.11)	2.2%	0.19* (0.08)	1.3%
24. Openness popularity	0.38** (0.11)	3.2%	0.35** (0.09)	3.5%
25. Grades activity	0.27** (0.08)	4.3%	0.24** (0.06)	4.2%
26. Grades popularity	0.12 (0.07)	2.6%	0.17** (0.06)	4.0%
27. Female (ref = male) activity	−0.48** (0.14)	2.1%	−0.29** (0.13)	1.5%
28. Female (ref = male) popularity	−0.34* (0.14)	2.3%	−0.34* (0.12)	2.5%
Homophily effects				
29. Agreeableness	0.57 (0.42)	1.2%	0.08 (0.31)	0.2%
30. Extraversion	0.60 (0.45)	1.2%	0.39 (0.38)	0.9%
31. Conscientiousness	0.32 (0.37)	0.8%	0.12 (0.29)	0.3%
32. Neuroticism	0.02 (0.56)	<0.0%	0.43 (0.33)	1.0%
33. Openness	0.29 (0.31)	0.8%	−0.00 (0.26)	<0.0%
34. Grades	1.99** (0.65)	5.3%	2.22** (0.45)	6.5%
35. Gender	0.72** (0.13)	6.2%	0.54** (0.12)	5.2%

Note. *SE* = standard error; RI = relative importance; * $p < 0.05$; ** $p < 0.01$.

4.5.2.2 PFC Network Selection

The results of the evolution of the PFC network are presented in Table 4.4. We infer from the results that agreeableness homophily is not a defining feature for forming relations, even though agreeableness (defining altruism is expected to be an attractive characteristic for a potential collaboration partner. We find that openness to new experiences – open-mindedness, creativity, and willingness to embrace new things and adventures – is a dominant feature for activity and popularity in PFC nominations. The relative importance of FFM traits estimates in the PFC network comprises approximately 12%, of which nearly 5% are attributed to openness.

Furthermore, similar directional effects as in the friendship selection model are found for the structural network effects and control variables in PFC selection. This is unsurprising given the 66% overlap in friendship and PFC network relationships. For instance, we find reciprocity (estimate = 2.48, *SE* = 0.21) and transitivity (estimate = 0.45, *SE* = 0.06) in the PFC network. Again, the block of structural PFC network effects is most important for explaining PFC selection – nearly 60% - compared to other SAOM estimates. Also, students with the same gender (estimate = 0.54, *SE* = 0.12) and grade (estimate = 2.22, *SE* = 0.45) tend to seek similar others out for PFC relations.

4.5.2.3 Multiplex network dynamic

The cross-network estimates in Table 4.4 shed light on the multiplexity of network relations. Including cross-network effects explains approximately 5% of tie changes in both selection models. The results show that friends are preferred for collaboration (estimate = 0.64, *SE* = 0.31) and that students who consider each other to be preferred collaboration partners tend to become friends (estimate = 0.88, *SE* = 0.19). Surprisingly, these tendencies are not strengthened when the existing relationships are reciprocal. A reciprocated PFC relationship does not significantly affect forming a friendship relation, and a reciprocated friendship does not significantly affect tie formation in the PFC network. This could be because we have a high degree of reciprocity in both networks already, so there is possibly little additional variance explained by the cross-network reciprocity effect.

Notably, both being popular in the PFC network makes the formation of a friendship relationship more likely (estimate = 0.15, *SE* = 0.05). This effect suggests that students perceive PFC popularity as valuable in forming a mutually beneficial friendship relationship. At the same time, we see grade homophily present in both networks, and we know that the networks substantively overlap. Thus, PFC popularity may be associated with high grades, and students with high grades tend to select each other as friends and PFC partners. Grade segregation may thus be a potential source of increasing overlap in network relations. At the same time, our results do not support that popularity in the friendship network spills over to the formation of PFC relationships. This is consistent with the interpretation that high grades make an individual popular in the PFC network, while friendship popularity does not contribute to higher grades.

4.6 CONCLUDING REMARKS

Different types of network relations can facilitate the transition from secondary to higher education, and personality traits are believed to play a crucial role in this process. Our research investigates the extent to which certain personality traits are more related to network formation in one network type than the other and the extent to which relations in one network induce relations in the other. We explicitly studied multiplexity to determine whether personality traits are related in different ways to the interdependent formation of friendship and preference-for-collaboration (PFC) network relations. Drawing on the Five-Factor Model, we find that creative and open-minded students are more likely to form friendships and PFC relations and receive more network nominations. Even so, this chapter finds that all effects of personality traits are qualitatively (i.e., directional and in terms of significance) the same in the friendship and PFC networks. Hence, we do not find that traits relate differently to the distinct friendship and PFC network. A potential reason for this is that we observe that the majority (60%) of changes in network relations are driven by structural relational choices, such as choices increasing relational reciprocity or transitivity. Approximately 35% of relational choices in both networks are based on personality traits, gender, and grades.

Our multiplexity findings indicate, furthermore, that networks cannot be studied in isolation from each other. Students who are friends may also be PFC partners, and relationships within the PFC network may extend to the friendship network. Our findings reveal that highly popular preferred collaborators who are connected are more likely to form friendships, suggesting that clustering of popular preferred collaborators and friends occurs in both domains. Although including multiplexity considerations accounts for a small, yet noteworthy, proportion of tie changes, especially the Venn diagrams provided clear evidence of structural overlap between the friendship and PFC networks, underscoring the importance of understanding multiplexity in network relations. More importantly, it is noteworthy that multiplex network relations can provide well-connected students with a range of options for accessing information from multiple sources, while more isolated students may not have this advantage. This suggests that multiplexity can serve as valuable support for well-connected students in their educational endeavors.

This study is not without limitations. First, the data was gathered among students in one study program and one study year. This limits the generalizability to other programs in different fields. For example, prior work shows that economics students tend to be less altruistic than students in other fields (Marwell & Ames, 1981). Our non-economic student sample may thus be more collectively oriented, meaning that relational choices may be directed to ensure community building instead of pursuing their self-interest and pursuing higher grades (and thus making relational choices to realize higher grades). To increase relatedness and applicability to the overall student population and other programs in higher education, we suggest remedying this situation by implementing a similar approach in multiple academic programs.

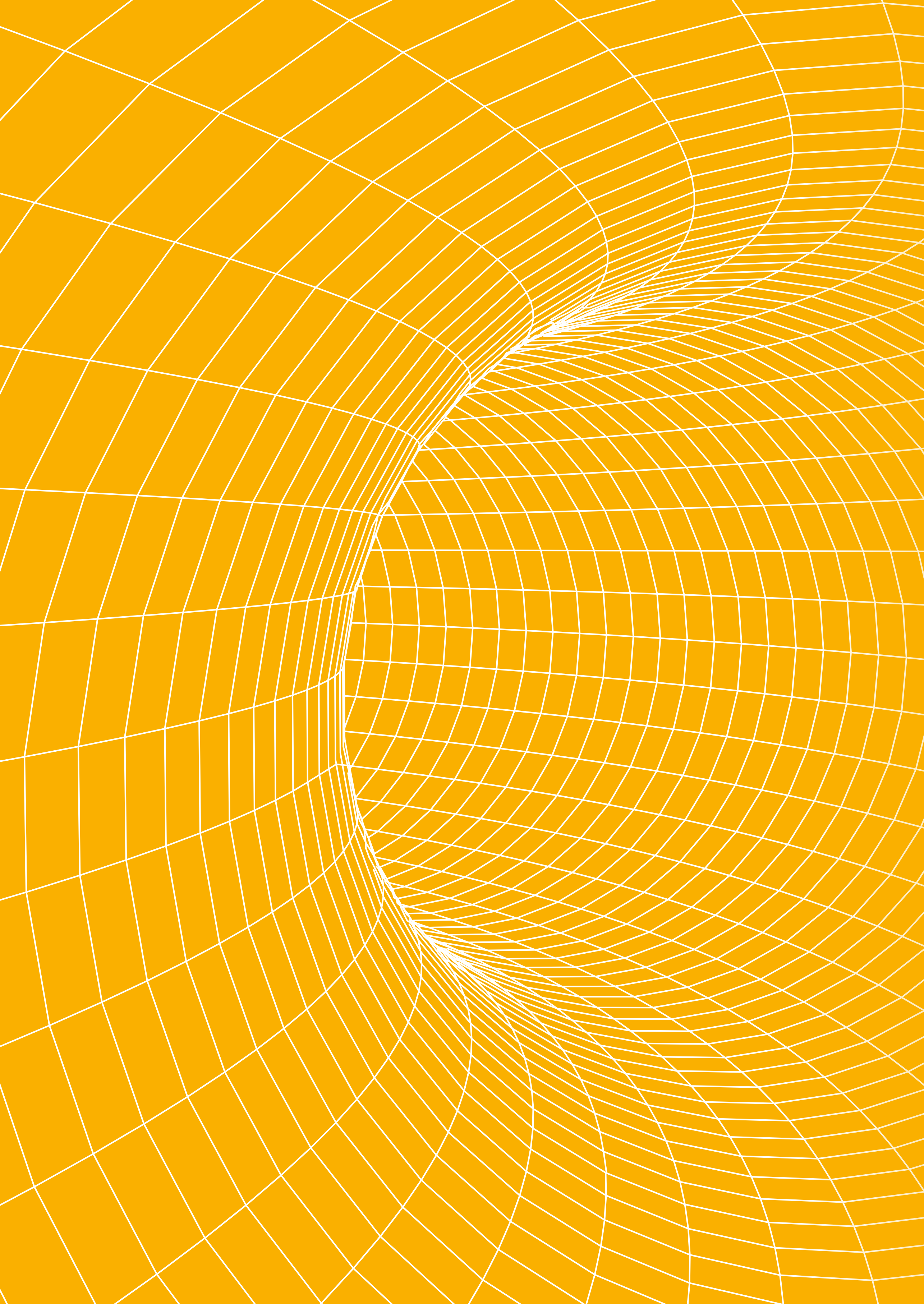
Second, our longitudinal approach incorporates only two discrete time points. Ideally, more information on network changes in an academic year would be available to assess the strength of selection in SAOMs more comprehensively. For example, more time points may explicate whether selection mechanisms found are robust in the long haul or that some features need more time to arise. To solve this issue, future work may consider more data collection time points with less time between waves.

Third, another limitation pertains to students indicating their preference to work with someone and not the extent to which they collaborated. Other social factors may come into play why a student prefers to collaborate with one over another. For example, promising a beer after successfully working together may be a reason to prefer collaborating with someone, whereas some may be more socially expressive in their collaborative qualities than others. Another potential source of information is whether collaborating in study teams is indicative of forming network relations and vice versa (Chapter 3; de Matos Fernandes et al., 2022b; Hoffman et al., 2023). In this chapter, we already advance by including multiple individual attributes and controlling for friendships spilling over to the PFC network.

Yet, the first-year students in the current sample may not have had enough experience with others' collaborative behavior to be able to judge their value as potential collaborators. Prior research on trustworthiness as input for homophily shows, for example, that information on others' behavior is essential to preferentially form friendship ties with others (Winter & Kataria, 2020). We believe that future work needs to consider incorporating behavioral or actual indicators of collaboration (think of combining behavioral indices with network experiments [Melamed et al., 2020] or a personality traits measure specifically designed to capture cooperative considerations [Chapter 2; de Matos Fernandes et al., 2022a]). Accounting for the multiplexity of social life already provides information on the multiple types of relations students have and with whom. Yet, SAOMs allow including collaborative behavior as an individual attribute.

A fruitful future endeavor would be using the micro-level effects for selection as input for SAOM simulations (Kretschmer & Leszczensky, 2021; Stadtfeld, 2018). An empirically-calibrated SAOM allows researchers to test how the strength of, for example, the homophily estimate of an attribute of interest affects the friendship or PFC network structure. This is what I do in Chapter 5. Attesting to previous work on multi-dimensional homophily examining different dimensions – e.g., gender, behavior, or attitudes – in which individuals can be similar (Block & Grund, 2014; Hooijsma et al., 2020), we also stress that empirically-calibrated models can be used to further our understanding of the interrelatedness of personality traits, gender, and grades on the micro-level and segregation on the macro-level (Stadtfeld, 2018).

Finally, stochastic actor-oriented models provide a powerful approach to disentangling underlying reasons for selection in multiplex networks. Our analyses and findings demonstrate the importance of a longitudinal approach and reveal that simultaneously assessing the evolution of friendship relations and PFC ties contribute to further our understanding of how personality traits and endogenous network features, such as reciprocity, shape multiplex networks. Higher education is a context suited to study this issue because students with different traits face challenges daily in which social and academically-related efforts are required. But the relevance of friendship and PFC networks reaches far beyond the context of higher education. In many organizational, collective, and societal settings, individuals are confronted with the need to navigate many (in)formal relations on different levels within the same group of people, for example, in a work organization in which colleagues have both social and professional interpersonal relations. Chapter 4 points to the importance of accounting for multiplex network relations in such settings.



Chapter 5

Cooperation in an artificial world[1]

The consequences of our actions are always so complicated, so diverse, that predicting the future is a very difficult business indeed.

—A quote by Albus Dumbledore in J. K. Rowlings's *Harry Potter and the Prisoner of Azkaban* (2004)

[1] This chapter is based on joint work with **Andreas Flache** and **Dieko Bakker** which is currently submitted to a peer-reviewed journal under the working title "*Injecting realism in simulation models: Do selection and social influence jointly promote cooperation?*".

ABSTRACT

Chapter 5 builds upon the research established in Chapter 4 by examining the extent to which social influence can interfere with exclusionary mechanisms. Social influence is a complex process in which individuals are influenced by their peers to behave in a particular way. I investigate how social influence can potentially affect the effectiveness of exclusionary mechanisms in promoting cooperation. I employ an empirically-calibrated model and use network data from a sample of 95 students to simulate the interplay between exclusion and social influence, taking into account other empirically established behavioral and social mechanisms. The results provide insights into the effectiveness of exclusion in promoting cooperation and how social influence affects exclusion. Specifically, in Chapter 5, I examine how actors who are already socially influenced to cooperate may not require exclusion, while those who are socially influenced to maintain their defection may fail to deviate from their behavior. Thus, social influence interferes in some simulation conditions with the positive effects of exclusionary mechanisms on cooperation.

5.1 INTRODUCTION

Individual and collective interests clash in cooperation problems. Individuals' rational pursuit of their interests may hamper collective success, eventually to everyone's detriment. Experiments and theoretical modeling have shown that network mechanisms can be key in solving cooperation problems (Baldassarri, 2015; Rand & Nowak, 2013; Simpson & Willer, 2015; Chapter 3). Especially two network mechanisms have been identified as key for cooperation to thrive, namely partner selection and social influence (Fehl et al., 2011; Fowler & Christakis, 2010; Melamed et al., 2018; Rand et al., 2014; Schuessler, 1989; Simpson & Willer, 2015). Partner selection, also referred to as network reciprocity (Nowak, 2006), allows cooperators to select similarly cooperative others as partners, which can force defectors to change their behavior to prevent exclusion from beneficial relationships with cooperators. In addition, partner selection protects cooperators from exploitation by defectors. We label this mechanism henceforth as cooperation selection (CS). Social influence (SI) refers to individuals assimilating to their peers' behavior (Coleman, 1990; Friedkin, 2001). Influence processes operate via network relations as social conduits through which individuals influence, motivate, or imitate others (Axelrod, 1984; Ehlert et al., 2020; Fehr & Gächter, 2002; Helbing & Yu, 2009). However, for sustaining cooperation, SI processes are a double-edged sword. On the one hand, cooperators might influence defectors to become more cooperative. On the other hand, cooperators may also be influenced by defecting counterparts. While both CS and SI have been shown to promote cooperation in theoretical analyses and experimental studies, it is unclear how effective they are in empirically more realistic settings where both mechanisms operate simultaneously or differently than assumed theoretically, alongside other processes.

5

In the real world, CS and SI may reinforce or interfere with each other, and their effects on cooperation may be distorted by many other processes known to affect people's behavior and relational choices in networks (Steglich et al., 2010; Steglich, 2018; also see Chapter 1). First, individuals do not always behave rationally; we make errors and do not always choose the best course of action in the long term. In this chapter, we model the effects of CS and SI on cooperation, assuming that agents myopically strive to optimize their cooperation behavior and relational choices, ignoring longer-term and strategic implications of their choices (Axelrod, 1984), as well as insights from learning or diffusion models (Macy & Flache, 2002; de Matos Fernandes et al., 2022a; Chapter 3). Myopic stochastic optimization[2] assumes that behavior is probabilistically based on what is considered best at this point, the best option in the short term based on the behavior of others around them. To be clear, myopic optimization assumes that actors do not pursue the highest payoff possible or behave in such a way based on insights from the past of themselves or those around them.

2 Myopic stochasticity builds on present literature describing non-rational approaches to behavior, stating that decisions, in general, are affected by stochasticity, bounded rationality, error, inertia, limited information, or uncertainty (Coleman, 1990; Mäs & Nax, 2016; Simon, 1982; Udehn, 2001; Wittek et al., 2013), and so are decisions about cooperation or defection.

Second, many social processes operate in real life and are generally not included in experiments and theoretical models, which tend to focus on CS and SI alone. For instance, a persistently strong driver for network formation is relational reciprocity, suggesting that the mere presence of a nomination makes one more inclined to return the nomination, overwriting other mechanisms. For example, reciprocity may undercut CS if a defector befriends a cooperator. Similarly, the empirically documented tendency to pursue transitivity in social relations may spur network relations between cooperators and defectors who are friends of the cooperator's friend, again countering CS.

Third, one important reason CS is argued to promote cooperation is that the exclusion from valuable cooperative relationships serves as a punishment for defection, forcing the defector to change their behavior. However, this does not have to be the case in real life. Defectors may even prefer to relate to defecting partners who value things more similarly than cooperators. Prior research suggests that seeking similar defectors as partners can be strategically rational because it allows defectors to avoid sanctions from cooperators they might face otherwise (Takács et al., 2008). For instance, a defector as a partner may be more attractive if (s)he provides resources or support that a cooperator cannot give. While this exacerbates the cooperation problem the collective faces because the defector is not punished for his behavior, it can be beneficial from the point of view of the defector's network partner (Bianchi et al., 2020; Flache & Hegselmann, 1999). For cooperators, CS may resemble a tendency for homophily (McPherson et al., 2001). Homophily implies that actors similar to each other in terms of behavior are more likely to form network relationships. As argued, the same may hold for homophily by cooperation in which cooperators and defectors preferentially seek similar others out. This differs from CS commonly implemented in experiments and models in that defectors prefer to connect to other defectors in our model. We come back to this issue in the discussion.

These complications for CS and SI to promote cooperation set the stage for the contribution of Chapter 5. To our knowledge, we provide one of the few studies assessing systematically which mix of SI and CS might optimally promote cooperation under empirically realistic conditions. We use an empirically calibrated stochastic-actor-oriented model (SAOM) as an agent-based computational model (ABCM). The empirical calibration consists of importing data on which ties agents initially have at all, as well as the degree of tie reciprocity and transitive closure they exhibit in the simulated network evolution, based on an estimation with a SAOM of the corresponding model parameters in the empirical data set. The empirical dataset on the basis of which we calibrate our ABCM contains longitudinal friendship and cooperation data from 95 students in higher education. We first calibrate the model, then use it to conduct counterfactual theoretical simulation studies in which we vary the relative strength of SI and CS to explore how this would affect the emergence and sustainability of cooperation among the 95 students according to our model.

5.1.1 Abstracting from real life

If we abstract away the complexity of multiple interfering processes of network formation and behavioral dynamics, whether the interplay of CS and SI successfully promotes cooperation depends on the exact mix of SI and CS as well as on further initial conditions (i.e., the initial prevalence of cooperation and the initial network structure). In more abstract and simplified theoretical models, we can theoretically expect three potential outcomes of the interplay between CS and SI (see Figure 5.1). First, cooperation *spreads* if SI operates in a network where the majority is cooperative and CS does not undercut the social relations through which cooperators influence defectors, or CS motivates defectors sufficiently to change their course through the threat of exclusion. Second, cooperation *dies out* if SI operates under less-than-ideal circumstances, for example, in a network where the majority is initially defecting. In such a world, even if cooperators strive to avoid relating with defectors due to CS, due to the double-edged nature of SI, they may still have too many initial encounters with defectors to avoid being "infected" by their counterparts' defection. In the end, cooperation may slowly (or quickly) die out. Third, cooperation *segregates* when cooperators and defectors form separate clusters due to CS processes, while SI ensures in each cluster the continuation of the behavior that is dominant in the respective cluster.

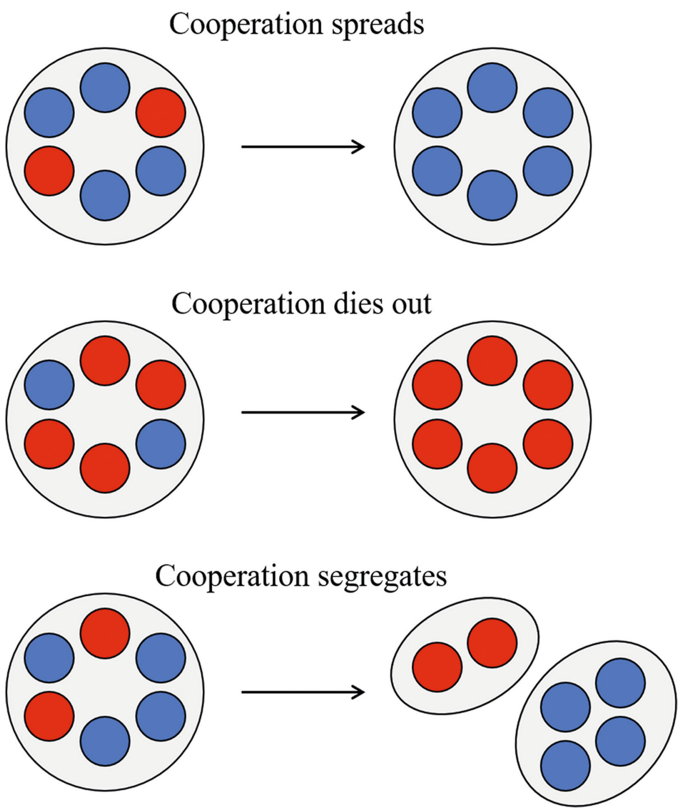


Figure 5.1: The three potential outcomes of cooperation visualized in a simplified example in which 6 agents are present. Blue circles are cooperators and red circles are defectors.

5.1.2 Do we need an empirically-calibrated model?

Field studies and lab experiments provide limited opportunities to inspect the theoretical consequences of changes in SI and CS's relative strength and presence. In contrast, theoretical models typically neglect many complicating factors relevant to empirical cooperation problems. To circumvent the limitations of both approaches, we rely on a model based on empirical calibration. More precisely, we employ SAOMs (Snijders et al., 2010; Snijders, 2017; Steglich et al., 2010). SAOMs assume myopic stochastic optimization, so unlike in many theoretical models of cooperation, actors are not assumed to anticipate future outcomes or learn from or be selected for successful behavior from the past. Moreover, actors make behavioral and relational choices based on what is considered best from their point of view.

Moving beyond the common practice in the mainly empirically oriented SAOM literature, we use empirically calibrated simulations (adams & Schaefer, 2016; Bianchi et al., 2020; Boda et al., 2020; Kretschmer & Leszczensky, 2022; Snijders & Steglich, 2015; Stadtfeld, 2018; Stadtfeld et al., 2020; Steglich, 2018; Steglich & Snijders, 2022). Specifically, we simulate behavior and network decisions and outcomes based on parameters empirically estimated with SAOMs from longitudinal observations in a specific empirical setting. The main benefit of our approach is that we can assess the effects of the relative strength of SI and CS on cooperation based on an empirically realistic representation of the network context under consideration. This chapter can be seen as an implementation of the strategy of "decreasing abstraction" advocated by Lindenberg (1992) for cooperation models (see Chapter 1, for more information). Simple analytical models provide clear tractable results but may lack empirical relatedness, whereas overly complex models may be realistic but intractable. Using existing theoretical models of cooperation as a benchmark, we decrease abstraction by inserting realistic relational and behavioral choices as well as contextual assumptions in our model. Our empirically-calibrated model provides us with realistic descriptiveness but, more importantly, also the possibility to understand what is happening "under the hood" of the model when interpreting its findings (for a discussion of inserting realism in theoretical, computational models in a step-wise fashion, see Flache & de Matos Fernandes, 2021).

The agents in our empirically-calibrated model take more considerations into account than the "standard" theoretical model of cooperation when making behavioral and relational choices. Theoretical cooperation models generally include two possible decisions: An agent decides to (*i*) cooperate or defect and (*ii*) maintain or sever the network tie (e.g., de Matos Fernandes et al., 2022a; Fehl et al., 2011). In many experiments and models, cooperation is often either a 0 (defection) or 1 (cooperation) or measured in terms of the points allocated to contribute to a public good. Behavioral and relational choices depend mostly on what others did in the past ("did they cooperate or defect"), past and prospective payoffs, and the behavioral mechanism implemented (see footnote 2).

We, however, model cooperation not as a dichotomous state but as a multinomial feature. Cooperation in real life is often unobservable and relative. That is why we include whether one is more neutral, defection-oriented, or cooperation-oriented as a non-fixed behavioral construct. An agent optimizes myopically by choosing a behavior the more likely, the higher its expected utility in a multinomial logit function given the network and behavioral state at that time. Input for the logit function modeling behavioral change comprises (*i*) social influence effects ("what others do"), (*ii*) a linear preference for neutral, defection-oriented, or cooperation-oriented, and (*iii*) whether one's behavior is reinforced (or not) by one's currently shown behavior (interaction effect of behavior on behavior). Our model thus includes contextual influences – as "common" theoretical cooperation models – but also myopic stochasticity.

Furthermore, relational choices are governed by a separate multinomial logit model. Input for this function modeling relational change comprises (*i*) the structure of the network ("where you are and whom you are connected to impacts how beneficial it is to form or sever a tie to a specific other"), (*ii*) dynamics of the network in which agents can select others around them to form ties with based on reciprocity or transitivity considerations, and (*iii*) cooperation selection effects in which similar actors are more likely to form ties than dissimilar ones. Our empirically-calibrated model extends cooperation models in which networks are dynamic by incorporating the three listed features.

In what follows, we first introduce our approach before describing the data and the empirically calibrated model. We then present our findings as well as explore additional scenarios of interest. We end this chapter with some concluding remarks.

5.2 OUR "WHAT IF" APPROACH

We explore counterfactual situations our model permits us to implement (Steglich, 2018). For example, a counterfactual condition can be generated by altering the strength of CS and/or SI relative to what has been observed in the real empirical setting under study and then exploring the expected impact on cooperation and network structure via simulations. In our empirical setting, the role of SI and CS in network formation and cooperation behavior is very weak at best. This makes it especially relevant to ask how cooperation would be affected if, for example, network intervention measures (Steglich, 2018; Steglich & Snijders, 2022) would strengthen these processes relative to each other and other mechanisms. In what follows, we discuss the three counterfactuals or "what if" situations that arise when we assume, respectively, a much stronger role of SI, CS, or both relative to the empirically observed situation.

5.2.1 Counterfactual 1: *I am under your spell*

We first consider what happens if SI is much stronger than we have observed empirically. Suppose peers actively influence agents via social learning (Ehlert et al., 2020), imitation (Helbing & Yu, 2009), contagion (Fowler & Christakis, 2010), or tit-for-tat (Axelrod, 1984) dynamics. The network will then likely "tip" towards the initial majority behavior. In other words, your network partners and what they do influence what you do. The double-edged nature of SI points to two potential outcomes of counterfactual 1. On the one hand, if most are cooperative, cooperation eventually prevails. This is the "cooperation spreads" outcome. In this scenario, cooperation spreads because defectors assimilate to the cooperative majority in their network. On the other hand, if the majority is defecting, these defectors influence others also to defect. Cooperation then "dies out."

Whether cooperation spreads or dies out with SI as the prevalent mechanism depends on the initial distribution of cooperation in the network and the inclusion of further empirically observed processes and contextual conditions. For instance, individuals may still engage in network selection, but not only based on CS considerations. The clustering generated by such selection processes may prevent convergence of the entire network on either cooperation or defection due to SI. For SI to occur, ties must be present: One individual can only influence another if they are connected. Moreover, cooperative behavior may change over time due to non-SI reasons. For example, an individual may optimize myopically – shortsighted and without concern for future or broader implications – behavior to change from cooperation to defection or vice versa. Behavior is sometimes suboptimal, for instance, due to bounded rationality or heterogeneity in unobserved preferences. We assess in counterfactual 1 the impact of these complicating behavioral and social factors on the effectiveness of SI as a prominent process for spreading cooperation.

> **Counterfactual 1** assesses whether strong SI promotes more cooperation than in the empirically observed condition of SI, while other processes affecting network formation and cooperation match those observed in our empirical setting.

5.2.2 Counterfactual 2: *Pikachu, I choose you*

In counterfactual 2, we inspect a condition where actors selectively form or dissolve network ties based on others' behavior (CS) without SI. Ideally, CS allows cooperators to preferentially form ties with other cooperators, protecting themselves from exploitation and motivating defectors to change their behavior (Fehl et al., 2011; Helbing & Yu, 2009). However, CS may also have the effect that defectors end up in parts of the network where defection predominates, without access to the benefits of relations with cooperators. As often argued in the literature using theoretical models, CS would induce the segregation of cooperators and defectors (Epstein, 1998; Gross & Dreu, 2019; Nowak & May, 1992; Schuessler, 1989; Waldeck, 2013). In addition, as we argued earlier, there can also be benefits for defectors to form ties to similar others (e.g., being shielded from social scaffolding), so it is important to explore the consequences of a preference of defectors to relate to similar others.

The effects of CS may, in addition, interfere with other network selection mechanisms. For instance, Melamed et al. (2020) showed that cooperation is superseded as input for network selection by demographics unrelated to cooperation, such as university affiliation. A potential partner of a similar university was more attractive as a network partner in Melamed et al.'s cooperation experiment than a random cooperator. McPherson et al. (2001) list many more features, e.g., attitudes, socio-demographics, ethnicity, and gender, affecting network selection via homophily. Thus, CS may fail to motivate defectors to change to cooperation because exclusion or the threat thereof is countered by other network processes, ensuring that defectors maintain cooperative network partners (potentially not leading to segregation).

To inspect whether CS induces segregation of cooperators and defectors, also when additional empirically observed network mechanisms are considered, we inspect the consequences of increasing the strength of CS relative to other present network mechanisms. The interplay between CS and other selection mechanisms is thus of interest to us: Will cooperation still segregate when CS is combined with other network formation processes as we observe them in our empirical setting? If it does not, why not?

> **Counterfactual 2** addresses whether strong CS gives rise to more segregated cooperation over time than in the empirically observed condition of CS, while other processes affecting network formation and cooperation match those observed in our empirical setting.

5.2.3 Counterfactual 3: *It takes two to tango*

Accounting for CS and SI as interdependent and simultaneously active decision-making mechanisms is a key real-life element of the "empirical world" in which multiple mechanisms operate simultaneously. While SI ideally pushes agents in the network towards imitating the behavior of the initial majority, CS can prevent this process by isolating the initial minority from "outside" influences. If CS operates next to SI, it can either prevent a defecting minority from being changed into cooperators or prevent a cooperating minority from being changed into defectors. As such, CS may thus counter the trend of SI for behavior to converge. It may be the case that the combination of CS and SI fosters the spreading of cooperation because it leads actors towards quickly forming a relationship with formerly defecting others who, perhaps accidentally, cooperated for once. This suggests that initial defectors can escape from defection and clusters in which defection is common if they cooperate at some point. However, the parallel occurrence of SI and CS may similarly suggest that CS and SI fuel a dynamic in which defectors pull cooperators who defect into their clusters and socially influence them to defect. Whether cooperation spreads, dies out, or segregates when strong CS and SI operate simultaneously depends on how strong each process is relative to the other, what the initial conditions are, and the other social processes affecting network formation and cooperation observed in this chapter.

> **Counterfactual 3** assesses whether the combination of CS and SI spreads or founders cooperation more over time, or gives rise to more segregated configurations than in the empirically observed condition of CS and SI, given that other processes affecting network formation and cooperation match those observed in our empirical setting.

5.3 THE MODEL

5.3.1 Data

Our analysis builds on data obtained from 95 first-year sociology students in a bachelor's program at a university in the Netherlands (same data as in Chapter 4). The longitudinal dataset comprises variables on friendship relations and cooperation. Students answered a 20 to 30-minute computer-based questionnaire twice in an academic year. Friendship network and cooperation data were collected at the end of the first semester (wave 1) and second semester (wave 2). Friendship network data is collected by students evaluating their fellow students on a scale from 1 ("best friends") to 6 ("I don't know who this is"). This scale is converted to a dichotomous variable representing friendship ties to analyze the friendship network using stochastic actor-oriented models. This is done as follows: 1 = "best friends," 2 = "friend," and 3 = "friendly relationships" are coded as 1 (a friendship tie). Whereas 4 = "neutral, not much in common," 5 = "only known from face or name," and 6 = "I don't know who this is" are coded as 0 (no friendship tie).

Data on cooperative behavior is hard to collect in fieldwork because the cooperativeness of an act is not easily determined compared to experiments or theoretical models. In such experiments and models, cooperation is often either a 0 (defection) or 1 (cooperation), or the points an individual allocates to contribute to a public good. More points mean that one is more cooperative. Cooperation in real life is often unobservable and relative. To circumvent this problem, we utilize information from a preference-for-cooperation question. Students indicated "I would like to collaborate with [name]" via a 5-point Likert scale, ranging from 1 ("strongly disagree") to 5 ("strongly agree"), with the option of 6 ("I do not know"). Categories 1, 2, 3, and 6 are re-coded in the matrix as "0" and 4 and 5 as "1". A "0" refers to not perceiving the other as cooperative, whereas a "1" points to the opposite. We then aggregated nominations cumulatively per student. Higher scores indicate that one is more attractive as a cooperative partner. For interpretation purposes, we re-scale the aggregated variable to a 3-point scale. Scores lower than 5 are re-coded to "1", scores 5 and 6 as "2", and higher than 6 as "3". The question thus captures attractiveness as a cooperator instead of behavior.

This measure partially captures cooperation considerations but also how disciplined students are in advancing in their studies and how others assess the quality of their work. The way we measure cooperation is thus not cooperative behavior as commonly measured in experiments and theoretical models. Without any other behavioral information, we rely on this reputation to indicate students' cooperativeness. The variable indicates whether students are considered less attractive as a cooperator (a score of 1, a defector if you will), a more neutrally attractive student as a cooperator (a score of 2), or perceived as a cooperative student by many (3). At wave 1, we have $n = 32$ with a score of 1, $n = 24$ more neutrally attractive students (a score of 2), and 39 with a score of 3. The mean at wave 1 is 2.07 ($SD = 0.87$). At wave 2, we have $n = 37$ with a score of 1, fewer with a score of 2 ($n = 16$), and $n = 42$ with a score of 3. The mean at wave 2 is 2.05 ($SD = 0.92$). In what follows, we refer to this variable as cooperation for interpretation purposes. Still, we want to stress that it does not refer to cooperation commonly assessed in experiments and theoretical work. A score of 1 refers to a defector here, a score of 2 is a neutral actor, and a score of 3 represents a cooperative actor. The major advance of Chapter 5 lies in capturing CS and SI in such a way as to accommodate social and behavior mechanisms found in real life and not to further cooperation research per se.

5.3.2 Stochastic actor-oriented models

We rely on stochastic actor-oriented models (SAOMs) to analyze the data. SAOMs allow us to assess how individual features and networks co-evolve from one point in time to another (Snijders et al., 2010; Snijders, 2017; Steglich et al., 2010). This requires data from at least two waves. Using SAOMs, we can identify network reasons – such as reciprocity and transitivity – for students to become friends with each other. Also, SAOMs allow us to inspect whether the formation of network ties depends on cooperation and whether peers influence behavior. Core features of SAOMs are the use of longitudinal network and behavioral data and the use of ABCM to simulate changes in-between waves. Succinctly, changes in the friendship network and cooperation between waves are simulated via mini-steps. An example of a mini-step is forming, maintaining, or dissolving a single relationship. An example of a behavioral mini-step is changing from cooperation to defection. Changes in friendship and behavior are the result of agents' decisions.

Actors myopically optimize one network relation or cooperation state at a time. How many mini-steps an actor can take is modeled via the rate function, whereas the objective function determines which actions to take in a mini-step. SAOMs assess which theoretically postulated mechanism can best generate a sequence of mini-steps that explains changes between observed waves. The objective functions capture network and behavioral factors that affect network tie formation behavior (selection) or whether friendship affects behavioral changes in cooperation (influence). The objective functions of network selection ($f_i^{\mathbf{net}}$) and behavioral changes ($f_i^{\mathbf{beh}}$) for an agent (*i*) are respectively defined as:

$$f_i^{\mathbf{net}}(ß,x,z) = \Sigma_k ß_i^{\mathbf{net}}\, s_{ik}{}^{\mathbf{net}}(x,z) \tag{5.1}$$

$$f_i^{\mathbf{beh}}(ß,x,z) = \Sigma_k ß_i^{\mathbf{beh}}\, s_{ik}{}^{\mathbf{beh}}(x,z) \tag{5.2}$$

Actors pursue the best possible $f_i^{\mathbf{net}}$ and $f_i^{\mathbf{beh}}$ state, given the state of the network (x) and the individual attribute of interest (z) at a certain mini-step. $s_{ik}{}^{\mathbf{net}}(x,z)$ and $s_{ik}{}^{\mathbf{beh}}(x,z)$ are the effects, think of CS, SI, reciprocity, and tendencies towards transitive closure. The $ß$'s are the weights of the s parameters, similar to generalized linear models. $f_i^{\mathbf{net}}(ß,x,z)$ and $f_i^{\mathbf{beh}}(ß,x,z)$ capture the value of the objective function of network selection and influence, respectively. In any given mini-step, the action that an actor chooses is more likely to be chosen, given the objective function, the higher the value of the network state resulting from that action. This discrete choice results in severing or forming ties (or maintaining the network tie) or changing (or maintaining) behavior. The objective function shows via a multinomial logit choice model how attractive a friendship network state or change in cooperative behavior for a student is, thereby controlling for various structural friendship network (e.g., reciprocity and transitive closure) parameters. The model is actor-oriented because it takes *i* as the locus of modeling. The rate and objective function of both influence and selection operate simultaneously to control for the inherent co-evolutionary interdependence. Finally, SAOMs estimate network and behavioral changes on the micro-level, i.e., the local network structure (selection) and whether those friends affect cooperation (influence). The SAOM simulation algorithm, in combination with statistical methods for model fitting and model selection, is used to assess which parameter values for the rate function and the effects constituting the objective function included in a model yield the best match with the empirically observed changes in network relations (selection) and behavior (influence).

5.3.3 Specifying effects in a SAOM

The SAOMs are assessed in R (R Core Team, 2021) using the package Simulation Investigation for Empirical Network Analysis (*RSiena*) (Ripley et al., 2021). In *RSiena*, we specify the effects that correspond to assumptions we make about which network properties or their association with individual characteristics (such as cooperation) actors pursue in choosing their behavioral and network actions. In our approach here, we want to focus on the effects of the relative strength of CS and SI, which leads us to specify a model that is as simple as possible yet still yields a reasonable fit to the observed empirical data in the calibration phase of our study. In what follows, we introduce effects included in the selection and influence model. *RSiena* terms are mentioned in italics.

First, we discuss the effects included in the selection model. The rate effect indicates how frequently actors can change ties over time. Three friendship network effects are included, comprising outdegree, reciprocity, and transitive closure. The outdegree (*density*) indicates the general tendency of actors to form ties in the network ($i \rightarrow j$). Reciprocity (*recip* in *RSiena*) indicates if there is a tendency towards reciprocal relations ($i \leftrightarrow j$). We model the likelihood of transitivity via *TransTrip*; e.g., if $i \rightarrow j$ and $j \rightarrow k$, then including this effect assesses how much this increases the chances that a tie $i \rightarrow k$ is initiated or maintained by i. For cooperation as an individual attribute related to homophily, we add *simX* in the objective function of the selection model. *simX* captures tendencies of unrelated students (i^* and j^*) to befriend others similar in cooperativeness ($i^* \rightarrow j^*$). We label this effect cooperation selection (CS). Note that most theoretical models studying cooperation selection assume that cooperators are preferred as friends, yet, we argued that homophily by cooperation might similarly guide relational choices. That is why we implement alternative scenarios in which we model cooperation popularity. The specifics are provided in the results section.

Next, we include the following effects in the influence model to inspect changes in cooperation. *Rate* indicates the occurrence at which changes in cooperation are likely to occur. The *linear* shape effect elucidates if there is a linear trend towards higher (positive effect) or lower (negative effect) values across waves. The quadratic shape effect (*quad*) models dispersion and accounts for potential non-linear re-enforcement effects (positive estimate) or self-correcting effects (negative estimate) of changes in cooperation. The quadratic effect operates next to the linear shape effect. For example, a positive significant quadratic shape effect with a non-significant linear shape effect induces an increasingly bimodal distribution of behavior over time. If the linear and quadratic shape effects are insignificant (so-called null effects), then the behavior is not subject to many changes, ceteris paribus. Social influence (SI) is estimated via *avSim*, indicating whether students align their cooperative behavior to the average level of friends' cooperative behavior ($i^* \rightarrow j^*$). Note that both CS and SI can induce a pattern of cooperation homophily in which connected actors are more similar ($i^* \rightarrow j^*$) than non-connected ones, either because they select similarly cooperative others as partners or because they adjust their cooperative behavior to resemble that of their network partners.

From *RSiena*, we can extract each effect's relative importance (RI) in the objective function (Indlekofer & Brandes, 2013). RI percentages refer to the degree relational and behavioral choices in mini-steps can be attributed to a particular effect in the model. Combining all percentages in the selection and influence model cumulatively leads to 100 percent. In what follows, we mention the RI of the empirical and empirically calibrated SAOM.

Table 5.1: SAOM findings for selection and influence in a friendship network.

	Selection model				Influence model		
Parameter	**Est. (SE)**	***p***	**RI**	**Parameter**	**Est. (SE)**	***p***	**RI**
Rate	12.75 (0.97)	<0.01		Rate	3.18 (0.90)	<0.01	
Outdegree	−2.51 (0.07)	<0.01	49.9%	Linear sh.	0.05 (0.19)	0.78	1.1%
Reciprocity	1.88 (0.13)	<0.01	29.7%	Quadratic sh.	1.55 (0.30)	<0.01	98.0%
Transitivity	0.24 (0.02)	<0.01	18.0%	SI	1.18 (1.06)	0.27	1.0%
CS	0.22 (0.24)	0.35	2.4%				

Note. Est. = estimate; *SE* = standard error; RI = relative importance; sh. = shape; RI percentages are rounded; CS = cooperation selection; SI = social influence.

5.3.4 Results of the "empirical" SAOM

We run a SAOM in which selection and influence are estimated in friendship networks with the data and *RSiena* effects as described above. The negative outdegree coefficients in Table 5.1 – the more negative the outdegree, the sparser the network is – indicate that there is a general tendency not to nominate another agent as a partner unless the nomination brings with it desirable consequences, such as reciprocating an incoming tie or forming a relationship based on CS. The sparsity of the network is also reflected in the low and decreasing degrees in the friendship network: From 6.12 at wave 1 to 4.99 at wave 2. The rate parameters indicate that in between the two waves, the model assigns almost 13 opportunities to change network ties to the actors. In contrast, they have only approximately three chances to alter their cooperation score. Students further tend to reciprocate friendship nominations and strive for transitivity. Shape effects suggest that actors tend to pull to either defection or cooperation.

We find little to no evidence for CS and SI. This is also reflected in the relative importance (RI) percentages in Table 5.1. Namely, CS and SI account for 3.8% and 1.0% of changes in network relations and behavior, respectively. Most changes in the network are driven by the outdegree estimate (54.8%), reciprocal considerations (24.6%), and the pursuit of transitive closure (16.8%). Changes in cooperation are driven nearly exclusively by the quadratic shape effect (98%), indicating that students tend to adjust their cooperativeness levels towards one of the two extreme ends of the scale. In our empirical context, RI percentages suggest that other social and behavioral processes are stronger than CS and SI. In what follows, we adjust the CS and SI effect to inspect whether minor changes compared to observed empirically affect cooperation levels, aiming the empirical calibration at the outdegree, reciprocity, transitivity, and shape estimates. Note that *RSiena* results model local connections (selection), i.e., in dyads and triads, and local SI from peers on changes in cooperation (influence). The model is well-fitted, given that the SAOM overall convergence ratio of 0.09 is well below the critical threshold of 0.25 (Ripley et al., 2021).

5.3.5 Measuring network segregation by cooperation

We assess segregation by cooperation in the network via the Moody gross-segregation index (de Matos Fernandes et al., 2022a; Moody, 2001; Chapter 3). The segregation index divides the sum of ties between similar actors on a certain attribute (controlled for all possible ties between dissimilar actors) by the sum of ties between dissimilar actors on a certain attribute (controlled for all possible ties between similar actors). The outcome is an odds ratio. Succinctly, odds ratios> 1 indicate network segregation, whereas odds ratios below 1 indicate integration. A score of 1 denotes that the chances for a link between individuals with similar and dissimilar scores are equal. The higher the odds ratio, the more likely the network is segregated by cooperativeness. For example, if the segregation index = 2, actors are twice as likely to form ties with similar than dissimilar others. Analyses show that network segregation by cooperation increased from 1.86 at wave 1 to 2.13 at wave 2. Interestingly we find no evidence for CS and SI over time but do find increasing network segregation by cooperation on the macro-level.

5.3.6 Towards an empirically calibrated ABCM

5.3.6.1 Decision-making model

Empirically calibrated SAOMs generate output networks conditional on estimated parameters (Snijders & Steglich, 2015; Steglich & Snijders, 2022). In this chapter, we extract the empirical estimates outlined in Table 5.1. We start with a network of 95 actors, mirroring the empirical situation. We simulate an additional two waves, representing two academic years in total. To keep it as simple as possible and to be able to interpret the impact of both effects as clearly as possible, we set the CS and SI parameter to either zero or thrice the empirical estimate, which is a common implementation using the current methodology (adams & Schaefer, 2016; Steglich, 2018). Table 5.2 shows the resulting models, composed of the empirically estimated parameters of all effects except for those of CS and SI, while the effects of CS and SI are replaced with their values for the counterfactuals. The relative importance of CS and SI increased – ceteris paribus – by approximately 6% and 2%, respectively, by adopting the higher values. The relative importance percentages are based on a re-calibration of the model after increasing the empirical values of CS and SI times three.

Table 5.2: Parameters of the empirically calibrated ABCM.

Selection Parameters		**Influence Parameters**	
Fixed	Estimate	*Fixed*	Estimate
Rate	12.75	Rate	3.18
Outdegree	–2.51	Linear shape	0.05
Reciprocity	1.88	Quadratic Shape	1.55
Transitivity	0.24		
Manually adjusted		*Manually adjusted*	
Cooperation selection	[0, 0.66]	Social influence	[0, 3.54]

Table 5.2 shows the decision-making model numerically, whereas equations 5.1 and 5.2 show it as a function. The probabilistic choice model assumes that the better the outcome brought about by a given behavioral or tie change from the agent's perspective, the more likely this change will be pursued. There are roughly 13 chances to alter network connections and roughly 4 occasions to alter cooperation (rate parameters in Table 5.2). If the opportunity arises to update actors' network relations, the "best" network state is determined by combining the outdegree, reciprocity, transitivity, and CS parameter. Seeking similar others in cooperativeness is either turned off (estimate = 0) or thrice the empirical case (estimate = 0.66). The latter will account for 6 percent in network changes in the simulation, cf. 2.4% in the empirical baseline. Next, for cooperation, changes are determined by the linear shape, quadratic shape, and SI parameter. Adjusting cooperation to the average of peers' cooperation levels is either absent (estimate = 0) or thrice the empirical size (estimate = 3.54). The increase in SI will account in the simulations for approximately 2 percent in behavioral changes. In sum, each agent makes two choices: (1) to drop, form, or maintain a network tie based on the selection function, and (2) to increase, decrease, or maintain cooperative behavior based on the influence function. The difference per agent is the local context. To whom one is tied, what peers' cooperation scores are, and whether one is embedded in network parts in which cooperating or defecting prevails.

5.3.6.2 Simulation algorithm

For our empirically calibrated ABCM, we rely on the *RSiena* algorithm commonly used to assess longitudinal network and behavior data (Ripley et al., 2021). In short, we build on the simulation procedure of the *RSiena* algorithm and tweak the SI and CS parameters to assess our counterfactuals. The approach taken in this chapter consists of several steps visualized in Figure 5.2. Each simulation run models cooperation and network relations changes on two additional occasions (*M* = 2). This roughly equals 4 semesters or two academic years, using the empirical configurations as year 1 (waves = 1 and 2) and the simulated configurations as year 2 (waves 3 and 4). The configuration of the prior wave (*M* – 1) is used as input for the concurrent wave. The parameter setup is always fixed, with the sole exceptions being CS and SI. Input for the model is the empirical network and cooperation distribution. We then simulate two additional waves based on the composition of the network and cooperation data in the prior wave. For more information on creating *RSiena* data formats, *RSiena* effects, and the siena07 algorithm, we refer to Ripley et al. (2021). The full code of our model is available in our Open Science Framework (OSF) online repository.[3]

Algorithm 1: Pseudocode of a simulation run	
initialization	
input	number of agents = *n*, waves = *M*, rate selection = *rate*, density parameter = *dens*, reciprocity degree = *rec*, transitivity tendency = *tt*, cooperation selection = *cs*, rate influence = *rate.b*, linear shape = *lin.b*, quadratic shape = *qu.b*, social influence = *avsim.b*
include	empirical network (*n* x *n* matrix) and cooperation (*n* x 2 matrix) data in **RSiena data** format
include	*rate, dens, rec, tt, cs, rate.b, lin.b, qu.b,* and *avsim.b* as **RSiena effects**
simulate	*m* = 1 cooperation and network data based on **RSiena data** and **RSiena effects** using **siena07 algorithm**
for	*m* = 1 to *M* waves
simulate	*m* + 1 cooperation and network data based on **RSiena effects** specified and network and cooperation **RSiena data** at *m* using **siena07 algorithm**
set	*m* (*m* + 1)
report	network (*n* x *n* x *M*) and cooperation (*n* x *M*) matrices for all simulated waves

Figure 5.2: A visualization of the flowchart of the simulation procedure.

5.3.6.3 Analytical approach

We set CS and SI as the baseline condition to 0 (Table 5.2). To assess counterfactual 1, we manually set the CS parameter to 0 and turn the SI effect “on.” Next, counterfactual 2 is assessed by assigning CS 0.66 and turning SI “off.” Finally, we set CS and SI “on” to assess counterfactual 3. We run 50 independent simulations per condition, realizing 200 simulation runs, all other things being equal. We contrast cooperation levels and network segregation by cooperation observed in our model to those observed empirically at wave 2.

3 The simulation code is available via https://doi.org/10.17605/OSF.IO/RKG7B. The code used is based on a R-script freely available at https://www.stats.ox.ac.uk/~snijders/siena/NetworkSimulation.R.

5.4 FINDINGS

5.4.1 "What if" conditions: Varying the strength of CS and SI

Figure 5.3 shows the aggregate level of cooperation in the empirical situations at waves 1 and 2 and the simulation experiments at waves 3 and 4. The mean and 95% confidence intervals are shown. The mean at wave 1 is 2.07 (95% CI [1.90, 2.25]) and lowered to 2.05 at wave 2 (95% CI [1.87, 2.24]). Next, in the baseline condition without CS and SI, the blue line in Figure 5.3 shows that the simulated level of cooperation lowered a bit from wave 2 to wave 4: From 2.01 (95% CI [1.98, 2.04]) in wave 3 to 1.86 (95% CI [1.84, 1.89]) in wave 4.

Based on current information on CS and SI in the literature, we worked out three potential outcomes for the three counterfactuals we inspect: Cooperation (*i*) spreads, (*ii*) dies out, or (*iii*) segregates. We find that with the inclusion of the behavioral (myopic stochastic optimization) and social (other competing social mechanisms as well as CS and SI competing with one another) empirical regularities is not so clear-cut which of the different outcomes is most likely to arise in which counterfactual.

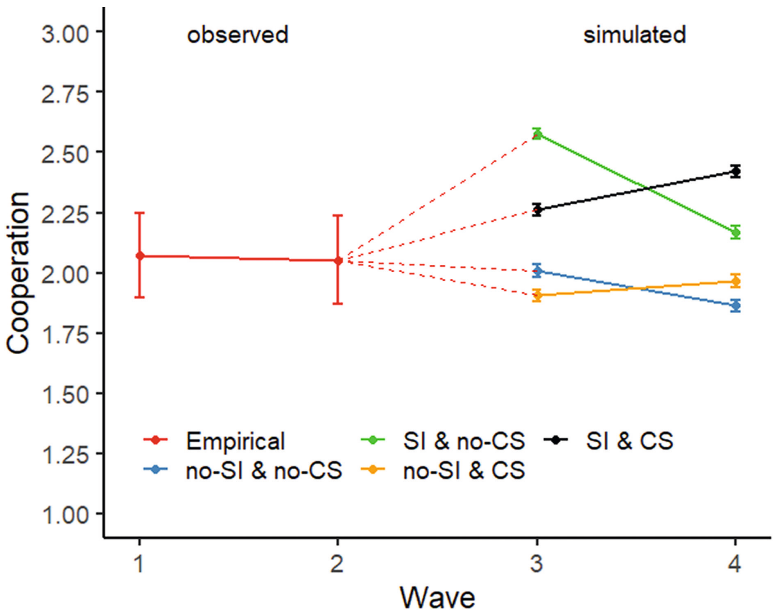


Figure 5.3: Visualizing the mean and 95% confidence intervals of cooperation levels generated via 50 simulation runs per proposition. The empirical distribution of cooperation is included in wave 1 and wave 2 (red). Dashed red lines are inserted to link the empirical outcome at wave 2 to the starting point of the simulations in wave 3. Note: SI = social influence; CS = cooperation selection.

In the SI and no-CS condition (counterfactual 1), we infer from Figure 5.3 that cooperation is highest in wave 3. Specifically, the mean of cooperation is 2.58 (95% CI [2.56, 2.60]). Yet, cooperation decreases to 2.17 (95% CI [2.14, 2.19]) in wave 4. Strong SI promotes cooperation at first

– as argued in counterfactual 1 – but then tends to die out. The question is whether network mechanisms or behavior shape effects (or both) affect the upwards-downwards distribution of cooperation over time.

The next "what if" scenario is the condition with only CS (counterfactual 2). We see in Figure 5.3 that cooperation declines from wave 2 to wave 3. Yet, we see a steady increase from 1.91 (95% CI [1.88, 1.93]) in wave 3 to 1.97 (95% CI [1.94, 1.99]) in wave 4. For counterfactual 2, we proposed segregated cooperation as the most probable outcome in the only CS condition. However, we cannot infer whether cooperation is segregated from the mean value of cooperation alone. That is why in section 5.4.2, we dive deeper into the dispersion of cooperation.

Finally, counterfactual 3 considers the combination of SI and CS and the potential consequences for cooperation. Cooperation is immediately higher in wave 3 than it was found empirically: Mean cooperation = 2.26 (95% CI [2.24, 2.29]). Notably, cooperation is even the highest – black line in Figure 5.3 – among all simulation scenarios at wave 4: Mean cooperation = 2.42 (95% CI [2.40, 2.44]). The dual process of SI and CS is thus the ideal condition for cooperation to spread compared to the other conditions, meaning that overall cooperation levels are highest.

Even so, the question is *why* we see current cooperation levels. Why can we not infer clearly whether cooperation spreads, die out, or segregates? What drives the effects in Figure 5.3? In the following section, we highlight the underlying drivers of the findings in Figure 5.3.

5.4.2 Explanation of simulation findings

We explain the findings in Figure 5.3 by inspecting (*i*) network connections on the micro (pairs of actors) and macro (network segregation) level, (*ii*) how the shape effects govern changes in behavior, and (*iii*) our explanations for the effects we found by inverting the initial cooperation levels taken from the empirical data (and inspecting the consequences on cooperation and segregation levels).

The first feature encompasses the network structure, affecting the SI scope. On the micro-level, we assess the local subnetwork by inspecting network pairs. We inspect which actors are most likely selected as network partners. According to the literature, many network ties among cooperatively oriented actors are key for cooperation to arise more readily in networks (Chwe, 1999; Gould, 1993; Granovetter, 1978; Kim & Bearman, 1997). We evaluate whether this intuition holds in our model. For SI to promote cooperation, moreover, a key feature is that defectors connect to cooperators and neutral actors. A byproduct of inspecting pairs of actors at the micro-level is finding out whether cooperators and defectors differ in their contribution to network segregation by cooperation. Segregation is measured via the Moody segregation index (Moody, 2001). The structure of the network on the micro and macro levels affects SI. Namely, cooperation tends to be sustained in clusters of cooperators, whereas the same holds for defection in clusters where this behavior prevails.

The second explanatory feature entails inspecting which response agents will most likely choose at a given time. The combination of the decision-making model's linear and quadratic shape effects suggests that cooperation follows a bimodal distribution in which actors are pulled to either defection or cooperation (and not the neutral stance). Thus, cooperation tends to polarize. To be clear, there is a difference between cooperation polarizes (or segregates) and network segregation by cooperation. When cooperation polarizes, cooperation rates become, over time, distributed bimodally, irrespective of the network structure. Instead, network segregation by cooperation considers whether network partners tend to be similar in cooperation. If network partners tend to be similar, we arrive at a segregated network in which cooperators, neutral actors, and defectors tend to have ties to similar others.

We test our tentative explanation of the findings for the third explanatory feature by inverting the cooperation data. In the empirical data from which start the simulation, cooperative actors initially have more network ties and are generally well-connected with similar others, contrary to defectors. Because the initial network is dense for the "cooperative group," it affects prospective outcomes. For example, suppose it is mainly SI that drives an increase in cooperation in the simulations. In that case, this outcome should be reversed if the initial majority are defectors rather than cooperation. That is why we artificially reverse the behavior to inspect whether we find similar patterns or whether outcomes change in a way consistent with our explanations of the observed outcomes. For example, we inspect how it affects simulated cooperation rates if previously cooperating actors are turned into defectors connected by a highly dense subnetwork.

5.4.2.1 With whom are you connected on the local and global levels?

The findings reported in Table 5.3 – showing the simulated data in waves 3 and 4 pooled – are key for explaining the changes in cooperation reported in Figure 5.3. A tentative explanation for why cooperation first soars and then declines in the SI and no-CS condition may be that cooperators have initially (after wave 2) more ties and thus more influence on others. This would mean we should see increased cooperation from wave 2 to wave 3. However, reluctant defectors – who still form ties to cooperators in wave 3 due to reciprocity and transitivity – may pull others back into defection if cooperators cannot isolate themselves from defectors. Cooperators could sever links to defectors, reducing their exposure to influence from defectors, but this is less likely to happen without CS. Including CS next to SI would suggest that cooperators not only influence others to cooperate but can also cut ties to defectors so that they form dense local clusters with other cooperators, sustaining cooperation. In addition, CS implies that defectors preferentially select other defectors to befriend, which, combined with reciprocity and transitivity, further reduces their possibilities to influence cooperators via reciprocal friendship ties. Here, we test whether our explanations for the cooperation findings pan out based on clustering on the micro and macro level.

Table 5.3 shows local sub-networks of cooperator-cooperator, defector-defector, and dissimilar pairs.[4] The results in Table 5.3 show the average from the simulated waves 3 and 4. A pair is considered dissimilar for cooperators if it includes a tie nomination to defector or neutral types. For defectors, a tie with cooperators or neutrals forms a dissimilar pair. Irrespective of which counterfactual is studied, Table 5.3 shows that defectors fall behind in the tie formation processes. They have fewer ties and are less likely to be selected as network partners, therefore being isolated from influences on their cooperation behavior. Conversely, Table 5.3 shows that in all counterfactuals, cooperator-cooperator nominations are disproportionately more likely to occur, which is in line with the empirical case. Empirically, we find that nominations among cooperators (2.94) are more likely than among defectors (1.26). Interestingly, defectors are more active in sending ties to dissimilar others in the empirical situation. The empirical network provides a subset where we initially have relatively more cooperator-cooperator pairs than defector-defector pairs.

Table 5.3: Mean number of network ties across waves of cooperators and defectors in subnetworks with similar and dissimilar others.

Condition	C-C pairs	D-D pairs	C-*other* pairs	D-*other* pairs
Empirical	2.94	1.26	2.61	4.29
no-SI & no-CS	3.12	0.75	2.55	2.07
SI & no-CS	3.09	0.74	2.52	2.05
no-SI & CS	3.55	0.83	2.84	2.46
SI & CS	4.53	1.02	3.43	3.35

Note. C = cooperator; D = defector; SI = social influence effect; CS = cooperation selection effect; other = agent of other types.

The simulations without CS show declining cooperation levels from wave 3 to wave 4 in Figure 5.3. Cooperation does not tend to spread so easily because cooperators are still connected to other types of agents. C-*other* and D-*other* pairs remain numerous in Table 5.3 in the no-SI and no-CS, and SI and no-CS conditions. Similar C-*other* and D-*other* levels can be seen in the CS conditions. However, the problem is that agents cannot make network decisions based on cooperative considerations in the no-CS conditions. In such instances, reciprocity and transitivity may prevent cooperators from cutting ties to dissimilar agents, thereby not insulating them from SI from their dissimilar counterparts. Agents in the CS conditions face a similar issue

4 Empirically, the average outdegree is 6.12 at wave 1 and 4.99 at wave 2. The no-SI and no-CS condition generates a mean outdegree of 4.96 (95% CI [4.84, 5.07]) in wave 3 and 4.28 (95% CI [4.05, 4.50]) in wave 4. A similar distribution is found in the SI and no-CS condition: mean outdegree in wave 3 = 4.89 (95% CI [4.78, 4.99]) and wave 4 = 4.26 (95% CI [4.04, 4.48]). Outdegree is higher in the CS and no-SI condition, irrespective of wave: mean outdegree is approximately 5.20 with a 95% CI of roughly 0.2 above and below the mean. Finally, most tie nominations are found in the CS and SI condition: mean outdegree is 5.94 (95% CI [5.81, 6.07]) in wave 3, increasing to 7.26 in wave 4 (95% CI [6.91, 7.61]).

potentially arising out of reciprocity and transitivity. Still, they do have the opportunity to form ties that may be behaviorally more beneficial for them.

Figure 5.3 shows that the drop in cooperation – cf. the no-CS conditions – is countered by including CS, with or without SI. Although ties among dissimilar types may be present due to reciprocity or transitivity, CS considerations in the decision-making model allow cooperators to make network decisions based on the behavior of others. This is reflected in the simulation scenario in which both CS and SI are strong. Counterfactual 3 leads to the highest activity in sending out nominations among cooperators. For instance, on average, cooperators send 4.5 nominations to cooperators and 3.4 to dissimilar others. Defectors tend to maintain at least one nomination to similar others. This suggests that, to some degree, defectors shield themselves from SI from their cooperative counterparts by maintaining ties with other defectors. Although a defector prefers to select a defecting partner, the defector cannot find many of them, given the sparse defector-defector sub-network in Table 5.3. An interesting case is the presence of relatively many asymmetric pairs across all conditions (*C*-other and *D*-other). Other network selection mechanisms maintain the presence of dissimilar pairs. However, Figure 5.3 shows that cooperation does not die out when CS considerations are included in the simulation conditions. The scope of SI is limited when CS is included.

The question is whether local ties to similar types give rise to network segregation by cooperation. Axelrod (1997) was among the first to model how local similarities lead to global segregation patterns through social influence and homophily. A segregated network is, in our context, an outcome in which defectors, neutral actors, and cooperators form clusters of similar types. To inspect this, we visualized network segregation by cooperation per condition in Figure 5.4. The distribution of segregation over time is visualized in Figure C5 in Appendix C[5], showing no major upward or downward decline of segregation from waves 3 to 4 except in the SI & CS condition. The empirical cooperation distribution in wave 2 exhibits a slight tendency towards segregation, per the red line, at an odds ratio of 1.44. Actors are thus more likely to form network ties with similar than dissimilar others.

Figure 5.4 shows that networks are most segregated when SI and CS are present, ranging from segregation indices 2 to 8. This means that cooperators, neutral actors, and defectors are two to eight times more likely to be connected to similar than dissimilar others. Even so, Figure 5.3 explicates that cooperation is generally not hampered in segregated networks in which CS and SI are dominant mechanisms. Relatedly, the implementation of CS and no-SI, as well as no-CS and SI, generally leads to network segregation by cooperation. SI allows agents to pull network partners towards behaving similarly, whereas CS allows agents to preferentially select similar others, giving rise to locally converged behavior but global segregation patterns (Axelrod, 1997). The SI and no-CS condition shows similar network segregation by cooperation degrees to the CS

5 The simulation code as well as the Appendix C of Chapter 5 is freely accessible online via https://doi.org/10.17605/OSF.IO/RKG7B.

and no-SI condition. A reason for this is that in both conditions, linked agents tend to be alike. SI fosters connected agents to become alike, whereas CS promotes similar others to form ties. Even so, combining insights from Figure 5.3 and Figure 5.4, we see a tendency of cooperation dying out in networks wherein actors are indiscriminatory in with whom they want to link (no-CS conditions), whereas cooperation keeps increasing in networks in which agents preferentially select similar others (CS-conditions) and which exhibit some degree of network segregation.

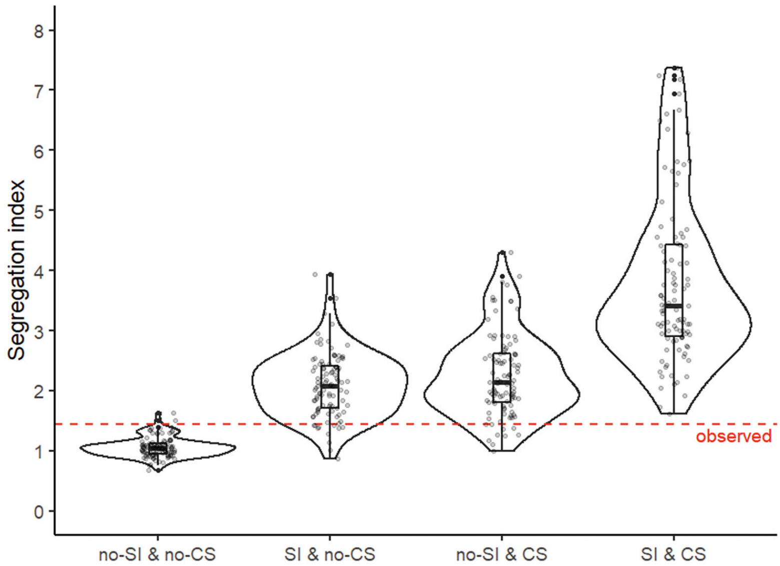


Figure 5.4: Visualizing network segregation by cooperation in waves = 3 and 4 combined. 50 simulation runs per condition are shown in the violin plots, box-plots, and via jittered data points. The dashed red line is the segregation level (1.44) at wave 2 found empirically. Note: SI = social influence; CS = cooperation selection.

To illustrate network segregation by cooperation, we present four networks in Figure 5.5. We show the empirical network with cooperation scores colored per node (white = defector, blue = neutral, and black = cooperator) at waves 1 and 2. The same coloring scheme applies to the two simulated networks at waves 3 and 4. The simulated networks build on a decision-making model in which both CS and SI are present and are based on output from a single simulation run. Figure 5.5 shows an onion-layered structure of the network with cooperators in the center and defectors hovering around them. This reflects findings reported in Table 5.3, showing that cooperators have many ties to other cooperators, whereas defectors tend to form ties to dissimilar types. Especially, cooperators tend to benefit from CS and SI given the many ties among cooperators, whereas defectors generally have fewer ties. Although the mere presence of segregation is not problematic for all, some defectors fall, so to say, through the network (Stadtfeld et al., 2020). Figure 5.5 is a visual corroboration of the findings reported in Table 5.3, showing how cooperators form dense local clusters, whereas defectors do not. The question is whether network segregation patterns also give rise to cooperation segregation in all simulation conditions. We set out to answer this question in the next section.

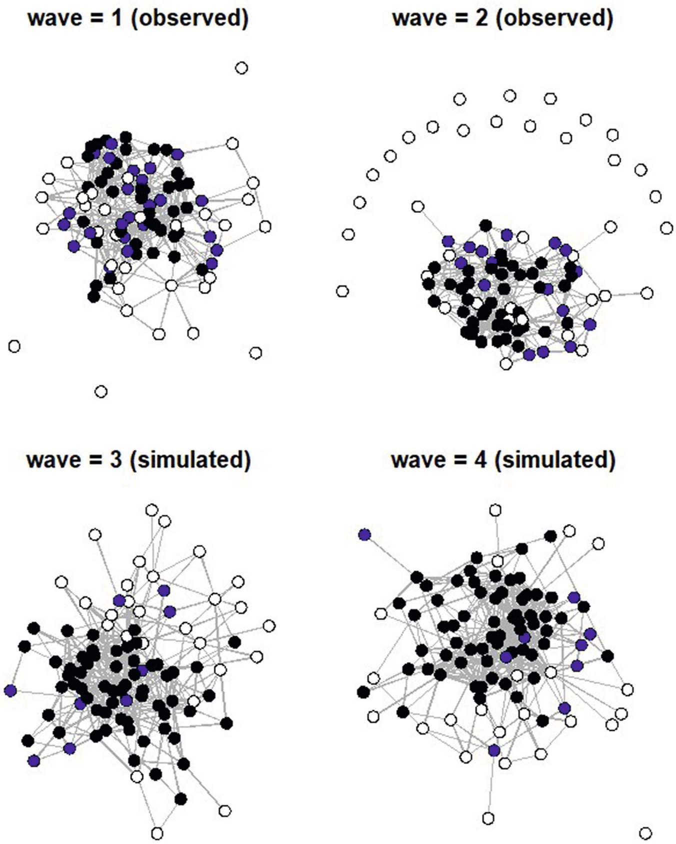


Figure 5.5: Visualizing four networks of 95 agents. White nodes are defectors, blue are neutral ones, and black nodes are cooperators. The simulated networks are generated with CS set = 0.66 and SI = 3.54. Segregation increased from 1.86 (wave 1), 2.13 (wave 2), 3.99 (wave 3), to 4.36 (wave 4). Average level of cooperation increased from 2.07 (wave 1) to 2.42 (wave 4).

5.4.2.2 Impact of the behavioral shape effects on cooperation

A core feature affecting cooperation comprises behavioral shape effects. Especially the quadratic shape effect gives rise to the bimodal distribution in cooperation; that is, cooperation polarizes. How each of the two effects influences the distribution of cooperation cannot be understood in isolation from the other effect. For example, higher behavioral scale values become more likely if the linear effect is strongly positive and the quadratic shape effect is as well. In our case, the quadratic shape effect is significantly positive, and the linear shape effect is almost null. This induced a bimodal shape of the utility function – albeit agents derive slightly more utility from moving to a positive value (cooperation instead of defection) due to the minor positive linear shape effect.

The quadratic shape effect has the most profound effect on changes in cooperation, according to the relative importance indices reported in Table 5.1. Figure 5.5 is a testament to the importance of quadratic shape effects. Percentage-wise, there are 39.0% defecting types, 16.8% neutral, and 44.2% cooperating types at wave 2. This bimodal distribution remains present at wave 3 (30.5% defecting, 18.9% neutral, and 50.5% cooperating). Yet, the bimodal picture of cooperation becomes starker in wave 4, in which there are only a few neutral types left (7.4%), alongside many defectors (31.6%) and cooperating actors (61.1%). The group of defectors remains sizeable, while neutral agents become rare. The problem is that defectors and cooperators are roughly 5 times more likely to connect to similar others, thereby providing defectors little opportunities – either via CS or SI – to change in behavior. Thus, the shape effects are key for cooperation rates.

Empirically, on average, cooperation hovers around 2.00 at wave 1 and wave 2. Percentage-wise, we find that 39.0% is defecting, 16.8% is neutral, and 44.2% is cooperating in wave 1. There are thus slightly more cooperators than defectors initially. We see a slight decline in cooperation empirically because the percentage of defectors increased more from wave 1 to wave 2 than the share of cooperators. Namely, in wave 1, we have $n = 32$ defecting types, $n = 24$ neutral, and 39 cooperating types. At wave 2, we have more defecting ones ($n = 37$), fewer neutral individuals ($n = 16$), and $n = 42$ cooperating ones. Especially the quadratic shape effect fuels polarization in cooperation.

In the CS and no-SI condition, CS assumes that cooperators tend to select cooperators, neutral actors select similar neutral ones, and defecting actors tend to select other defectors. CS ensures that new ties are based on cooperativeness scores. Without SI, cooperation is only affected by shape effects and with whom an agent is locally connected. Cooperation tends to disperse in a bimodal fashion. Most agents are thus either defecting or cooperating, and only a few are neutral. In wave 3, we have 45% defectors, 19% neutral actors, and 36% cooperators. We see a minor shift to wave 4: 44% defectors, 15% neutral actors, and 41% cooperators. The share of neutral actors decreases, ensuring a more polarized distribution of cooperation at wave 4. That is why simulated networks without SI do not exhibit a major linear increase or decrease in the average level of cooperation but rather show polarization of cooperation. Indeed, cooperation tends to polarize if CS is included as an active mechanism. In the network generated by no-CS & SI, we see less segregation by cooperation because network selection mechanisms (reciprocity and transitivity) are not driven by cooperation or defection considerations. Defectors are thus able to form ties to cooperators and, thus, influence others to follow suit. Although the model in which agents are solely driven by CS processes runs the risk of "writing off" clusters of actors who are not, by definition, incorrigible defectors, this may thus eventually happen in networks in which actors can only engage in CS and no-SI.

The simulations reveal in Figure 5.3 that, in wave 3, most agents initially cooperate in the SI and no-CS condition. Notably, the proportion of cooperators increased from 44% to 74% for wave 2 to wave 3. This trend toward all-out cooperation did not continue. Namely, the proportion of cooperators decreased from wave 3 (74%) to wave 4 (48%) in the SI condition. Intuitively, the reason for the current finding – first an increase and then a decrease in the proportion of cooperators – is the combination of other network selection mechanisms and behavioral shape effects. To start, there are many isolates in the empirical network at wave 2, and those isolates are all defecting (see, for example, Figure 5.5). A potential consequence of isolated defectors forming relations with well-connected cooperators via reciprocity or transitivity considerations – combined with SI– is the push towards more cooperation due to SI in those relations. SI operates if actor *i* nominates actor *j*. Then *j*, in turn, influences *i*. Defectors tend to send network nominations to other types (see D-*other* in Table 5.3), enabling SI to occur from cooperators to whom defectors are linked. Cooperators are the *j*, and defectors are the *i* in this SI story. We see such a spreading-of-cooperation dynamic happening from wave 2 to wave 3 in conditions including SI.

However, we also see a change in cooperation from wave 3 to wave 4. Why? There are two reasons. First, reciprocity and transitivity still link cooperators to defectors (see C-*other* in Table 5.3). This makes cooperators still susceptible to SI from defecting counterparts. This time, defectors are the *j*, and cooperators are the *i* in this SI story. Second, defectors, neutrals, and cooperators also tend to change behavior based on shape and SI effects. From wave 3 to wave 4, well-connected defectors (see D-*other* in Table 5.3) can socially influence their cooperative counterparts to follow suit. Given that a change from cooperation to defection is more appealing than a change from cooperation to neutrality due to the quadratic shape effect, we see a shift in cooperation rates. Eyeballing a single run corroborates this intuition (Figure C6 in Appendix C).

5

In the SI and CS conditions, we see an increase in the proportion of cooperators from 52% to 64%, respectively, for waves 3 to 4. The drop of neutral actors drives the increase in cooperation: From 23% to 14% in waves 3 and 4, respectively. Defectors comprise a sizeable amount: 25% (wave 3) and 22% (wave 4). The combination of CS and SI creates a situation in the current network where cooperation can flourish over time. Defectors are mostly isolated in wave 2, unable to influence others to defect. Cooperators, conversely, are generally well-connected. Cooperators tend to sever ties to defectors due to CS, which means that cooperators are less exposed to SI from defectors, further stabilizing cooperation. As such, CS and SI allow cooperators to maintain their relations with similar others while simultaneously ensuring via SI that the target of influence keeps cooperating. The pool for defectors to preferentially form a relationship with a similar defector decreases over time. Yet, this process is not deterministic: Actors may deviate from prior behavior or alter the local network structure via shape effects. A defector (or cooperator) still optimizes myopically actions over time due to the linear and quadratic shape effects.

5.4.2.3 Altering initial cooperation

To test our intuitions regarding the importance of local connections in the subnetwork, global levels of network segregation by cooperation, and the importance of shape effects, we test to which degree the results reported in Figure 5.3 are conditional on the empirical situation of an initial, dense cluster of cooperators as well as the level of initial cooperation. The empirical network has a relatively large group of cooperators, whereas defectors tend to fall through the network. We study whether we find similar results for defectors if the tables are turned. To this end, we artificially alter the data: Actors who are cooperators in the empirical situation of wave 2 become defectors, defectors become cooperators, and neutral actors remain similar. A consequence of swapping cooperation for defection scores is that defectors now have more ties in the initial situation and are part of more dense local subnetworks. We test the tentative explanations purported in section 5.4.2. In what follows, we discuss the outcomes of 100 simulation runs per condition – no-SI and no-CS, no-CS and SI, CS and no-SI, and SI and CS – using the decision-making model presented in Table 5.2.

Results of the simulations with reversed input data are shown in Figures 5.6 and 5.7 and Table 5.4. Akin to Figure 5.3, Figure 5.6 shows that the no-SI and no-CS and the no-SI and CS conditions tend to favor the minority group – the minority being either defectors or cooperators – in the network. In Figure 5.3, defectors are in the minority, whereas cooperators are in the minority in Figure 5.6. Figure 5.6 supports our earlier finding that the SI and CS condition promotes the behavior that the majority adheres to. Also, the SI and no-CS simulation condition shows that SI first furthers the behavior of the majority but then promotes the spreading of the other behavior – in the case of Figure 5.6: Cooperators can still influence others to behave differently if CS is absent. Numerically, the proportions of cooperators, defectors, and neutrals remain largely similar, as reported in the previous section. The only exception is that due to the larger share of defectors initially, we see that proportions tend to favor defectors.

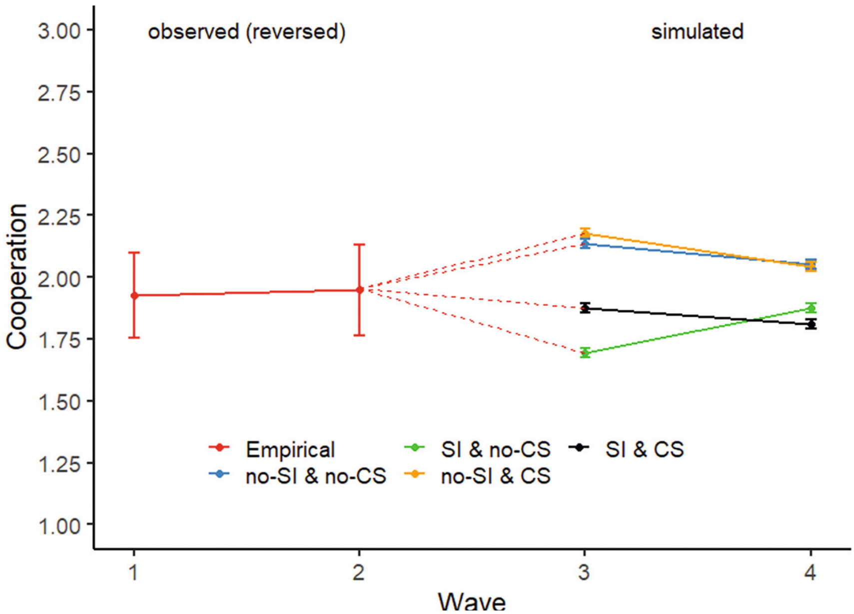


Figure 5.6: Visualizing the mean and 95% confidence intervals of altered cooperation data generated via 100 simulation runs per proposition. The cooperatoin distribution of the emprical data is reversed and included in wave 1 and wave 2 (red). Dashed red lines are inserted to link the empirical data at wave 2 to the starting point of the simulations in wave 3. Note: SI = social influence; CS = cooperation selection.

Table 5.4 shows local sub-networks of cooperator-cooperator, defector-defector, and dissimilar pairs (average of waves 3 and 4 combined), but this time for the reversed input data. Our findings stress the importance of the input data for future outcomes. Defector-defectors nominations are disproportionately more likely to occur than cooperator-cooperator pairs (opposite of findings in Table 5.3). Conditions without CS show that defection does not spread because dissimilar agents are still connected and able to influence each other as well as to form ties based on reciprocity and transitivity (see *D*-other and *C*-other columns in Table 5.4). Simulation conditions with CS and SI allow defectors to connect to similar others preferentially. For instance, on average, defectors send 4.7 nominations to defectors and 3.0 to dissimilar others. Cooperators tend to maintain at least one nomination to similar others. Present *D*-other and *C*-other relations in the SI and CS condition allow dissimilar types to influence each other to behave differently. The story of Table 5.4 is fairly similar to the one told before in Table 5.3, with the sole exception that reversing the data resulted in defectors being now the well-connected types in the network.

Table 5.4: Mean number of network ties across waves of cooperators and defectors in pairs with similar and dissimilar others using reversed input cooperation data.

Condition	C-C pair	D-D pair	C-*other* pair	D-*other* pair
Empirical (modified)	1.26	2.94	4.29	2.61
no-SI & no-CS	1.81	2.04	2.30	2.25
SI & no-CS	0.94	2.97	2.07	2.57
no-SI & CS	1.90	2.37	2.52	2.58
SI & CS	0.88	4.74	3.45	2.96

Note. C = cooperator; D = defector; SI = social influence effect; CS = cooperation selection effect; other = agent of other types.

To visually show network segregation by defection, we present four networks in Figure 5.7. The simulated networks are based on one simulation run. Figure C7 in Appendix C shows network segregation by cooperation of all the simulation runs. As suggested in Table 5.4, Figure 5.7 shows that defectors have many ties to other defectors, whereas cooperators squander around with fewer ties in less dense parts of the network.

5.4.3 Extending our simulation scenarios

We inspect two additional scenarios: *i*) negative SI and CS estimates and *ii*) cooperation selection modeled via a cooperation popularity effect rather than a cooperation similarity effect as in CS. We inspect the consequences for cooperation levels; that is, does cooperation spread, die out, or segregate in the simulated networks?

First, a negative CS parameter points to heterophily, i.e., the tendency to select dissimilar others as friends preferentially and thus to distance oneself from similar others. Correspondingly, negative SI comprises distancing from peers' cooperativeness, for instance, when someone strives to dissent from the norm among one's peers or rebel against the prevailing norm in the peer group. Chapter 3 shows theoretically that cooperators switch to defection because of social influence received from their defecting counterparts. This may fuel unhappiness with the situation at hand because one may be innately motivated to cooperate but forced to defect. A potential consequence is that some agents may want to rebel against the now prevailing unwanted norm that one privately opposes. One way to sway others to cooperate could be to behave oppositely, hoping to set off a cascade of cooperation. The opposite also holds: Cooperators could randomly explore whether defection sits well with the maintaining cooperators around them, exploiting the benefits of others cooperating. This captures the tendency of "they cooperate, I defect; they defect, I cooperate." We set CS to –0.66 and SI to –3.54.

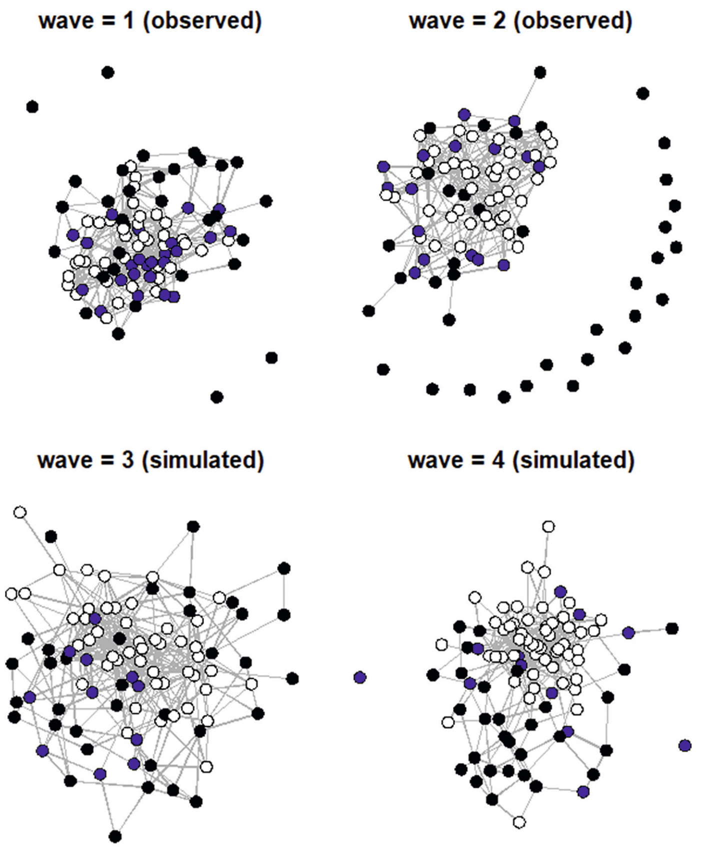


Figure 5.7: Visualizing four networks of 95 agents with input cooperation data in wave 1 and wave 2 reversed. White nodes are defectors, blue are neutral ones, and black nodes are cooperators. The simulated networks are generated with CS set = 0.66 and SI = 3.54. Segregation increased from 1.86 (wave 1), 2.13 (wave 2), 3.15 (wave 3), to 6.83 (wave 4). Average level of cooperation decreased from 1.93 (wave 1) to 1.74 (wave 4).

Second, CS resembles homophily, but what would be the effect if everyone, including defectors, were striving to connect to the "best" possible and most cooperative partner? Such a behavioral mechanism would better reflect the assumption that everyone benefits from having a relationship with a cooperative other, making cooperative others a popular and competed-for target for forming friendship relations. Cooperation popularity (CP) is the effect of interest in this scenario. To assess CP accordingly, we include *altX* in the SAOM and run the model using empirical data. *altX* models whether cooperating actors are more likely to be selected for friendship nominations than neutral or defecting agents. The CP estimate is 0.02 (SE = 0.08, p = 0.84). We set CP thrice the empirical estimate: 0.06. To maintain the model as simple as possible, we keep the other effects listed in Table 5.2 the same.[6] The only difference is that we swap CP for CS.

5.4.3.1 Negative CS and SI estimates

For the additional negative estimate scenarios, we run 50 independent simulations per condition, all other things being equal. This means that we have a total of 250 simulation runs. We implement all potential comparisons: (*i*) negative-CS and negative-SI, (*ii*) negative-CS and positive-SI, (*iii*) negative-CS and no-SI, (*iv*) negative-SI and positive-CS, and (*v*) negative-SI and no-CS. Results are visualized in Appendix C, Figure C1, showing that negative mechanisms are, as expected, forces for cooperation levels and network segregation to reckon with. Heterophily and social distancing generally dampen cooperation. Figure C1 in Appendix C shows that cooperation levels are near 2 across all model comparisons, elucidating that agents are largely indifferent between choosing cooperation over defection and defection over cooperation. Figure C2 in Appendix C shows largely integrated networks. It is not a surprise that preferring to connect to dissimilar others and distancing oneself from similar others in cooperation increases integration. Heterophily and social distancing are innately destined to foster ties among dissimilar agents.

5.4.3.2 Cooperation popularity as an alternative estimate for CS

To explore combinations of CP vs. CS with and without SI, we run 50 independent simulations per condition (a total of 100 simulation runs). Results are visualized in Figures C3 and C4 in Appendix C.

First, CP without SI shows that cooperation on average increased over time: From 1.82 (95% CI [1.80, 1.85]) in wave 3 to 2.04 (95% CI [2.02, 2.07]) in wave 4. Cooperation is still lower than the effect found empirically at wave 2 (2.05) and under model conditions with CS and SI (2.41). This model configuration produces a bimodal distribution in wave 3: Defectors = 48%; neutral = 21%; cooperators = 31%. For wave 4, the picture is similar except that there are more cooperators and fewer neutral ones: Defectors = 42%; neutral = 12%; cooperators = 46%. Next, most simulations report a segregation index near 1 (also see Figure C2 in Appendix C), meaning that defec-

6 Estimates in models with CP or CS are largely the same (Table 5.2 estimates in brackets): Rate selection = 12.76 (12.75), outdegree = −2.50 (−2.51), reciprocity = 1.88 (1.88), transitivity = 0.24 (0.24), rate influence = 3.11 (3.18), linear shape = 0.04 (0.05), quadratic shape = 1.57 (1.55), SI = 1.21 (1.18).

tors, neutrals, and cooperators are generally indifferent to with whom they form network ties. Mean outdegree scores suggest that cooperator-cooperator nominations (3.14) are likelier than defector-defector nominations (0.78). Defectors and cooperators nominate dissimilar others: 2.09 and 2.58 on average. Thus, striving to connect to the best cooperative partners reduces cooperation on average the most compared to all model configurations incorporating CS, SI, or CS and SI (Figure 5.4). Behavior is primarily influenced by the quadratic shape effect, providing agents the highest utility on the outer ends of the behavioral scale; agents derive the most utility either as a defector or cooperator.

Second, including SI next to CP paints a different picture of cooperation and segregation than the CP and no-SI condition (Figures C3 and C4 in Appendix C). Cooperation, on average, decreased over time, from 2.47 (95% CI [2.45, 2.50]) to 2.17 (95% CI [2.15, 2.21]), suggesting that all-out cooperation is more readily achieved in wave 3 (proportion types in wave 3: Defectors = 19%; neutral = 15%; cooperators = 66%) but reversed its course in wave 4: Defectors = 35%; neutral = 13%; cooperators = 53%. We find, furthermore, that cooperators are more active in sending out ties (C-C mean outdegree = 3.51, C-other mean outdegree = 2.71). Defectors, conversely, are less active and less inclined to nominate similar others (mean outdegree = 0.77). Akin to previous D-other pair outdegrees, defectors are more likely to select dissimilar others for network relations (mean outdegree = 2.40). The CP and SI condition does not lead to higher cooperation levels than the CS and SI condition (mean cooperation = 2.41). CP implies that many want to form a relationship with cooperators.

The expectation beforehand was that including CP might lead to higher cooperation levels, but we do not see that. Why? One possible explanation is that CP, together with reciprocity, transitivity, sizeable C-*other* and D-*other* relations from the outset, and a relatively fast pace of network changes relative to behavioral changes, entails that the strong tendency to nominate cooperators leads to the subsequent emergence of nominations from cooperators to defectors because reciprocity and transitivity effects override the aversion cooperators have to befriend defectors that follow from CP. Consequently, despite CP, there is still social influence from defectors on cooperators, which may explain the decline of cooperation observed in the counterfactual from wave 3 to wave 4 for the CP & SI condition. In other words, despite CP, cooperators cannot insulate themselves from social pressures from dissimilar others due to multiple operation network selection mechanisms.

5.5 CONCLUDING REMARKS

Although recent studies suggest that cooperativeness can be a relatively stable individual trait when individuals act in isolation (de Matos Fernandes et al., 2022b; Chapter 2), networks have been found to affect cooperation crucially (Melamed et al., 2020; Simpson & Willer, 2015; Chapter 3). In this chapter, we studied the co-evolution of networks and behavior, extending prior research using empirically calibrated stochastic actor-oriented ABCM that studied network evolution (Snijders & Steglich, 2015; Stadtfeld, 2018). We moved beyond current literature by studying cooperation and networks in a setting based upon an empirically observed situation in many ways but also sufficiently abstract to study the relative contribution of different mechanisms in counterfactual simulation scenarios. The empirical calibration in Chapter 5 incorporates reciprocity and transitivity based on empirically observed network changes, assuming myopic stochastic optimization in modeling changes in both cooperation and the formation of network relations. We focus on the effect of social influence (SI) and cooperation selection (CS), particularly on whether cooperation dies out, thrives, or segregates under more realistic conditions. We find that the parallel occurrence of SI and CS fosters network segregation by cooperation but, more importantly, generates overall higher cooperation levels than in all other conditions we inspected.

Starting from an initial majority of cooperators (based on the empirical setting), peers socially influence each other to adopt the behavior that prevails in their peer networks, leading in our simulations to an initial spreading of cooperation. However, cooperation would not be maintained in this situation if cooperators remain socially connected to defectors because, otherwise, these defectors will influence cooperators to reduce their cooperation levels – as we observed in simulations in which social influence was not accompanied by cooperation selection. It is here where cooperation selection becomes crucial by allowing cooperators to sever their ties with defectors, insulating them from defectors' social influence and thus stabilizing increasing cooperation levels in the population. An important underlying feature of these findings is that cooperators are more active in forming network relations and can form relations with similar others than defectors. We also observed a downside to this successful combination of SI and CS: Many defectors end up in segregated networks insulated from the social influence of cooperators and may thus refrain from cooperating altogether. If the defectors had maintained their social ties to cooperators somewhat longer (without "infecting" them with defection), a higher level of cooperation could have been achieved in the population.

There are limitations to our empirically calibrated model. To stay as close as possible to the empirical setting, we only altered CS and SI in simulating two additional waves. We chose three times the estimates found empirically to explicate the impact of CS and SI. This choice is based on prior research that utilizes either twice or thrice the effects found empirically (Schaefer, adams, & Haas, 2013; Steglich, 2018). One may question whether altering the parameter strength still resembles the empirical situation, whereas a question remains whether CS and SI in our counterfactual scenarios represent how strong these processes are in experimental research and theoretical models.

Another limitation pertains to the limited set of *RSiena* effects that we included in our SAOM. Most studies using *RSiena* include more effects to capture selection and influence mechanisms more comprehensively (Steglich et al., 2010). We only zoomed in on a few *RSiena* estimates to explicate the role of reciprocity, transitivity, and myopic stochastic optimization on CS, SI, and cooperation levels as clearly as possible. This approach of including a limited set of estimates in a model is common (Stadtfeld, 2018), but it remains an open question to what extent our results would be robust if we had based our counterfactuals on a richer specification of the empirical model. Relatedly, we could not directly model the assumption that defectors are deterred from continuing their defection by exclusion. Whether agents are more or less motivated to cooperate via (the threat of) exclusion is input for future work. Intuitively, the corresponding effect may represent a tendency like "the fewer incoming ties a defector has from cooperators, the more a defector should be inclined to cooperate" or conversely, "cooperation is more attractive, the more friendship nominations one receives." A prerequisite to include such an effect is that agents need to be able to learn which behavior leads to a satisfactory outcome and which behavior does not. To our knowledge, *RSiena* does not offer such an implementation.

Furthermore, a limitation of this chapter pertains to the non-significant effects of CS and SI in the empirical SAOM, suggesting that CS and SI are not dominant mechanisms in the empirical setting on which our simulations are based. However, we believe that this may make our study even more relevant because we could inspect how cooperation rates could be best fostered if interventions succeeded in increasing the strength of CS or SI or both relative to the empirically observed scenario. Furthermore, the relational and behavioral assumptions inherent to the SAOM framework are not necessarily theoretically or empirically the most convincing choices but must be relied on when applying the framework. This limits the scope of our ABCM. We could not deviate too much from the *RSiena* framework, as it was not our intention.

Also, embeddedness in multiple contexts is a complicating feature to promote cooperation (Bakker, 2019). Not cooperating in the workplace but actively volunteering means one is viewed as a defector in one context while being a cooperator in another. In one context, purported defectors may cooperate at a football club or with family members or friends. Promoting cooperation in one social context may backfire in another context. We leave including multiple, potentially competing contexts for future work.

As always, we must be aware of model artifacts if we alter too much in an ABCM. Additional analyses we reported revealed, for instance, that cooperation and connectivity outcomes from the empirically-calibrated simulations depend on the initial input data. We reversed the input data artificially, turning defectors into cooperators and cooperators into defectors in the initial situation. Network positions did not change. Well-connected defectors are more likely to keep defection ongoing due to having more ties and having more ties to similar others than cooperators. This allowed us to underpin our explanation for the results we observed in the counterfactual simulations, but it also showed that these results could not be easily generalized to empirical settings without considering possible differences in the initial situations between empirical and simulated scenarios.

Our analyses showed that it is essential to fully grasp the underlying mechanisms driving the results in a simulation study and test tentative explanations with suitably designed simulation experiments (Flache & de Matos Fernandes, 2021). Otherwise, it is impossible to fully understand how and why the ABCM produces what it produces. That is why empirically calibrated SAOMs build on three ways to ensure that we – as modelers – understand the underlying mechanics of the model and, more importantly, understand what the model is generating. First, some studies compare the impact of different decision models on behavioral and network outcomes (adams & Schaefer, 2016; Kretschmer & Leszczensky, 2022; Schaefer et al., 2013). Chapter 5 falls into this category. Second, some studies go a step further and re-calibrate the SAOMs based on the manually altered effects (Snijders & Steglich, 2015; Steglich et al., 2010). For example, reciprocity and outdegree parameters are then re-fitted to accommodate the impact of the "new" CS effect. In this way, one tries to stay near the empirical data and inspect under which conditions we find simulated outcomes resembling those observed empirically. We explicitly aimed to inspect alternative scenarios, taking the empirical case as starting point for our simulations. Third, one can keep the decision model as is and only alter the input network or behavioral data (Stadtfeld, 2018). For instance, we could have employed our empirically estimated model on a simulated network data set with 500 agents. One could pursue this road if it is expected that increasing the network size has a dominant effect on the mechanism or tentative explanation of the simulation results.

We opted for the first approach, given that the latter two cannot build on the empirical setting at hand while simultaneously exploring "what if" situations using different values for SI and CS. We use the empirical network as starting point and question, "What happens to cooperation if CS and/or SI are turned off and/or on while behavioral and social regularities estimated empirically are present?" The empirically calibrated part of this chapter is that we derived estimates of behavioral and relational choice processes from a student friendship network. The goal of Chapter 5 was to explore counterfactuals rather than to mimic the effects found empirically in re-calibrating the model to find similar distributions of cooperation or a similar network configuration. We chose not to rely on an initially simulated network to maintain relatedness to the empirical setting at hand. Adjusting both the decision-making model and input network or cooperation configurations poses a challenge to understanding what feature is driving the effects. One can then question whether the input network, input cooperation distribution, CS and SI presence or absence, or combinations of the previously mentioned conditions drive the findings. Although we implemented additional simulation scenarios, we leave other potential model adjustments for future work. In the current way, we keep track of changes in the decision-making model and what were our outcomes.

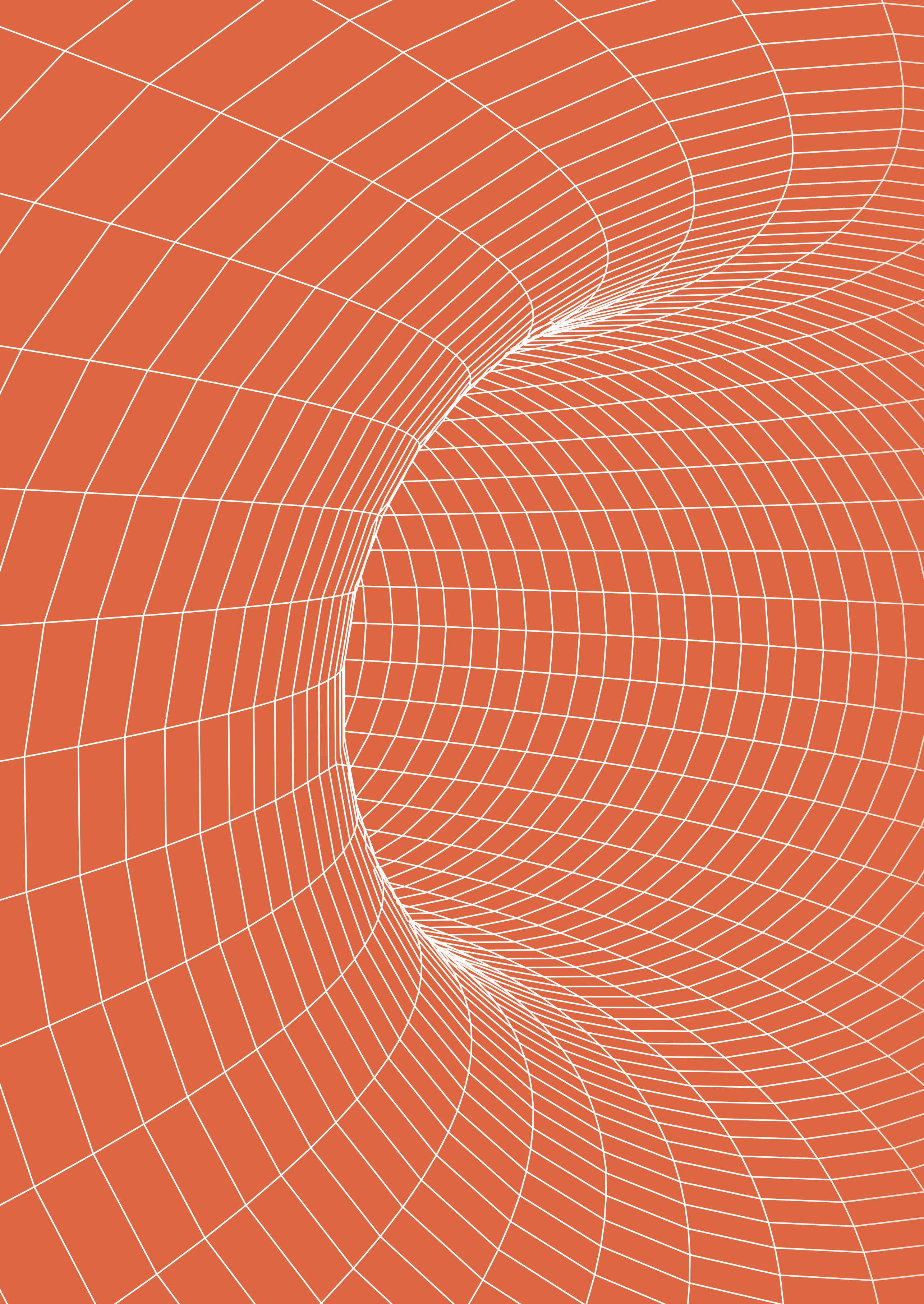
A potential extension left untouched for now for the reasons listed above relates to the rate effects of network selection and the behavioral evolution functions. The rate function models the speed of changes – "how many opportunities an agent gets to change" – in network relations and cooperation. In our model, actors have, on average, roughly 13 opportunities in-between waves to change network connections. A change comprises severing a connection, forming a new tie, or maintaining a network relation. For cooperation, actors have approximately 3 opportunities to change in behavior. A change in behavior entails sticking to the current shown behavior, adopting a higher score (for defectors and neutrals), or scoring lower (for cooperators and neutrals). A potential consequence of the difference in speed between selection and influence is that network changes occur irrespective of cooperation and/or defection. A tie nomination may already be reciprocated before a partner changes behavior, perhaps in an unsatisfactory manner. This may lead to SI operating in heterogeneous local subnetworks because tie changes are not driven by cooperation similarity in the no-CS condition. Alternatively, the consequences of CS and SI may be more quickly noticeable in the network because agents have more opportunities to pursue similar others than changing their behavior due to SI. If agents quickly ensure a local subnetwork with similar others, then SI may thereafter ensure behavioral homogeneity in the newly configured local subnetwork. Even so, prior research notes that increasing network speed falls flat as a driving feature for network segregation due to propelling effects of stronger homophily on network tie changes (Steglich, 2018). Increasing the speed of network formation is less important for increasing than strong homophily effects. This suggests that network changes are more sensitive to input from the objective than the rate function. Follow-up studies could dive deeper into the role of altering the rate function in selection and influence SAOMs that are empirically calibrated.

Another prospect for future research relates to including a richer model of homophily to study whether homophily based on other attributes than cooperation interferes with the effects of SI and CS we observed. Notably, homophily on individual features is another dominant process that leads social actors to seek, for example, same-gender or same-attitude others as network partners (McPherson et al., 2001). Moreover, suppose some of these processes have led to network relations between cooperators. In that case, relational norms of friendship may override the effects of SI, leading cooperators to accept their partners' defection while maintaining the social relation simultaneously (Flache & Macy, 1996; Flache, 2002). Attesting to previous work on multi-dimensional homophily stressing different dimensions – e.g., gender, behavior, or attitudes – in which individuals can be similar (Block & Grund, 2014; Hooijsma et al., 2020; McPherson et al., 2011), we propose that an empirically-calibrated SAOM can be used to further our understanding of the interrelatedness of features such as gender, opinions, or ethnicity on the micro-level and network segregation on the macro-level.

Following the method of decreasing abstraction, the tentative "thought experiments" in the introduction of this chapter represent a simple cooperation model. We started with a simple, abstract representation of the problem and gradually incorporated more complexity, details, and realism. By starting with a simpler model, researchers can better understand the basic mechanisms underlying the problem and avoid creating overly complex models that are difficult to interpret or analyze. This is already mostly done in prior research (see, for example, the extensive overview of Rand & Nowak, 2013). Here, the empirically calibrated SAOM ABCM starts relatively complex, leaving simpler ABCMs for the future. Our model is more detailed and incorporates more realistic features compared to prior cooperation models studying CS and SI, including features affecting individual-level behavior and other contextual factors. We tested how these factors affect the overall dynamics of the system. One could have started simpler with a SAOM, including only CS or SI. Yet, such a model still builds on assumptions such as myopic stochasticity and basic network features affecting network formation. Here, we tried to show that an empirically calibrated ABCM is a valuable approach for studying complex systems in which cooperation and network relationships are interrelated. More importantly, we show that including real-life features introduces interfering features that prevent and promote the effectiveness of CS and SI, otherwise remaining unknown in simpler models.

The approach taken in Chapter 5 builds on the macro-micro-macro, social complexity, and social mechanism approach (Coleman, 1990; Hedström & Ylikoski, 2010; Mäs, 2021; Chapter 1). Two components, in particular, stand out. The first component is the dependence of macro-level network configurations, such as segregation or polarization, on micro-level decision-making. In particular, in socially complex systems, it is typically misleading to assume that there is a straightforward aggregation of the actions and intentions of individuals to the resulting macro-level outcomes. Well-known examples are residential segregation (Schelling, 1978), opinion polarization (Flache et al., 2017b), and gender segregation (Stadtfeld, 2018). Indeed, we theoretically explored the implications of different assumptions concerning CS and SI preferences on the micro-level for cooperation levels and network configurations on the macro-level and found several counter-intuitive outcomes, such as the result that cooperation rates can decline if the strength of social influence is increased, even when cooperators are initially in the majority. The second core tenet is that the interplay of multiple social processes and actors at the micro and macro-level generates outcomes that cannot be understood by analyzing each process separately. Chapter 5 showed how network segregation by cooperation depends on the specific configuration of network relations, local clustering, cooperation, and CS and SI as social mechanisms.

To conclude, we advanced the literature by controlling for the parallel occurrence of influence and selection and accommodating the "messiness" of real life in explaining conditions under which defectors readily cooperate in networks. We learned from Chapter 5 that understanding the interplay of CS and SI improves our understanding of cooperation in a real-life network context with multiple interfering behavioral and social mechanisms. Cooperation depends on the relative strength of SI and CS, the initial prevalence of cooperation, network structure, the strength of other behavioral and social mechanisms, and the empirical context of whether cooperation spreads, dies out, or segregates. The effectiveness of SI and CS in promoting cooperation may be less clear if we account for empirical features such as reciprocity, transitivity, and myopic stochasticity. The benefit of the current model lies in maintaining relatedness to the empirical context while simulating alternative conditions that apply to the context under consideration. We believe to have demonstrated that the common "as is" empirical approach can be extended with a "what if" analytical and model-based procedure.



Chapter 6

There is no I in TEAM, but there is a M-E in there[1]

Guys were complaining to me and said,
'Shaq, Kobe is not passing the ball.'
I said: 'I'll talk to him.'
I said: 'Kobe, there's no I in team.'
And Kobe said, 'I know.
But there's an M-E in that m***f*****.'**

—An excerpt of the speech Shaquille O'Neal (former basketball player) gave during *A Celebration of Life for Kobe and Gianna Bryant* on February 24th, 2020, dedicated to honoring the late basketball player Kobe Bryant.

1) This chapter is based on joint work with **Marion Hoffman** and **Jasperina Brouwer**, which is currently under review at a peer-reviewed journal under the working title "*Antecedents of student team formation in higher education*".

ABSTRACT

Chapter 6 explores the role of exclusion in team formation, where joining one team instead of another can separate individuals. First, Chapter 6 draws on the theoretical analysis in Chapter 3 by providing an empirical examination of the factors that influence team formation based on cooperative considerations. Second, Chapter 6 utilizes the preference-for-collaboration (PFC) measure from Chapter 4 and applies it as a popularity index for each student in Chapter 6. As such, Chapter 6 investigates whether individuals who are more reputable as cooperative type are more likely to form teams with each other or whether other factors such as competence, gender, or friendship relationships play a more significant role in team formation. To achieve this, I draw in Chapter 6 on empirical data from a sample of students required to form project teams. In doing so, this chapter contributes to a better understanding of what shapes teams, using the novel tool of Exponential Random Partition Models (ERPMs) tailored for studying the antecedents of team relations.

6.1 INTRODUCTION

The power of teams in generating valuable output is well-known. Yet, while prior research shows that the composition of a team is critical for its efficiency (Bell, 2007; Mathieu et al., 2008; Chapter 3), research on how and why specific team configurations come about – notwithstanding recent advances (Bailey & Skvoretz, 2017; Kaven et al., 2021) – is scant. Understanding what drives team compositions can provide valuable insights to stakeholders, such as managers of healthcare teams (Leggat, 2007), teachers implementing distributed leadership (de Jong et al., 2023), or project team leaders during the COVID-19 pandemic (Tannenbaum et al., 2021). The same holds for forming teams in higher education. In small-group teaching curricula, teachers often ask students to select fellow students to create project teams by themselves rather than impose pre-defined groups (e.g., through a random assignment). Such a self-organizing approach is frequently preferred over random assignments because it can lead to better outcomes, such as achieving the team's goals, fostering pride in the collective result, limiting conflicts, and increasing the levels of students' enjoyment (Chapman et al., 2006).

Students exhibit varying approaches toward team formation (Bailey & Skvoretz, 2017; Bercovitz & Feldman, 2011; Kaven et al., 2021). Some students adopt a result-oriented approach and prioritize forming teams with high-grade or cooperative students. Conversely, others prioritize familiarity and forming teams with students they know, regardless of their academic performance. Students may optimize team formation based on other factors unrelated to academic performance, such as convenience or social signaling. For example, Bailey and Skvoretz (2017) note that positive past experiences, the mere presence of a friendly face, similarity in attributes (e.g., gender, competence, interests, attitudes), and popularity may influence the choice of teammates. Students may choose to team up with friends for reasons beyond academic performance, such as the desire to spend time with them. There are abundant reasons to join project teams, but empirical studies examining the features driving team compositions are rare.

In this study, we aim to investigate the mechanisms of team formation in a higher education context. More specifically, we examine project teams in a cohort of first-year bachelor students. The data encompasses 70 students asked to self-form teams to carry out a semester project. Chapter 6 dives into three underlying antecedents of project team formation: friendships, familiarity with fellow team members due to previous teamwork and interactions, and similarity of team members regarding gender, grades, and preference-for-collaboration (PFC) popularity. The PFC popularity measure relates to Chapter 4, in which we utilize PFC as a network, allowing us to indicate who prefers to form a PFC relationship with whom. Here, we treat the PFC network as a popularity indicator, stressing that students may want to join a group with more PFC-reputable others.

Because teams are a group-level relation, we utilize a model tailored for group data which can be used to test these determinants together; that is, the Exponential Random Partition Model (ERPM; Hoffman et al., 2023). To our knowledge, this study is the first to employ ERPMs to study team formation in higher education.

We use a multifaceted approach to examine the determinants of team compositions, adding a range of other explanatory variables beyond PFC popularity. *Friendships* are usually formed based on ‘attractiveness’ and personal affect (Byrne, 1971; Verbrugge, 1977). In the context of friendship making, ‘attractiveness’ refers to the degree to which an individuals’, for example, personality traits, behaviors, and appearance are perceived as appealing and desirable by others. The same reasons that lead individuals to find others attractive as friends can explain their choice of teammates (Bercovitz & Feldman, 2011)– we can thus expect that individuals are more likely to form teams with friends. Moreover, *familiarity* from previous teamwork may play a role in the composition of teams. Students have already learned how to work together, which may explain the repetition of working together. Namely, past interactions with potential team members may provide information on others’ work ethics, who they are, and how to work together smoothly (Bailey & Skvoretz, 2017). A positive past interaction may particularly further the odds of repeating a collaboration. Finally, students may form teams based on *similarity* preferences. Homophily (i.e., the tendency to associate with similar others) is a pervasive finding in networks (Blau, 1977; Lazarsfeld & Merton, 1954; McPherson et al., 2001). Individual attributes relevant to similarity preferences in the context of student networks comprise gender, ethnicity, age, religion, socio-economic status, and grades (Brouwer et al., 2022; McPherson et al., 2001; Stadtfeld et al., 2019). We thus expect the role of similarity in gender, grades, and PFC popularity, generally found in empirical networks, to generalize to the team formation context.

Our study uses a novel tool called *exponential random partition modeling* (ERPM; Hoffman et al., 2023). Here, we define a partition as a set of teams within a one-year cohort. This method allows us to study team formation’s individual and social antecedents in the current context instead of relying on correlational studies. Fundamental differences exist between the ERPM and network modeling frameworks, such as the stochastic actor-oriented model (SAOM; Snijders et al., 2010) or the exponential random graph model (ERGM; Lusher et al., 2013) that make the ERPM much better suited for our analyses. First, network models are used to represent dyadic relationships (e.g., “that is my friend”), whereas ERPMs model group-level relationships (e.g., “that is my team”). This group specificity allows us to go beyond simple dyadic mechanisms (e.g., “I prefer someone similar to me”) and model group-level mechanisms (e.g., “I prefer to be in a homogeneous group”). Second, classic network models can be used to model group relationships in the form of two-mode (or bipartite) networks, in which individuals (i.e., nodes of the first mode) can form ties to groups (i.e., nodes of the second mode). However, group nodes, in that case, need to be defined in advance, and individuals are not constrained regarding their membership to different groups (i.e., they may belong to none, one, or several groups at a time). In contrast, the ERPM is tailored for data where groups are not defined in advance and where individuals can be part of one and only one group. The ERPM’s logic thus departs from network models for one-mode or two-mode networks. The tool corresponds much better to the data regarding its applicability to study factors that explain team compositions.

The following section discusses possible features affecting team formation and forms hypotheses. We then describe the data used to test these hypotheses and explain our modeling strategy. We finally present our findings and finish with some concluding remarks regarding the contributions of this chapter.

6.2 THEORETICAL FOUNDATIONS OF TEAM FORMATION

Chapter 6 zooms in on three pieces of the project team formation puzzle: How do friendships, familiarity, and similarity in gender, grades, and PFC popularity affect the compositions of teams? In this section, we reason in what way the three mentioned factors are essential to consider when studying team formation. More colloquially, we invest factors that affect team formation in many ways. Students care about each other (friendships) and may want to help each other get good grades; students want to avoid conflicts in the team; students know each other well so they can more easily work together (maybe they already often study together; familiarity); and students might have similar interests and ways of working (similarity in various attributes).

6.2.1 Friendships: *Teaming up with friends*

Friendships play an essential role in shaping students' academic lives. Friends help each other, provide each other with social support, and share personal information. Dunbar (2018, p. 32) notes that "[f]riends provide moral and emotional support, as well as protection from external threats and the stresses of living in groups, not to mention practical and economic aid when the need arises." Friendships enable smoother interactions (Uzzi, 1997), serve as a motivator to put your best foot forward to achieve the best for your friend (Granovetter, 1985), foster social cohesion (Coleman, 1990), and engage more easily in beneficial exchanges (Homans, 1974). In the student context, friendships are shown to be a source of resources affecting academic performance (Brouwer et al., 2018, 2022; Stadtfeld et al., 2019).

Previous work has shown that friends stick together in self-organized teams (Bailey & Skvoretz, 2017; Kaven et al., 2021). Based on previous arguments, collaborating with friends may be easier than working with students because they are not affectively related to them. Yet, there may be downsides to working with friends who shield each other from social repercussions (Flache & Macy, 1996). For example, if not contributing to a team project is punished (e.g., by being assigned extra work), a friend may offer support or help, possibly diminishing the punishment's effect. Friends may not correct one another if an error is made instead of pursuing the team's best interests. Having friends in a project team may thus have unintended negative consequences for its success along the way. Yet, the potential downsides of working with friends may not be apparent in advance (or may never materialize). In this chapter, we argue the benefits of forming a team with friends should outweigh the possible drawbacks. Therefore, **Hypothesis 1** states that *friends are more likely to be found in the same team*.

6.2.2 Familiarity: *Teaming up with known others*

Familiarity may influence the likelihood of students joining the same team. Here, familiarity means that prospective team members had prior interactions (potentially positive). Familiar

6

students are expected to be aware of prospective teammates' working approaches and could thus be better potential partners as they have already learned to work together (Bailey & Skvoretz, 2017). For familiar team members, the expenses and effort required to communicate and share information may be lower (Bercovitz & Feldman, 2011; Gulati, 1995). Among previous interactions, those perceived as positive should be excellent predictors of renewed collaborations (Bailey & Skvoretz, 2017). Studies have shown that positive prior collaborations are indeed a key determinant to continuing to work together in teams (Kaven et al., 2021; Lungeanu et al., 2014), as well as the formation of new teams (Bailey & Skvoretz, 2017; Bercovitz & Feldman, 2011; Huckman et al., 2009; Lungeanu et al., 2014; Maynard et al., 2019; Muskat et al., 2022).

In this chapter, we use two experiences during the previous semester to define familiarity between students: (*i*) learning communities and (*ii*) coworking experiences in a course. First, at the beginning of the year (5 months before the project course we are examining), students are divided into small learning communities that meet (bi-)weekly for approximately 45 minutes. Learning communities allow students to meet other students regularly (Brouwer et al., 2018, 2022; Smith et al., 2004; Zhao & Kuh, 2004). Learning communities are designed so students stay together for the first semester for all courses, and then they can switch groups within courses. The project course we study starts when learning communities' influence on team formation decreases. Second, the prior project course ran simultaneously with the learning community meetings. Prior project teams were based on the composition of learning communities, albeit in teams of a maximum of three students. Students were then provided intensive guidance by staff to succeed in their projects. Therefore, we expect prior encounters in both settings to be mainly neutral to positive. Following our previous arguments, we argue that the opportunities to get to know one another provided by learning communities and the previous project course should foster the likelihood of students ending up in the same team. **Hypothesis 2** states that *students familiar with each other due to previous experiences are more likely to be found in the same team*.

6.2.3 Similarity: *Teaming up with similar peers*

We explore three attributes inducing homophily in teams: gender, grades, and PFC popularity. The phenomenon of homophily, where individuals tend to form connections with those with similar attributes, has been well documented in the literature (McPherson et al., 2001). Gender is an observable attribute, but even "invisible characteristics" of individuals, such as their attitudes and values, can significantly impact their likelihood of forming a connection (van Duijn et al., 2003). Such invisible traits might be indirectly perceived via observable signals or cues (e.g., language, clothing, cultural activities) to inform others about unreadily observable individual characteristics (Gambetta, 2009).

There are multiple reasons why students have homophilous relations. One potential source of similarity (the outcome of homophily) is social influence (see Chapters 3 and 5), meaning that peers influence each other to behave similarly. For example, students who are friends may influence each other to work hard in their studies to realize high grades. The opposite could also be true: Students may influence others to slack off and focus on non-academic activities (e.g., attending parties instead of studying). In what follows, we merely point to network or group-related mechanisms that foster homophily because our analytical tool builds on selection mechanisms. Moreover, we highlight multiple mechanisms that may give rise to the outcome of peers teaming up with similar others. Yet, with our data and tool, we cannot distinguish which process would drive the result of similarity. We merely provide a theoretical overview of how the similarity-in-teams outcome (if found) may have occurred.

The first potential source is that individuals want to be related to individuals with the same grade or gender (McPherson et al., 2001). Having similar grades or behaving similarly in terms of cooperation can be a signal of shared approaches to higher education, study behaviors, and "who they are." For example, a high grade may convey to others that one is hardworking and serious about succeeding in higher education. Similarly, being popular as a collaborator can be a signal of shared approaches to higher education, study behaviors, and "who they are." Similar others may provide each other with resources – think of receiving social support, sharing social values, or feeling a 'sense of belonging' – which dissimilar actors cannot give or are unwilling to give (as shown, for example, in theoretical work by Bianchi et al., 2020, and Flache & Macy, 1996). This argument applies to all three attributes.

A potential second source of homophily – particularly applicable to grade and PFC popularity homophily – is that although lower-scoring students may want to form relationships with higher-scoring students, higher scorers reject such relationships and primarily interact with similar higher scorers. Lower-scoring students – for example, students with lower grades – are left with the only option to interact with other lower-scoring students (as found theoretically by Bianchi et al., 2020; Flache & Hegselmann, 1999). For example, this tendency occurs in networks where actors must cooperate (Simpson & Willer, 2015). Cooperators reject ties to defectors, forcing defectors to form ties to similar others. They do this because it smoothens working together and protects cooperators from defectors who tend to reap the benefits of their cooperative efforts. We saw a similar tendency in Chapter 3, where the matching mechanism allows cooperators to reject defectors. For non-PFC popular students and those with lower grades may aspire, propose, and join teams with their more successful peers to, for example, learn from them or reap the benefits of their success. Yet, they may be rejected by their 'higher' peers who prefer to interact with similar others.[2]

A third potential source of homophily on a particular attribute is a byproduct of similarity in other dimensions (Hooijsma et al., 2020). The similarity in grades or PFC popularity may arise as

2 Status asymmetry – i.e., pursuing ties with higher-status others, albeit higher achievers or PFC popular students – may interfere with seeking similar others (Snijders & Lomi, 2019). Aspiration plays an important role in the team formation process. Such teams may tend to be more popular in the sense that more students want to be part of such a team. Understanding the relationship between pursuing similarity and aspiration is crucial to comprehend the dynamics of PFC popularity and grades in team formation. We control for this 'aspiration effect', stating that students with higher grades or PFC popularity might end up in larger groups.

a byproduct of, for example, gender homophily. Recent work shows, for instance, that women are more cooperative than men (Höglinger & Wehrli, 2017), while economics students are more likely proselfs than other students (Marwell & Ames, 1981). In other words, personal predispositions toward certain forms of cooperative behavior are more likely to occur for individuals with similar characteristics or socialized in some similar way. Building on the intuition that the tendency to preferentially connect to socio-demographically similar others may lead to similarity on other dimensions, teams may be more homophilous regarding grades and PFC popularity.

6.2.3.1 Forming similarity hypotheses

For gender, the attraction paradigm set forth by Byrne (1971) stresses that seeking similar others is appealing because others are "like me." Indeed, gender similarity has been observed in many contexts and serves as a fundamental organizing principle in forming network relations (McPherson et al., 2001). The same holds for the perseverance of gender homophily in students' friendship networks (Brouwer et al., 2018, 2022; Stadtfeld et al., 2019). Previous research shows that pursuing same-gender others in the team context is a prevalent mechanism (Bailey & Skvoretz, 2017; Kaven et al., 2021), suggesting that being of the same gender breeds a team connection. We study whether this is the case for students and state the following in **Hypothesis 3a**: *Students of the same gender are more likely to be found in the same team*.

Competence in the form of grades is a feature to account for during students' friendships (Brouwer et al., 2022) and team formation (Kaven et al., 2021). Students with similar grades are likelier to be friends (Brouwer et al., 2018, 2022). If the friendship network expresses grade homophily, the question is whether we find the same in teams. Therefore, we hypothesize that preferences for similarity in grades found in friendship networks should also exist in the context of our project teams. **Hypothesis 3b** thus states that s*tudents similar in grades are more likely to be found in the same team.* Intuitively, grade homophily (the same argument applies to PFC popularity homophily) may lead to lower-achieving students facing a double-edged sword: They may have less access to valuable information and qualities in networks and teams if similarity preferences drive assortment in both the friendship network and team context.

The third attribute that can elicit similarity preferences is PFC popularity. Network research shows that actors tend to be similar in cooperation (Rand & Nowak, 2013). The question now is whether this also holds for an empirical case in which students do not only have to form friendship relationships but need to form teams. Chapter 3 showed theoretically that teams tend to be homophilous regarding cooperation. Based on the reasons listed earlier and the findings in Chapter 3, we expect that *students similar in PFC popularity are more likely to be found in the same team* (**Hypothesis 3c**).

6.3 DATA

6.3.1 Higher education context

We analyze data obtained from $N = 70$ first-year bachelor students from a Dutch university. Participation in the data collection process was voluntary. Students answered a 30-minute computer-based questionnaire at the end of the first semester (November-December). In the last weeks of semester 1 (January), students were asked to choose a project team via an online platform. The course for which teams are formed ran in semester 2. We then extracted the compositions of the second-semester project teams from the online registration system. The data used in this chapter stems from the same dataset used in Chapter 4 and 5. The sample size is however smaller because not all students participated in the project teams.

6.3.2 Project team data

Our analyses focus on the composition of the teams formed by the students for a project course during the second semester. The goal of the course was to follow a complete research cycle, from reviewing the literature to collecting data, analyzing the data, and drafting a report. The students formed 12 teams, each focusing on active citizenship in a local municipality as a research topic. Each student selected their team on an online platform where they could see the choices of other students. The team selection had to be made individually on the platform, but students could discuss the composition of teams before the online registration. Although the size of a team was not fixed, there was a cap of a maximum of seven students per team. The final distribution was: two teams with four students, one team with five students, six teams with six students, and three teams with seven students. Each project team was assigned a tutor (teacher) who guided the students through the research process, but the students were responsible for the quality of the report. A single grade was awarded to the whole team.

6.3.3 Variables

6.3.3.1 Friendships

Friendships were collected via online questionnaires in which students rated their fellow students on a scale from 1 (“best friends”) to 6 (“I don’t know who this is”). To investigate whether a friendship relationship exists or not, we dichotomized the variable: 1 = “best friends,” 2 = “friend,” and 3 = “friendly relationships” were coded as 1, whereas 4 = “neutral,” 5 = “only known from face or name” and 6 = “I don’t know who this is” were coded as 0 (following the approach recommended by van de Bunt et al., 1999). A friendship is thus either absent or present. A visualization of the friendship network is provided in Figure 6.1. The density of the network is 0.09. The density statistic is based on the total number of nominations (418) divided by all nominations (70 * 69 = 4830). Students reported, on average, approximately six friendship nominations (418 / 70 = 5.97). Of 418 nominations, 127 are reciprocated relations, resulting in a reciprocity index of 0.61.

6.3.3.2 *Familiarity*

To assess familiarity, we collected two variables: the composition of learning communities and project teams in the first semester. Learning communities were created at the beginning of the first year and carried on until the start of semester 2 (when the project teams under study were formed). There were eight learning communities, with only two students not being part of any learning community. The size of learning communities ranged from 6 to 11 students. In addition, project teams were formed for a course in the first semester, unrelated to the second-semester project course we study here. We refer to these teams as first-semester project teams or prior project teams. Prior project teams comprised 1, 2, or 3 students. Both learning communities and project teams were defined by random assignment rather than self-organization, based on the time students enrolled in the study program.

We created a binary matrix to indicate whether two students were in the same project team and/or learning community. This familiarity matrix is symmetric, meaning that ties in the associated network are undirected. The density of the familiarity matrix is 0.06. The density statistic is based on the total number of undirected relations (273) divided by all possible undirected relationships (70 * 69 = 4830). Students are, on average, familiar with almost four other students (273 / 70 = 3.90).

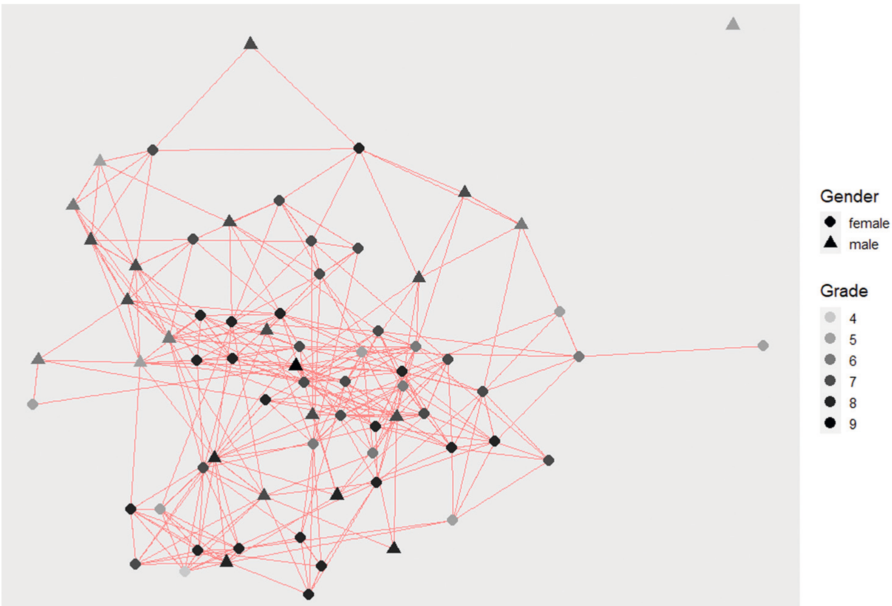


Figure 6.1: The friendship network of students with the shape of nodes set as gender and the darker nodes indicate a higher grade. A light-red line between nodes refers to a present friendship nomination.

6.3.3.3 *Similarity*

We include gender, grades, and PFC popularity as individual attributes. First, the sample contains 32% males (n = 22) and 68% females (n = 46). Second, we collected all the grades received by the students in their first semester and computed their averages. Each student's grade was weighed by the number of credits obtained for this course, divided by the maximum possible credit points. The grade variable was rounded to 0 decimals. This resulted in a minimal grade of 4, a maximal grade of 9, and an average grade of 6.93 (SD = 1.08). In the Dutch grading system, a 1 is the lowest grade possible, while a 10 is the highest. Third, PFC popularity was measured by asking students who they would prefer to collaborate with. Students rated the statement "I would like to collaborate with [name]" on a 5-point Likert scale, ranging from 1 ("strongly disagree") to 5 ("strongly agree"), with the option of 6 ("I do not know"). Categories 1, 2, 3, and 6 were re-coded as not popular to collaborate with (0), and 4 and 5 as popular to collaborate with (1). We dichotomized this variable because we are (*i*) interested in differentiating whether one is perceived as cooperative or not and (*ii*) to be able to construe whether others recognize a student as cooperative or not. We summed for each individual all nominations from others (i.e., a student's score corresponds to the number of other students perceiving them as cooperative). This variable ranges from 0 to 16, with an average of 6.93 (SD = 3.31).

Descriptively, we find a Pearson correlation of 0.433 (p < 0.001) between PFC popularity and grades, indicating that students with higher grades are more likely to be collaborators. Second, two ANOVAs show that PFC popularity (F[1,68] = 0.070, p = 0.792) and grades (F[1,68] = 0.013, p = 0.910) do not differ significantly per gender.

Table 6.1: Descriptive information per team numbered from 1 to 12.

Team	Size	Gender	Grades		PFC popularity		Friends	Familiarity ties
#	Count	Female (%)	M (SD)	Range	M (SD)	Range	Count	Count
1	7	7 (100%)	7.29 (1.11)	3	10.58 (2.82)	8	24	7
2	4	3 (75%)	7.00 (0.82)	2	7.00 (6.22)	14	4	1
3	6	6 (100%)	6.50 (0.84)	2	4.50 (1.64)	4	8	2
4	6	5 (83%)	7.67 (0.52)	1	8.00 (3.29)	9	8	2
5	6	6 (100%)	6.67 (1.37)	4	5.67 (1.51)	4	2	2
6	6	2 (33%)	6.17 (0.98)	2	6.33 (2.66)	6	10	2
7	5	2 (40%)	5.60 (0.89)	2	4.20 (1.48)	4	1	2
8	6	6 (100%)	7.17 (0.75)	2	7.83 (3.37)	8	8	1
9	7	3 (42%)	6.57 (1.51)	4	6.86 (4.53)	13	12	3
10	7	4 (57%)	7.14 (1.07)	3	8.71 (2.68)	8	3	1
11	6	5 (83%)	8.00 (0.00)	0	5.67 (1.21)	3	6	2
12	4	1 (25%)	6.50 (0.58)	1	6.25 (2.22)	5	2	1

Note. # = number; M = mean; SD = standard deviation; Range = the calculation of the max minus the min score in the same team.

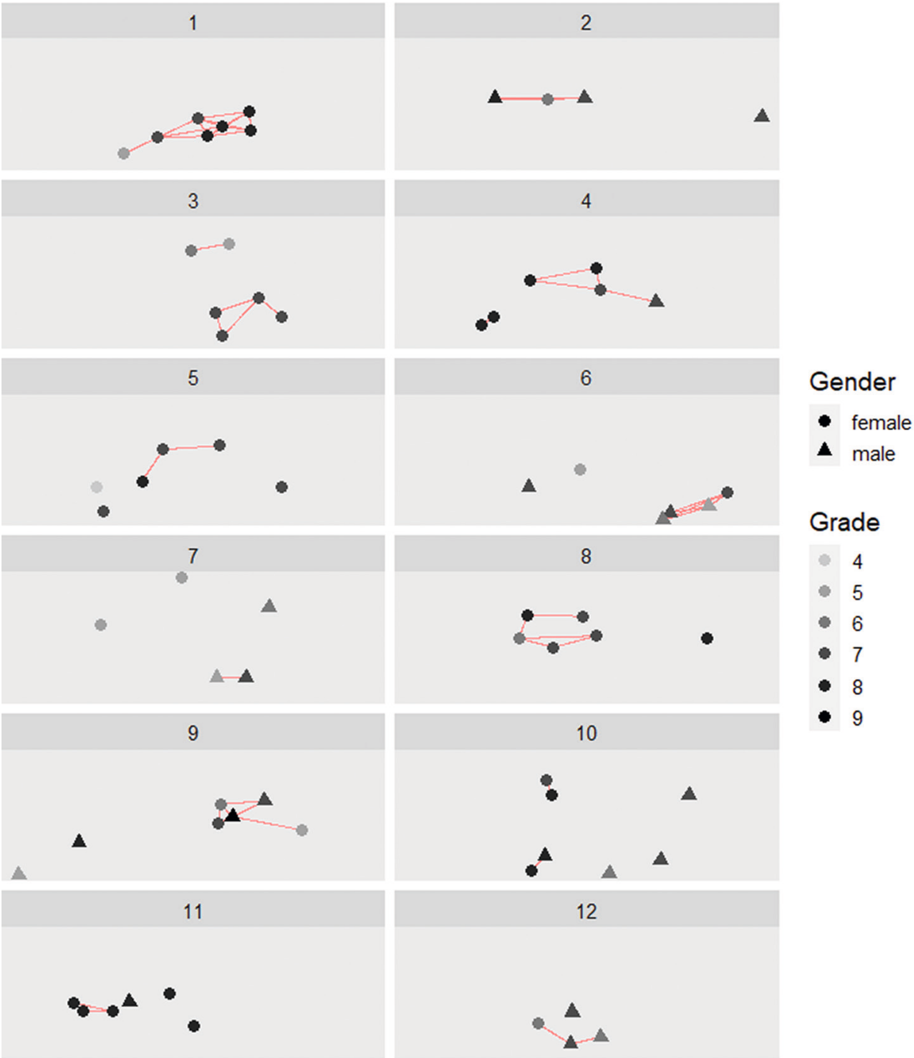


Figure 6.2: Visualizing friendships, gender, and grades per team. A light-red line between nodes refers to a present friendship nomination. Females are visualized in circles and males in triangles. A darker color refers to a higher grade of the student.

6.3.3.4 Missing data

We imputed missing data using the Multivariate Imputation by Chained Equations (MICE) package in R (van Buuren & Groothuis-Oudshoorn, 2011) to be able to specify the ERPMs. The imputed values were generated using predictive mean matching, estimating missing values by matching missing cases to the observed data. Missing data comprised two gender values and two grade values.

6.3.4 Descriptive information on team compositions

Table 6.1 provides for each team summary statistics related to gender, grades, cooperative reputations, friendship nominations, familiarity ties, and team size. Table 6.1 shows that the range of grades and PFC popularity differs across teams. The differences in within-group variance for PFC popularity and grades show that some teams comprise students with similar grades (e.g., teams 11 and 12) and PFC popularity (e.g., teams 5 and 11). In contrast, others contain students with widely different scores on PFC popularity (e.g., teams 2 and 9) and grades (e.g., teams 5 and 9). Most students have at least one friend on their team. 88 (21%) friendship nominations are distributed within the same team, and 330 (79%) friendship nominations are among students not on the same team. Twenty-six students have no friends on their team. Table 6.1 and Figure 6.2 show that some groups comprise many friendship ties, represented as red links between nodes.

6.4 METHODOLOGY

6.4.1 Model definition

We use an exponential random partition model (ERPM; Hoffman et al., 2023) to model the partition of students into teams and uncover the processes driving the formation of these teams. The ERPM framework allows us to explicitly model the number of groups (teams), their sizes, and their composition, considering that groups cannot overlap. This framework is more potent than permutation tests such as the quadratic assignment procedure (QAP; Krackhardt, 1988) that can only investigate group compositions, controlling for (and not modeling) size distribution. ERPMs are also better suited to our data than bipartite exponential random graph models (ERGMs; Wang et al., 2009; see also Lusher et al., 2013) or generalized location systems (GLS) models (Butts, 2007) because they do not require to define in advance the number and characteristics of the groups, unlike bipartite ERGMs and GLS models. In that sense, ERPMs consider the number of groups and their compositions fully emergent rather than exogenously imposed. A complete comparison between ERPMs and other models is provided by Hoffman and colleagues (2023).

6

The model defines a random partition $\boldsymbol{P}$ (i.e., a set of non-overlapping teams) over a set of actors $\{1,\ldots,N\}$. In our case, we have $N = 70$ actors (students). The probability distribution of $\boldsymbol{P}$ is given as an exponential family distribution over the whole set of possible partitions φ . Because team size was restricted (from 4 to 7 members), we restricted the set φ of possible partitions to partitions that contained teams within this size range. The probability of observing a partition p_{obs} is then expressed as:

$$Pr(P{=}p_{observed} \mid \alpha){=} \frac{\exp(\Sigma_k \alpha_k s_k(p_{obs}))}{\Sigma_{p\epsilon\varphi} e^{\exp}(\Sigma_k \alpha_k s_k(p))} \tag{6.1}$$

with $s{=}(s_k)$ a vector of k statistics and $\alpha = (\alpha_k)$ the parameter vector for the distribution. The statistics s_k can describe any characteristic of the partition, but we specifically use here sums of team statistics to make the interpretation of the model easier. These statistics aim to reproduce essential features of the partition and thus capture team formation processes. The maximum likelihood estimates of α can be obtained via stochastic approximation (see section 6.4.3). Because maximum likelihood estimation and the method of moments are equivalent for exponential families, we are essentially looking for the parameter values for which the distribution of the statistics $s(p)$ predicted by the modeled are centered around the observed statistics $s_{obs}(p)$.

Both models are used similarly to a logistic regression to predict ties or teams, considering different types of dependencies. One can note that Equation 6.1 mirrors the joint form of an ERGM (Lusher et al., 2013), with the support of the distribution being a set of partitions rather than a set of networks. While the intuition behind both models remains similar, there are some notable differences. The ERGM expresses the probability of a tie between two nodes being present or absent, depending on the characteristics of these nodes and the presence of other ties in the network. On the other hand, the ERPM expresses the probability of a team being formed, depending on the characteristics of its members and potentially the composition of other teams. Because individuals must belong to one and only one team, the probability of a team indirectly depends on what alternative teams could have been formed. Moreover, the sufficient statistics used in the ERPM differ from those used in ERGMs. Statistics in ERGMS usually measure tie configurations (e.g., the number of ties or the number of same-gender ties). In contrast, ERPM statistics measure team configurations (e.g., the number of teams or the number of same-gender teams).

6.4.2 Model specification

Statistics in the model should represent the factors explaining the composition of the observed partition. The factors of interest for our hypotheses are friendship, familiarity with one another, and similarity in three individual attributes (gender, grade, and PFC popularity). In what follows, we list all the statistics included in the model and highlight per statistic: (1) what the statistic represents, (2) how the statistic is defined, and (3) how the parameter associated with the statistic can be interpreted. We first describe the statistics related to our hypotheses and then define additional statistics used as controls.

The effect of friendship and familiarity are included in the ERPM with the “network tie” statistic, which counts the sum of relations (friendship or familiarity) present within teams in the partition. This statistic captures students’ tendency to form teams with friends and familiar partners. If we define $Z_{i,j}$ as a dyadic binary covariate indicating, for example, whether students are friends or familiar with each other, this statistic is defined as:

$$s_{network\ tie}(p) = \Sigma_{G \epsilon p} \Sigma_{i,j \epsilon G} Z_{i,j} \tag{6.2}$$

A positive parameter for this statistic indicates a tendency for individuals to form teams with peers they have a relationship with; in other words, a tendency to find many network ties within rather than between teams. We, therefore, test *Hypothesis 1* using this statistic with Z defined as the friendship network and *Hypothesis 2* with Z defined as the familiarity network.

We then represent the effect of similarity in gender, grades, or PFC popularity with two types of statistics, to represent the tendency to form homogenous teams regarding these attributes. We first use "dyadic similarity" statistics to capture the tendency for individuals to form teams with similar others– this is equivalent to saying that pairs of similar individuals are more often found within than between teams. For a categorical attribute, we define for each pair of actors *i* and *j* the variable $\boldsymbol{same_{i,j}}(a)$ as a variable taking a value of 1 if the two actors have the same value of the attribute $\boldsymbol{\alpha}$, and 0 otherwise. We use this notation for gender. For continuous attributes, we define $\boldsymbol{diff_{i,j}}(a)$ as the absolute difference between the attribute values of actors *i* and *j*. Lower values of $\boldsymbol{diff_{i,j}}(a)$ point to students being more similar (here, $\boldsymbol{a}$ can be either the grade or the PFC popularity attribute). The dyadic similarity statistic is then defined as a sum of either the same attribute pairs (for categorical variables) or absolute differences (for continuous attributes) within groups:

$$s_{dyadic\ similarity}(p)= \Sigma_{G\epsilon p}\Sigma_{i,j\epsilon G}\boldsymbol{same_{i,j}}(a) \text{ or } \Sigma_{G\epsilon p}\Sigma_{i,j\epsilon G}\boldsymbol{diff_{i,j}}(a) \tag{6.3}$$

In the case of a categorical attribute (here, gender), a positive parameter predicts that more pairs with the same attribute are found in the same teams. In the case of grades and PFC popularity, a negative parameter would indicate a tendency to find, within teams, pairs of individuals with low differences (i.e., similar values).

Second, we define "team similarity" statistics to represent the tendency for individuals to form homogenous teams. For a categorical variable (i.e., gender), the variable $\boldsymbol{same_G}$ takes the value 1 when all individuals in the team $\boldsymbol{G}$ are from the same category and 0 if they do not. For continuous variables (i.e., grades and PFC popularity), we define $\boldsymbol{range_G}$ as the difference between the highest and the lowest value in the group $\boldsymbol{G}$. High values of $\boldsymbol{range_G}$ indicate diversity in the team, while a value of 0 indicates that all members have the same attribute. The team similarity statistic is defined as the number of teams with only members of the same category or the sum of attribute ranges in all teams. Formally, we write:

$$s_{team\ similarity}(p)= \Sigma_{G\epsilon p}\boldsymbol{same_G} \text{ or } \Sigma_{G\epsilon p}\boldsymbol{range_G} \tag{6.4}$$

For gender, a positive parameter associated with this statistic indicates a tendency to form non-mixed teams. The interpretation of the individual level would be that students tend to form teams where everyone is from the same sex as them, in other words, where no one is different from them. A negative parameter for grades and PFC popularity corresponds to a tendency to form teams where the difference between the two most different individuals is low. At the individual level, a student would then tend to have a low difference from the student most different to themselves in the team.

We use dyadic and team similarity statistics for gender, grade, and PFC popularity as two alternative tests of *Hypotheses 3a, 3b,* and *3c*, respectively. The distinction between both types of statistics is essential. Dyadic similarity expresses that students select similar teammates, potentially tolerating some dissimilar ones, while team similarity expresses that students optimize

6

similarity among *all* teammates. Without any clear expectation as to why one would be a better specification than the other, we define models with either specification and compare them to understand which one better represents similarity selection in our context.

One should note that dyadic and team similarity are closely related and that both statistic types might lead to similar attribute distributions in the partition. To understand this, one can consider the case of a student leaving team A with dissimilar others to join team B with more similar others: with this change, team A and team B can become more homogeneous. Eventually, teams might become completely homogenous simply because individuals try to optimize similarity at a dyadic level. A similar effect is well-known in residential segregation models (Flache & de Matos Fernandes, 2021; Schelling, 1971), where dyadic similarity preferences can lead to highly segregated neighborhoods due to relocation cascades.

In addition to the parameters discussed above, our model also contains control statistics. The first control statistic, specific to the partition context, is the statistic counting the number of teams. If the parameter is negative, the model predicts fewer but larger teams than in a random partition. Conversely, a positive parameter expresses a tendency to form many small teams. This statistic thus controls the distribution of team sizes.

Other control statistics account for the popularity of students with higher grades or PFC popularity. Because popularity mechanisms may be confounded with similarity preference mechanisms (see section 6.2.3), we control for popularity specifically by considering that more attractive students may attract more peers in their teams and thus be found in larger teams. We thus include as a statistic the sum of each student's attribute multiplied by the number of their teammates. The statistic is equivalent (up to a multiplicative factor) to summing over teams the product of team size and the average value of the attribute in the team. If we define a_i as the attribute of individual i (their grade or PFC popularity) and $|G|$ as the size of the team G, the popularity statistic is defined as:

$$s_{popularity}(p) = \Sigma_{G \epsilon p} \Sigma_{i \epsilon G} (|G| - 1) a_i \qquad (6.5)$$

Positive parameters for popularity statistics thus express the tendency of individuals scoring high on a given attribute to be found in larger teams. This statistic also has an important influence on the group size distribution.

6.4.3 Analytical strategy

The estimation of the ERPM is implemented in R (R Core Team, 2021) using the R package *ERPM*.[3] The maximum likelihood estimation of the model parameters builds upon the Robbins-Monro algorithm (Robbins & Monro, 1951) initially proposed for ERGMs (Snijders, 2002) using a Markov

3 The package can be found at github.com/marion-hoffman/ERPM. The R-script used for this study is shared via the Open Science Framework (OSF) and accessible via https://doi.org/10.17605/OSF.IO/3GDBX.

Chain Monte Carlo (MCMC) scheme to sample from the distribution (as the calculation of Equation 6.1 is intractable). The statistical significance of the parameters is assessed via a Wald test using the standard error of the estimator. All details on the algorithm are provided by Hoffman and colleagues (2023).

We estimate two models to test our hypotheses. Model 1 uses similarity statistics defined at the level of dyads, while Model 2 includes an alternative statistic for similarity at the team level (see above). Table 6.2 provides an overview of the hypotheses linked to the expected sign of the statistic. Finally, we examine the goodness of fit of both models in Appendix D to assess whether data simulated from the model reproduce the observed data well.

6.5 FINDINGS

We investigate whether friendships, familiarity, and similarity in gender, grades, and PFC popularity affect project team compositions among first-year higher-education students. In what follows, we first discuss per hypothesis preliminary results in the form of the average of friendship nominations and familiarity ties within and across teams, intra-class correlation coefficients (ICC), and Blau indices. The ICC captures intra-team consistency in grades and PFC popularity. The Blau index (also known as the Gini-Simpson index) indicates to what degree teams are diverse regarding gender. After presenting averages of tie nominations, ICCs, or Blau results, we investigate using ERPMs whether students tried to optimize their team composition based on friendships, familiarity, and similarity. The ERPM results are presented in Table 6.3.

Table 6.2: A summary of the hypotheses and expected outcomes of the ERPM statistics.

Description of the hypothesis	Expected sign of the parameter	Actor-oriented explanation of the expected result
H1: Friends are more likely to be found in the same team.	+	A student is more likely to join a team with many friends.
H2: Students familiar with each other due to previous experiences are likelier to be in the same team.	+	A student is more likely to join a team with many familiar others.
H3a: Students of the same gender are likelier to be on the same team.	+	*Dyadic similarity*: A student is more likely to join a team with individuals of the same gender. *Team similarity*: A student is more likely to join teams in which the team is entirely of the same gender
H3b: Students similar in grades are more likely to be found in the same team.	–	*Dyadic similarity*: A student is more likely to join a team with others with low differences in grades. *Team similarity*: A student is likelier to join a team with a lower grade range.
H3c: Students similar in PFC popularity are more likely to be found in the same team.	–	*Dyadic similarity*: A student is more likely to join a team with others with low differences in cooperation reputations. *Team similarity*: A student is likelier to join a team with a lower range in PFC popularity.

6.5.1 Hypothesis 1: Friendships

Hypothesis 1 states that friends are likelier to be in the same team. Each student nominated, on average, 1.26 others as friends within teams and 4.71 between teams. The question of whether friendships are distributed equally between teams or that some teams have more-than-expected friends within the team is answered using the ERPM. We report the results of Model 1, given that both models report a similarly significant and directional effect. The highly significant friendship estimate of 1.11 (SE = 0.16), reported in Table 6.3, indicates that students tend to be part of project teams with friends. A partition in which two students have a friendship in a team is 3.02 ($e^{1.11}$) times more likely than one where they do not have a friendship. The presence of friends appears to be a key feature of group formation. We can thus corroborate *Hypothesis 1*.

Table 6.3: Estimated parameters for the ERPMs studying project team compositions.

	Model 1			Model 2		
	Estimate	**SE**	**OR**	**Estimate**	**SE**	**OR**
Number of teams	−6.76	7.80	<0.01	−6.45	6.51	<0.01
Popularity						
Grade	−0.03	0.07	0.97	−0.02	0.06	0.98
PFC popularity	0.02	0.04	1.02	<0.01	0.04	1.00
Friendships						
Friendship nomination	1.11**	0.16	3.02	1.17**	0.16	3.22
Familiarity						
Familiarity tie	−0.57	0.46	0.56	−0.54	0.45	0.58
Dyadic similarity						
Gender	0.24*	0.10	1.27			
Grade	−0.12	0.06	0.89			
PFC popularity	−0.03	0.03	0.97			
Team similarity						
Gender				1.77*	0.76	5.86
Grade				−0.75	0.42	0.47
PFC popularity				−0.15	0.17	0.86
Log-likelihood		−144.60			−143.89	

Note. Estimate = log-odds; SE = standard error; OR = odds-ratio; Convergence ratios for all presented estimates are ≤ 0.22.
* $p < 0.05$; ** $p < 0.01$.

6.5.2 Hypotheses 2: Familiarity

Hypothesis 2 proposes that students familiar with each other due to previous experiences are more likely to be found in the same team. At first, students have, on average, 0.37 familiarity relations with others in their team and 3.53 of such relations across teams. Most teams have a composition in which students do not share familiarity ties. We formally analyze the role of this feature on team compositions using the ERPMs. In both models, students do not end up in teams with those with whom they shared a learning community or prior project team in the past. For instance, the results in Model 1 suggest that a partition in which two students have a familiarity link is 0.56 ($e^{-0.57}$) less likely than sharing a familiarity link. If we turn this argumentation around using $e^{-0.57}$, students within the same team are 1.77 times more likely to not have a familiarity tie instead of having a familiarity tie. We, therefore, refute *Hypothesis 2*.

6.5.3 Hypotheses 3: Similarity

We hypothesize that students similar in gender (*Hypothesis 3a*), grades (*Hypothesis 3b*), or PFC popularity (*Hypothesis 3c*) are more likely to be in the same team. We find an ICC score of 0.310 and 0.262 for grades and PFC popularity, respectively. ICC scores below 0.5 indicate a poor inter-team alignment of grades and PFC popularity. This suggests that teams do not have a clear constituency with similar grades or PFC popularity but that teams vary in similarity and diversity. Furthermore, the high Blau index of 0.915 for gender indicates that teams do not tend to be diverse. The Blau index result suggests that males and females tend to be with similar others in teams. We use the ERPMs to formally test our similarity hypotheses. Model 1 considers a dyadic similarity measure, and Model 2 includes a team similarity statistic for gender, grades, and PFC popularity.

Model 1 provides evidence for gender similarity (est. = 0.24, *SE* = 0.10). Students are 1.27 ($e^{0.24}$) times more likely to end up in teams with more same-gender others. Students thus tend to preferentially end up in teams with more same-gender students instead of with dissimilar others. Our ERPM results corroborate *Hypothesis 3a*. The results presented in Model 1 refute *Hypotheses 3b* and *3c*. The PFC popularity (est. = –0.03, *SE* = 0.03) and grade (est. = –0.12, *SE* = 0.06) similarity estimates are nonsignificant. Grade differences tend to be low within teams as the *p*-value of the estimate is just above 0.05, namely 0.055. If we take the reverse of the negative estimate, we find that students are 1.13 ($e^{0.12}$) times more likely to end up in teams where grade differences tend to be lower.[4]

4 We estimated ERPMs without the network effects to inspect whether the effect of grade similarity was obscured due to multicollinearity. First, Model 1 without the friendship parameter (all other effects, including familiarity, are in the model) shows a slightly significant grade similarity effect: Est. = –0.13, *SE* = 0.07, *p* = 0.048. This is not surprising given that the friendship network expresses a tendency for grade similarity (Chapter 3): Friends thus tend to have similar grades. All other estimates and standard errors largely mirrored those presented in Table 6.3, Model 1. Second, excluding the familiarity parameter and including all other effects (including friendship) resulted in a grade similarity effect mirroring the one presented in Table 6.3, Model 1: Est. = –0.10, *SE* = 0.08, *p* = 0.191 (all other estimates and standard errors largely mirrored those presented in Table 6.3, Model 1). The inclusion of the familiarity parameter did not particularly interfered with the insignificant grade-team formation relationship we found in Table 6.3, Model 1.

Model 2 investigates whether there is a tendency for team similarity regarding gender, grades, and PFC popularity. Note that the higher log-likelihood of Model 2 indicates that the team similarity statistics fit the data slightly better than the dyadic similarity statistics.[5] For gender, there is a tendency for homogenous teams (est. = 1.77; *SE* = 0.76). Students are 5.86 times more likely to be in a homogenous instead of a more heterogeneous team when it comes to gender. Yet, as we inspected descriptively, all students neither entirely end up in homogenous nor perfectly mixed teams. 9 out of the 12 teams have a constituency in which at least two-thirds are of the same gender. 4 out of the 12 teams have an all-female team. This suggests that there are a lot of homophilous ties, i.e., with similar others, but also non-homophilous ties. Finally, although both estimates are nonsignificant, the negative team similarity estimates for grades (est. = –0.75; *SE* = 0.42), and PFC popularity (est. = –0.15; *SE* = 0.17) indicates that the range of grades and PFC popularity in teams tends to be low, indicating a tendency for team similarity.

Table 6.4: A summary of the hypotheses and expected outcomes of the ERPM statistics.

Hypothesis	Expected sign of the parameter	ERPM sign of the parameter (direction of insignificant parameters)	Description of finding
Hypothesis 1	+	+*	Students tend to join teams with many friends.
Hypothesis 2	+	0 (–)	Familiarity considerations do not drive team formation.
Hypothesis 3a	+	+*	Students tend to join teams with same-gender others.
Hypothesis 3b	–	0 (–)	Students do not tend to join teams with others who have the same grade.
Hypothesis 3c	–	0 (–)	PFC popularity is not a defining factor in team formation.

Note. * Statistically significant, $p < 0.05$ or $p < 0.01$, parameters.

6.5.4 Other ERPM statistics

We can infer from the negative number of teams estimates in Models 1 and 2 that there is a tendency to form fewer teams instead of many. This means that teams comprising more students are larger. The OR value of <0.01 indicates that a partition with a team splitting up into multiple teams is far less likely than a team *not* splitting up into more loose teams. For example, a team comprising six students is more likely to remain than splitting into two teams comprising three students each. The results in Table 6.3 also show that popularity – meaning

5 We studied variations of the team and dyadic similarity statistics to inspect whether the fit of the ERPM improved or not in comparison to Models 1 and 2 in Table 6.3. We first included dyadic similarity for grades and PFC popularity and team similarity for gender. This resulted in a slightly worsened log-likelihood of –155.57. The second variation was to include the dyadic similarity statistic of gender and the team similarity statistics of grades and PFC popularity. The fit of the model deteriorated to a log-likelihood of –248.72.

pursuing teams with constituents who have higher grades and PFC popularity – is not a dominant mechanism for team compositions in both models. We end this section by pointing to Table 6.4. We matched the expected sign with the sign of the ERPM results in both models. Two hypotheses are confirmed, and three are refuted.

6.5.5 Goodness of fit

Appendix D includes an assessment of the goodness of fit (GOF) for both models. We use several auxiliary statistics to evaluate the fit of the simulated to the observed data. Simulations for both models are based on 500 runs. Figures D1 and D2 in Appendix D demonstrate that Models 1 and 2 capture the team sizes and similarity effects well. For example, Figure D1 shows the count of teams with 4, 5, 6, or 7 students in the simulations being well-fitted with the observed data. Even though the simulations slightly over- and underestimate teams with 5 and 6 students, respectively. In Figure D2, we assess the auxiliary similarity outcome not included in the model. The model with dyadic grade similarity recovers well for the tendency of homogenous teams (team similarity). The opposite holds as well: The GOF of the model with team similarity provides a satisfactory fit with a dyadic similarity index.

Moreover, Figure D4 in Appendix D shows that the simulations satisfactorily capture the observed grade differences. For example, most students have a 7 as a grade, and there are thus more connections to similar 7-graders within teams. The simulations capture this tendency well. The ERPMs also effectively model the observed gender combinations, as illustrated in Figure D5. Gender combinations comprise female-female, male-male, and male-female relations. Next, Figures D6 and D7 in Appendix D show that both models fit the observed data regarding the average number of friends and familiar students within teams. Both models provide a satisfactory fit to the observed data, as evidenced by similar log-likelihoods.

6.6 CONCLUDING REMARKS

Forming teams is a crucial aspect of many organizations. Teams allow individuals to pool their skills, knowledge, and resources to achieve common goals and objectives (Bailey & Skvoretz, 2017; Bercovitz & Feldman, 2011). In our case, a team setting allows students to leverage their strengths and complement each other's weaknesses to produce better results than working alone. However, ensuring that teams are formed effectively and efficiently is crucial. We hypothesized that friendships, familiarity through prior teamwork, and the preference for forming teams with similar others play an essential role in team compositions. Akin to earlier findings in which the role of *friendships* next to other types of networks is investigated (Brouwer et al., 2022; Chapter 4), we find that friendships spillover to the team compositions context. *Familiarity* plays no decisive role in this context for the final team composition. The current project team allowed students to form teams of their choosing, thereby suggesting that students possibly did not prefer the initial configuration based on the enrollment date.

Furthermore, gender *similarity* is vital for establishing teams: Students with the same gender are more likely to end up in the same team. Yet, similarities in grades and preference-for-collaboration (PFC) popularity are not significant antecedents of team formation. Hence, this study points to non-academic factors (friendships and gender) as the most important antecedents of team compositions.

This research has practical consequences as social connections cross-cutting team boundaries are expected in many organizations. Whether the team succeeds in realizing the most valuable outcome is not a key question for team formation. We show that students want friends around, possibly indicating that students also “want to have a good time” and potentially find it “safe” to form a team with friends as well as with same-gender others (as found in Brouwer et al., 2018, 2022, and in Chapter 4). For a study director, it may be helpful to encourage students to interact with peers outside of their immediate friend group to form more diverse teams. What is more, organizations outside academia, such as businesses and nonprofits, can use these findings to better understand team compositions and account for factors like friendships and gender homophily when individuals themselves form teams. In such instances, realizing effective, diverse teams may need some managerial oversight if individuals tend to form teams comprising, for example, similar others who also happen to be friends.

Network research is essential for analyzing and addressing various social, economic, and political issues, ranging from segregation dynamics to disseminating misinformation. This tool can advance theory building on team formation in the field. Namely, for academic research, this study is a testament to the importance of networks. For example, research shows abundant information on features affecting network formation in higher education (Brouwer et al., 2018; Stadtfeld et al., 2019). We show that network relations affect team compositions. Collecting network and team data and analyzing the data with appropriate analytical tools helps us better understand how information, resources, and influence flow via the network into other domains, such as teams.

Chapter 6 presents a novel tool in exponential random partitions modeling (ERPM) for analyzing team compositions, contributing theoretically and practically to research on team formation. ERPMs allow researchers and practitioners to understand the complex and dynamic nature of team formations in various settings. The analytical tool provides a comprehensive approach to capturing the interplay between factors such as individual characteristics, social dynamics, and contextual factors affecting team compositions. We show that students do not form teams randomly but that they prefer to team up with friends and similar others in terms of gender. This helps to shed light on the underlying mechanisms that drive team formation and may, as a next step, provide valuable insights into the determinants of effective team performance.

This chapter comes naturally with limitations. First, the literature emphasized the importance of cooperative behavior as individual assets for network and team formation (Simpson & Willer, 2015). However, we do not find that more reputable collaborators stick together in teams or that teams with popular collaborators are more attractive. This discrepancy could be due to several reasons, such as using a crude measure, the lack of clarity in determining whether PFC popularity accurately reflects behavior, and the confounding effect of a high correlation with grades. We propose future studies on whether PFC popularity is important for forming teams in an empirical context to rely on more in-depth individual information regarding cooperation and collaboration. For example, one can rely on social value orientations as a measure explicitly designed to indicate how cooperative one is (Chapter 2).

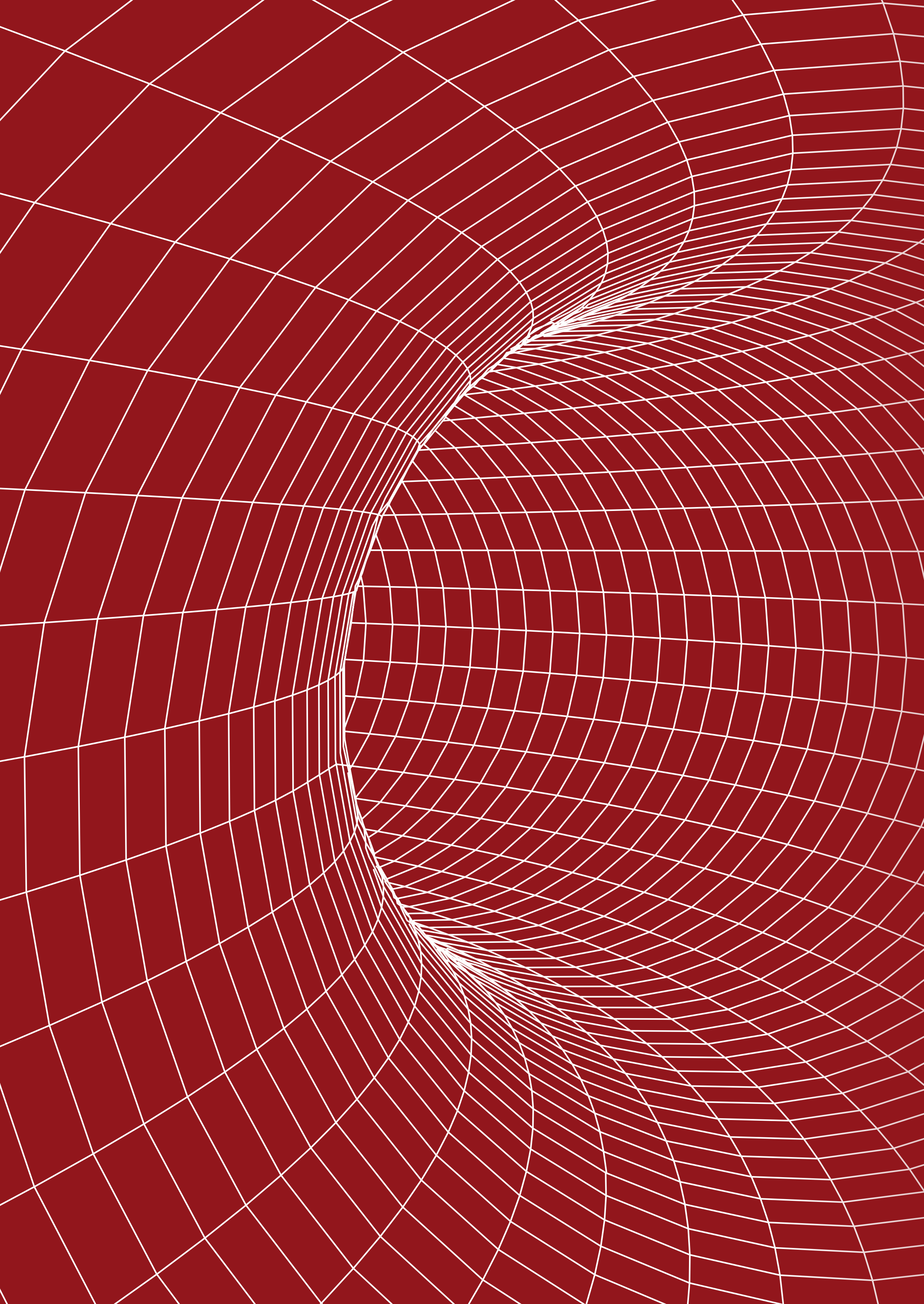
Second, the study was conducted during a crucial period of transitioning from secondary to higher education. Students need to adjust quickly to their new environment. Making friends is essential for first-year students who must adapt to a new social environment. Friendships are among the most important sources of support, help, or peer feedback to achieve academic success (Stadtfeld et al., 2019). Perhaps that is also why friendships are an important factor in the formation of project teams in the first year, whereas skills and abilities become more important later on.

Third, we unexpectedly find that similarity in grades does not foster joining the same team. Research studying how similarity in grades affects network formation shows that students similar in grades are more likely to form ties than students with different grades (Brouwer et al., 2022). On the one hand, a potential implication from this is that lower achievers may be able to learn from their high-grade peers in the team context. Social learning can take off for some students, given that students with similar abilities – reflected in their grades – are not teamed up. On the other hand, grades may become more critical later in their studies. “Having a nice time” may be more important than receiving the highest grade possible.

Some potential avenues for future research studying team composition could include examining the role of diversity and inclusion in team formation and academic performance; examining the role of individual characteristics and personalities in team composition; and exploring ways to optimize team composition. Some work stresses that diverse teams – in terms of gender, ethnicity, creativity, and ability – are better equipped to accommodate contemporary challenges (King et al., 2009; van Knippenberg & Schippers, 2007), whereas others are primarily inconclusive on whether diversity in teams fosters or hampers collective success (Bercovitz & Feldman, 2011; Campbell et al., 2013; Ilgen et al., 2005; Jackson et al., 2003; Yang et al., 2022). This line of literature particularly zooms in on the organizational and work context in which productivity and creativity are the main focus. Also, the diversity-similarity line of research zooms in on whether teams are more able to produce valuable outcomes. Here, we took a step back and studied what features affected the final team composition.

This study has demonstrated the significance of non-academic and academic factors in forming teams. We disentangled whether students form teams based on friendships, familiarity, and similarity in individual attributes, arguing that joining a team based on its ability to realize a productive outcome may, at times, be inferior to joining a team based on non-productivity reasons (e.g., many friends in the team). For some, joining a team with friends ensures that a "good time" is more easily achieved, or some may feel obligated to join teams in which friends reside. Thus, educators must consider both academic and non-academic factors when forming teams.

Effective team formation and small-group teaching aim to create an inclusive and collaborative learning environment that fosters academic growth and student success. Ideally, project teams provide a platform for students to collaborate, share ideas, and support each other, leading to a deeper understanding of the subject matter in the course. Whether the tendency of friends and same-gender students interjects with this is a question for future research. We leave the question of whether more diverse or similar teams generate more valuable outcomes for future research since we studied here the antecedents of team compositions in higher education, not the consequences for productivity.



Chapter 7

Summary in Dutch (Nederlandse samenvatting)

Ik kijk nu terug, en toch heb ik geen spijt
Het waren mooie jaren want wat ik deed
Nooit deed ik iemand kwaad ermee
Het is mijn eigen leven

—Een stuk songtekst uit de single *Ik leef mijn eigen leven* (1994)
van wellicht de beste Nederlandse zanger aller tijden:
Andre Hazes sr.

ABSTRACT

Hoofdstuk 7 vat Hoofdstukken 1 tot en met 6 samen in het Nederlands. Ik ga met name in op de argumenten en informatie die in mijn ogen essentieel zijn voor een lezer om meer te weten over samenwerking (coöperatie) en uitsluiting (exclusie). Ik schrijf dit hoofdstuk vooral voor lezers die geen zin hebben om al het academisch gebrabbel in het Engels in de Hoofdstukken 1 tot en met 6 te lezen. De lezer kan hier een algemene inleiding vinden over wat het samenwerkings-probleem precies is, waarom ik uitsluiting onderzoek en wat de paradox van uitsluiting inhoudt.

7.1 HET SAMENWERKINGSPROBLEEM

Coöperatie, of samenwerking, is van cruciaal belang voor het functioneren van groepen en netwerken. Studenten, activisten, werknemers, atleten en wetenschappers moeten bijvoorbeeld vaak samenwerken met anderen om iets voor elkaar te krijgen dat zij moeilijk, of niet, alleen kunnen doen. Denk aan het organiseren van een buurtbarbecue, werken aan een projectopdracht of het mobiliseren van mensen voor demonstraties of stakingen. Maar ook op globaal niveau is samenwerking essentieel. Ondanks verschillen in taal en cultuur slagen de landen van de EU erin om samen te werken op tal van sociaal-maatschappelijk en juridische terreinen. De VN biedt daarnaast een platform voor landen om samen te werken aan mondiale uitdagingen zoals klimaatverandering, armoede en terrorisme waar de oplossing veelal alleen bereikt kan worden door gezamenlijke afspraken te maken. In al deze voorbeelden moeten zo veel mogelijk betrokkenen een steentje bijdragen om de gewenste uitkomst te bereiken (Hardin, 1968; Komorita & Parks, 1996; Olson, 1965).

Als samenwerking zo essentieel is, en iedereen is zich hiervan bewust, waar zit dan het probleem? Het samenwerkingsprobleem ontstaat doordat iedere betrokkene zich geconfronteerd ziet met een sociaal dilemma (Apicella & Silk, 2019; Attari et al., 2014; Bianchi & Squazzoni, 2015; Dawes, 1980; Komorita & Parks, 1996; Nowak, 2006; Olson, 1965; Ostrom, 1990; Perc et al., 2017; Rand & Nowak, 2013; Simpson & Willer, 2015; Van Lange et al., 2013). Dit dilemma ontstaat wanneer het eigen belang tegenstrijdig is met het belang van het collectief. In deze situatie kan het voor individuen voordelig zijn om niet bij te dragen aan de collectieve inspanning, maar te profiteren van de bijdragen van anderen. Dit staat bekend als "free-riding." Als landen A en B flink minder uitstoten dan hoeft land C dat misschien niet meer te doen. Land C kan de industrie met rust laten, de economie op volle kracht laten draaien en ondertussen wel profiteren van de afgenomen uitstoot waar landen A en B zich voor hebben ingespannen. Dit free-rider gedrag kan echter leiden tot een situatie waarin niemand meer bereid is tot samenwerking. Landen A en B staan namelijk voor dezelfde keuze als land C. Ook voor deze landen is het verleidelijk de oplossing van het probleem aan anderen over te laten. Als alle landen zo redeneren betekent dit uiteindelijk dat geen enkel land de uitstoot verlaagt: een suboptimaal resultaat voor het collectief (Dawes, 1980; Kollock, 1998; Ledyard, 1995). Samenwerking is dus niet vanzelfsprekend omdat "individuele rationaliteit leidt tot collectieve irrationaliteit" (Kollock, 1998, p. 183, vrij vertaald). Ongeacht wat anderen doen, voelt elke betrokkene de verleiding van "free-riding," maar de gewenste collectieve uitkomst kan alleen worden bereikt wanneer voldoende betrokkenen die verleiding weerstaan.

7.2 MECHANISMEN OM HET SAMENWERKINGSPROBLEEM OP TE LOSSEN

Het oplossen van het samenwerkingsprobleem is van groot belang in verschillende domeinen, zoals de economie, sociaal-maatschappelijk domeinen, politiek, milieu en tijdens een pandemie (Greijdanus et al., 2020; Korn et al., 2020; Kraft-Todd et al., 2015; Ostrom, 2010; Van Lange &

Rand, 2022). Begrip van de mechanismen die een positief effect hebben op de ontwikkeling van coöperatief gedrag kan bijdragen aan het ontwerpen van effectieve strategieën om dit dilemma te overwinnen en tot optimale resultaten te komen voor het collectief. Er zijn verschillende strategieën of mechanismen onderzocht om coöperatief gedrag te stimuleren (Axelrod, 1984; Baldassari, 2015; Fehr & Gintis, 2007; Rand & Nowak, 2013; Simpson & Willer, 2015). Eén strategie is bijvoorbeeld om reputaties te gebruiken als middel om samenwerking te bevorderen en free-riding te ontmoedigen (Nowak & Sigmund, 1998). Individuen die bekend staan als betrouwbare samenwerkpartners zijn dan populairder dan diegenen die oncoöperatief gedrag vertonen. Een andere strategie is om straffen en/of beloningen uit de delen, waardoor het minder voordelig is om te free-riden (Fehr & Gächter, 2002; Flache et al., 2017a). Uitsluiting is ook zo'n mechanisme gericht op het stimuleren van coöperatie.

7.3 MIJN BIJDRAGE AAN DE LITERATUUR: DE ROL VAN UITSLUITING

In dit proefschrift richt ik me op gevallen waarin de samenwerking in *dynamische* groepen en netwerken – groepen en netwerken die kunnen veranderen – (mis)lukt. De mogelijkheid om groepen en netwerken te veranderen maakt het namelijk ook mogelijk om free-riders uit te sluiten (Fehl et al., 2011; Guido et al., 2019; Gunnthorsdottir et al., 2007; Rand & Nowak, 2013; Simpson & Willer, 2015).

Voorbeelden van uitsluitingsmechanismen zijn de toelatingsprocedures in het onderwijs, de werving van werknemers, de transfermarkt in het voetbal en de vorming van projectteams in organisaties. In de coöperatieliteratuur gebruik men veelal evolutionaire termen zoals partnerselectie, netwerkselectie of dynamische netwerken om uitsluitingsprocessen te beschrijven (Rand & Nowak, 2013; Simpson & Willer, 2015). Het "selectieperspectief" kijkt met name naar hoe betrokkenen die bereid zijn tot samenwerken gelijkgestemden kiezen. Wanneer betrokkenen die willen samenwerken vooral contact zoeken met elkaar betekent dit vanzelf dat zij *free-riders* uitsluiten. Ik benadruk hier dat de uitsluiting van anderen negatieve en soms onvoorziene gevolgen kan hebben voor het collectief. Vandaar de keuze voor de focus op uitsluiting.

Hoe werkt uitsluiting? Ik probeer dit te illustreren met een voorbeeld. Denk aan vier studenten die moeten samenwerken voor een vak. Zij moeten gezamenlijk een rapport schrijven. Om dat productief te doen verdelen ze de taken. Helaas vertoont één van hen oncoöperatief gedrag, hopende dat een andere student bijvoorbeeld een literatuuronderzoek doet. Als gevolg hiervan kunnen de drie coöperatieve studenten de free-rider uit de groep gooien en *uitsluiten* van de deelname aan toekomstige projectgroepen waar zij onderdeel van zijn. Zo zorgen de coöperatieve studenten ervoor dat alleen zij voordeel halen uit de moeite die zij in het groepsverslag hebben gestoken en zorgen zij er ook voor dat de free-rider niet weer in hun groepje terecht komt. In dit voorbeeld worden de goed samenwerkende studenten beloond voor hun inzet middels een goed cijfer. Dat is ook het doel van het uitsluitingsmechanisme (Rand & Nowak, 2013). Coöperatieve mensen worden beloond voor hun moeite en free-riders kunnen niet meer profiteren van de tijd en energie die de samenwerkende groepsgenoten hebben gestoken in het produceren van waardevolle collectieve output. Wellicht moet de free-rider extra werk verzetten om een voldoende te halen.

Onder ideale omstandigheden stimuleert uitsluiting de samenwerking. Idealiter is de dreiging van uitsluiting voldoende om een free-rider te motiveren om wel zijn of haar beste beentje voor te zetten. Zo kan een potentiële free-rider voorkomen dat diegene wordt uitgesloten. Deze (dreiging van) uitsluiting dwingt de (potentiële) free-rider om te kiezen voor coöperatief gedrag, om zo in de toekomst in aanmerking te komen voor een coöperatieve groep of netwerkrelatie. Coöperatoren beschermen zo het collectieve belang van de groep en motiveren zo potentiële free-riders om wel te blijven samenwerken (anders worden ze uit de groep gegooid). Daarnaast stelt de mogelijkheid tot uitsluiting coöperatieve mensen in staat om te bepalen met wie zij wel of niet een groeps- of netwerkrelatie willen vormen. Dit is voor coöperatoren handig als ze in de toekomst een groepje moeten vormen.

7.4 DE PARADOX VAN UITSLUITING

Het probleem is dat uitsluiting een paradoxaal effect kan hebben op het ontstaan van samenwerking. Uitgesloten free-riders kunnen “gevangen raken” in een vicieuze cirkel waaruit geen ontsnappen meer mogelijk lijkt. Zij hebben bij volledige uitsluiting geen mensen meer in de directe omgeving die coöperatief gedrag willen stimuleren. Het gevolg daarvan is dat een uitsluiting niet het gewenste positieve effect heeft – dat wil zeggen dat het de samenwerking stimuleert – maar dat het leidt tot meer oncoöperatief gedrag. Zo komt een doel van uitsluiting – bevorderen van zoveel mogelijk samenwerking – onder druk te staan. Deze paradoxale aard van uitsluiting wordt met name duidelijk als uitgesloten free-riders ervoor kiezen om geïsoleerd te blijven of zich aansluiten bij gelijkgestemden. In dit geval werkt het uitsluitingsmechanisme averechts en kan het leiden tot minder samenwerking dan gewenst. Free-riders zitten dan als het ware in een fuik.

Ik richt mij in dit proefschrift op de voorwaarden waaronder samenwerking kan ontstaan en in stand kan worden gehouden als de mogelijkheid tot uitsluiting een rol speelt. Daarnaast onderzoek ik ook onder welke omstandigheden uitsluiting niet effectief is in het stimuleren van samenwerking. De hoofdonderzoeksvraag is als volgt:

> **Hoofdonderzoeksvraag**: *Op wat voor manier en onder welke condities kunnen uitsluitingsmechanismen samenwerking bevorderen?*

Er zijn hoofdzakelijk drie redenen waarom ik onderzoek doe naar dit onderwerp. Ten eerste: Aanwezige groeps- of netwerk relaties tussen individuen spelen een cruciale rol in het faciliteren van de eerder genoemde mechanismen die samenwerking bevorderden (Rand & Nowak, 2013; Simpson & Willer, 2015). Maar als zo’n connectie niet aanwezig is dan beperkt dat de functionaliteit van het mechanisme. Een relatie moet bijvoorbeeld aanwezig zijn als groepsgenoten elkaar informeel willen sanctioneren. En een vertrekkende collega kan de negatieve gevolgen van zijn of haar oncoöperatief gedrag ontlopen. Bestaand onderzoek bouwt veelal op de aanwezigheid van connecties, maar die relaties kunnen afwezig zijn door uitsluitingsmechanismen. Daarom is het essentieel om eerst de invloed van uitsluitingsmechanismen op samenwerking te begrijpen (wat ik doe in deze dissertatie) alvorens de rol van mechanismen te bestuderen die expliciet bouwen op aanwezige groeps- of netwerkrelaties (dit laat ik achterwege en is onderwerp van huidig en toekomstig onderzoek).

Ten tweede: Uitsluiting hangt samen met het ongelijkheidsvraagstuk. De kiem van veel maatschappelijke problemen is te vinden in ongelijkheid. Een te hoge mate van ongelijkheid wordt algemeen gekarakteriseerd als disfunctioneel voor een gemeenschap (Piketty, 2014; Scheffer et al., 2017; Stiglitz, 2012; Wilkinson & Pickett, 2009). Een gevolg van uitsluitingsmechanismen in groepen en netwerken is dat ze ongelijkheid tussen samenwerkende groepsgenoten en free-riders kunnen vergroten. Eén “slechte” keuze van een free-rider kan leiden tot systematische ongelijkheid die niet gemakkelijk hersteld kan worden, omdat de free-rider vastzit in een fuik waar het vertonen van oncoöperatief gedrag is genormaliseerd. In dit soort gevallen kan het moeilijk zijn om free-riders weer tot samenwerking te stimuleren. Het paradoxale karakter van uitsluiting kan dus een nadelig effect hebben op de algemene (on)coöperatieve omgeving en het niveau van ongelijkheid.

Ten derde: In veel wetenschappelijk onderzoek wordt samenwerking en uitsluiting niet bestudeerd in empirische en realistische situaties waarin ook vele andere gedrags- en sociale facetten een rol spelen. Dit is niet noodzakelijk een probleem. Experimentele controle biedt juist de mogelijkheid om het effect van uitsluiting op samenwerking te isoleren. Als deze experimentele controle andere belangrijke en potentieel verklarende sociale mechanismen niet onderzoekt is dat echter wel problematisch. Met name het effect van uitsluitingsmechanismen op samenwerking kan worden verstoord door andere processen waarvan bekend is dat zij de gedrags- en relationele keuzes van mensen beïnvloeden. Dit illustreer ik op twee manieren.

Als eerste wijs ik naar de bevinding dat beslissingen en individuele handelingen veelal niet rationeel zijn (Macy & Flache, 1995; Simon, 1982; Udehn, 2001; Wittek et al., 2013). Mensen werken bijvoorbeeld niet altijd samen omdat ze niet over voldoende informatie beschikken of de gevolgen van hun acties niet goed kunnen inschatten. Het is met name problematisch als oncoöperatief gedrag dat leidt tot uitsluiting gebaseerd is op een individuele handeling die wordt beïnvloed door beperkte informatie. Individuele besluitvorming is beperkt en daar hou ik rekening mee tijdens het bestuderen van de effectiviteit van uitsluiting op het stimuleren van samenwerking.

Ten tweede verwijs ik naar het uitsluiten – klein grapje, moet kunnen – van andere sociale processen in experimenten en theoretische modellen. Dit is begrijpelijk omdat het de mogelijkheid biedt om analytisch uit te pluizen hoe en wanneer uitsluiting de samenwerking wel of niet stimuleert. De realiteit is uiteraard anders. Veel verschillende processen spelen tegelijkertijd een rol in onze beslissingen. Wij leven als mensen niet in een cocon waar maar één mechanisme aanwezig is en gedrag deterministisch is. We zijn allemaal ingebed in diverse groepen en netwerken. Ik geef een voorbeeld. Wederkerigheid komt in diverse vormen voor, maar in essentie gaat het om “jij helpt mij, ik help jou” (Kadushin, 2012). Als een free-rider bijvoorbeeld vrienden is met groepsgenoten dan kan wederkerigheid uitsluiting ondermijnen. Het is wellicht lastig om een goeie vriend uit de groep te gooien. Wederkerigheid kan er dus voor zorgen dat free-rider in een groep of netwerk worden getolereerd, waardoor uitsluiting even plat slaat als lauw bier. Er zijn dus meerdere redenen om in het “echte leven” connecties aan te gaan en om samen te werken met anderen. De vraag is of uitsluiting in meer realistische situaties nog steeds werkt als oplossing voor het coöperatieprobleem. In vijf hoofdstukken ga ik hier dieper op in.

7.5 DE DIEPTE IN: SAMENVATTING VAN DE HOOFDSTUKKEN

In welk opzicht zijn persoonlijkheidskenmerken – waarop iemand een uitsluitingskeuze zou kunnen baseren – van potentiële samenwerkingspartners stabiel over een langere periode? In Hoofdstuk 2 bestudeer ik in hoeverre *social value orientations* (SVOs) stabiel zijn in het meten van pro-socialiteit. SVO meet de mate waarin iemand pro-sociaal of meer egoïstisch is ten aanzien van samenwerking. Hoe pro-socialer iemand is, hoe meer geneigd die persoon is om samen te werken. Dit soort informatie kan dienen als input voor uitsluiting. In Hoofdstuk 2 kom ik – met behulp van paneldata met herhaalde metingen – tot de conclusie dat men vrij stabiel is in zijn of haar SVO. De analyses valideren SVO als een betrouwbare schaal en laat zien dat SVO potentieel een rol kan spelen in de beslissing om wel of niet uit te sluiten.

Hoe beïnvloedt imperfecte informatie de effectiviteit van uitsluitingsmechanismen in groepen? *Is het voldoende dat actoren stabiel zijn in hun pro-socialiteit*? *Werkt uitsluiting dan nog wel als oplossing voor het samenwerkingsprobleem*? Ik beantwoord deze vragen theoretisch in Hoofdstuk 3. Daarom spreek ik in dit hoofdstuk over kunstmatige *agents*. Ik onderzoek of groeperen (*meritocratic matching*) minder effectief is wanneer de input voor de groepering gebaseerd is op imperfecte informatie. Imperfecte informatie betekent hier dat men niet kan zien wat iedereen individueel heeft bijgedragen aan het collectief, maar alleen groepsuitkomsten kan zien. Het model laat zien dat pro-socialen in oncoöperatieve groepen dan niet te onderscheiden zijn van egoïstische types. Het gevolg hiervan is dat deze pro-socialen later geen toegang krijgen tot meer coöperatieve groepen. Informatie uit sociale netwerken biedt een uitkomst voor dit soort pro-sociale agents die bekend staan als oncoöperatief. De combinatie van informatie uit groepen en netwerken creëert een situatie waarin uitsluiting in groepen beter kan functioneren. Desondanks heeft de “netwerkoplossing” een keerzijde: Hoofdstuk 3 laat zien dat wanneer pro-socialen ontsnappen uit oncoöperatieve groepen, egoïstische types minder interactie hebben met pro-socialen, waardoor zij niet kunnen ontdekken dat samenwerken ook voordelen kan hebben. Hier zien we dus de paradox van uitsluiting.

Speelt uitsluiting op basis van coöperatieve overwegingen een rol in “echte” vriendschapsnetwerken? De theoretische resultaten in Hoofdstuk 3 laten zien dat uitsluiting in groepen de samenwerking stimuleert als de informatie over samenwerking in netwerken tijdens het groepsvorming proces wordt gebruikt. Maar de vraag is of individuen coöperatieve overwegingen meenemen bij het vormen van netwerkrelaties in een empirische situatie (hier: studenten in het hoger onderwijs). Hoofdstuk 4 onderzoekt of dit het geval is. Ik bestudeer wat de drijfveren van studenten zijn om vriendschaps- en samenwerkingsrelaties te vormen. Uitsluiting speelt in Hoofdstuk 4 een centrale factor: Hebben bevriende studenten ook de neiging om samenwerkingsrelaties aan te gaan, met als gevolg dat niet-vrienden worden uitgesloten van samenwerkingsrelaties? De netwerkrelaties – lees vriendschap- en samenwerkingsrelaties – worden dan gebruikt als uitsluitingsmechanisme. De resultaten laten zien dat vrienden eerder samenwerkingspartners worden (en dit geldt omgekeerd ook: Studenten die samenwerken tijdens de studie worden sneller vrienden). Bovendien worden studenten die allebei populair zijn als samenwerkingspartner sneller vrienden. Wellicht is populariteit als samenwerkingspartner een voedingsbodem voor een vriendschap omdat het aantoont dat beide studenten wel willen werken om goede cijfers te halen. Hoofdstuk 4 toont aan dat netwerken ook fungeren als uitsluitingsmechanisme: Vrienden hebben een voorkeur voor elkaar als samenwerkingspartner.

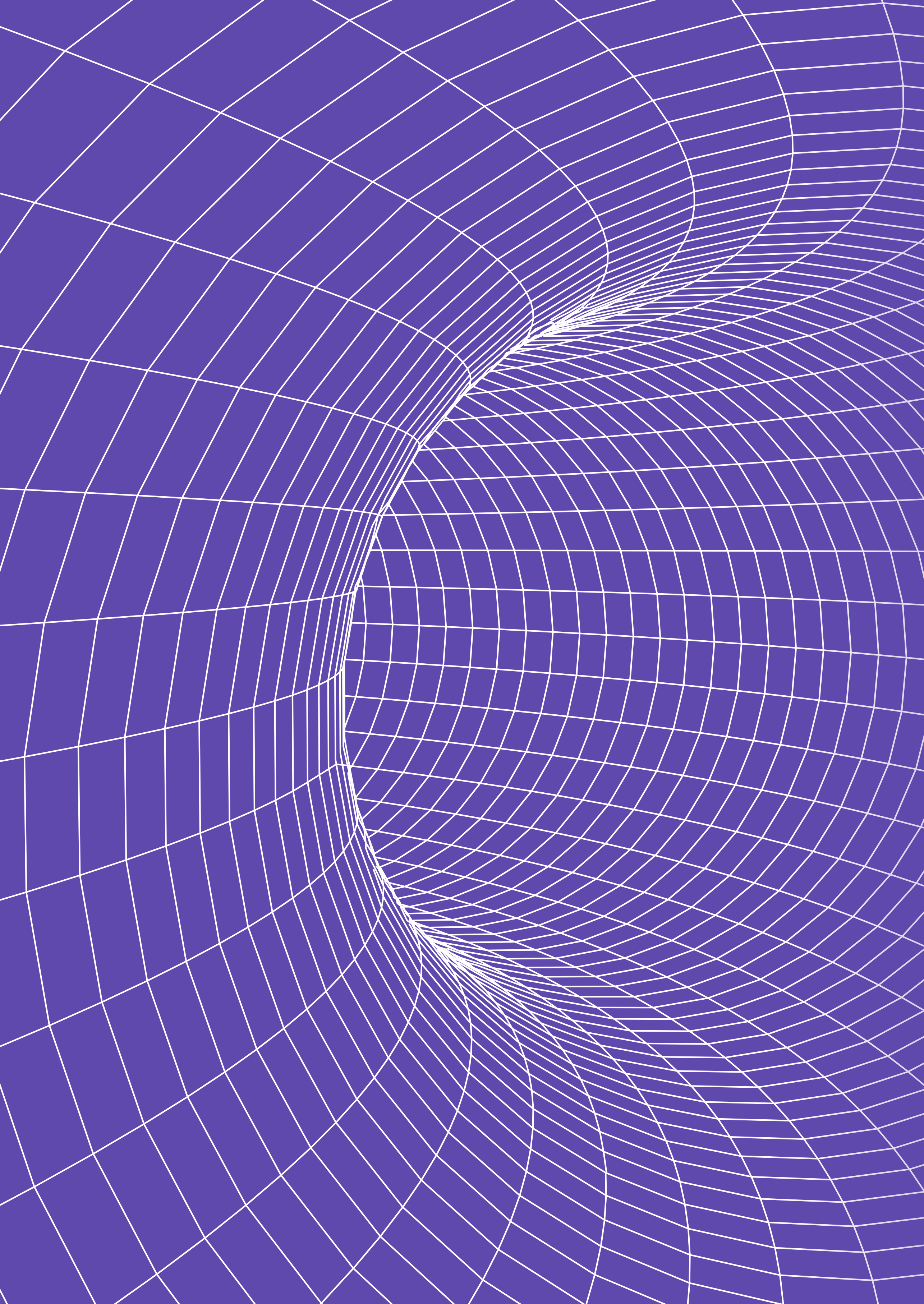
In hoeverre wordt samenwerking in een realistische setting beïnvloed als ook sociale beïnvloeding een rol speelt naast het uitsluitingsmechanisme? Met behulp van een empirisch gekalibreerd model ga ik in Hoofdstuk 5 een stapje verder dan in Hoofdstukken 3 en 4. Ik verken zogenoemde “wat als” situaties waarin ik de preferentie voor uitsluiting en sociale beïnvloeding kunstmatig versterk. Dit betekent dat de agents in het model sneller geneigd zijn om uit te sluiten op basis van samenwerkingsgedrag, en dat agents beïnvloedbaar zijn. We hebben in Hoofdstuk 3 gezien dat sociale beïnvloeding vaak een groter effect op coöperatief gedrag heeft dan persoonlijke voorkeuren (SVO; zie ook Hoofdstuk 2). Het simulatiemodel in Hoofdstuk 5 gebruikt empirische data als input. De simulaties laten zien dat coöperatie het best gestimuleerd wordt wanneer uitsluiting en sociale beïnvloeding tegelijkertijd in sterke mate aanwezig zijn. Een verklaring hiervoor is dat samenwerkers elkaar sneller vinden in het netwerk door uitsluiting en elkaar dan ook beïnvloeden om te blijven coöpereren. In navolging van de theoretische bevindingen uit Hoofdstuk 3 wijst het model in Hoofdstuk 5 opnieuw naar de paradox voor uitgesloten free-riders: Sommigen blijven in hun eigen oncoöperatieve fuik en worden niet gestimuleerd om samen te werken, waardoor hun kans om te “ontsnappen” uit de fuik afneemt.

Op welke manier beïnvloedt samenwerking de voorkeuren voor teamsamenstellingen van studenten in een omgeving waar kenmerken zoals vriendschappen, geslacht, cijfers en eerdere persoonlijke ervaringen met elkaar ook een rol spelen in het teamvormingsproces? Hoofdstuk 6 breidt de theoretische analyse van Hoofdstuk 3 uit door empirisch te onderzoeken of studenten daadwerkelijk teams vormen op basis van coöperatieve overwegingen. Hoofdstuk 6 bouwt daarnaast voort op Hoofdstuk 4 maar neemt de mate waarin een student een populaire samenwerkingspartner is mee als indicator. Het uitsluitingsmechanisme dat in Hoofdstuk 6 wordt onderzocht, bestaat uit de voorkeur van een student voor een bepaald team (wellicht met gelijkgestemden) met als gevolg dat anderen (en andere teams) worden uitgesloten. Ik controleer voor diverse factoren die mogelijk ook van invloed zijn op teamvorming. In Hoofdstuk 6 laat ik zien dat vrienden en studenten van hetzelfde geslacht geneigd zijn om zich bij hetzelfde projectteam aan te sluiten. Of iemand wel of niet bekend staat als een populaire coöperator maakt niet veel uit tijdens het teamvormingsproces.

7.6 EN NU?

Samenwerking is onderdeel van onze alledaagse interacties. In dit proefschrift heb ik mij gericht op hoe uitsluitingsmechanismen het samenwerkingsprobleem kunnen oplossen, maar ik heb ook gewezen naar de paradoxale aard van uitsluiting. Ik laat in dit proefschrift zien dat een interdisciplinaire benadering – waarbij inzichten uit de psychologie en sociologie worden gecombineerd – een stap in de goede richting is om samenwerkingsproblemen te onderzoeken dat dicht tegen samenwerking in de realiteit zit. Deze benadering heeft mij de mogelijkheid geboden om vanuit verschillende perspectieven naar hetzelfde probleem te kijken. Mijns inziens is toekomstig onderzoek gebaat bij (*i*) het meenemen van het "individu-perspectief" (Hoofdstukken 2 t/m 4), (*ii*) een benadering waarbij de dynamiek van groepen netwerken wordt meegenomen in het bestuderen van samenwerking (Hoofdstukken 3 t/m 6) en (*iii*) het gebruik maken van een combinatie van statistische methoden, theoretische modellen en empirische data die rekening houden met de interactie tussen het individu en de context waarin uitsluiting plaatsvindt (Hoofdstukken 2 t/m 6).

Tot slot, dit onderzoek laat zien dat uitsluitingsmechanismen een keerzijde hebben: Als een free-rider aan de "verkeerde kant" van de medaille staat dan is het moeilijk om "terug" te gaan. Als free-riders geen sociale druk ervaren om te coöpereren dan blijven ze doen wat ze hiervoor deden – oncoöperatief gedrag vertonen – en wat men om hen heen ook doet. De situatie wordt nog nijpender als free-riders geïsoleerd raken of zijn en dus geen enkele externe stimulans krijgen om hun gedrag te veranderen. Ik benadruk hier dat een "out" dus gepaard moet gaan met een "in". Niet alleen free-riders moeten werken om te leren samenwerken, maar ook zij die willen samenwerken moeten voorwaarden scheppen waarin free-riders wel tot samenwerking "kunnen komen."



Supplementary Material

A. APPENDICES TO CHAPTER 3

A1. Two additional matching rules

We took a step-wise approach to capture every small change in the matching rule and inspected how such minor alterations in matching rules affect model dynamics (Flache & de Matos Fernandes, 2021). We aim to preserve, on the one hand, the importance of rigorous model building while, on the other hand, favoring less information-heavy scenarios. See the Chapter 3 descriptions of rules 1-3. The two new rules are as follows:

> ***Rule 1 adjusted.*** We omit the pre-game in this rule to move away from the complete information assumption under rule 1. Agents are initially randomly grouped, allowing for initial mismatching. Still, agents have, during the game, unlimited cognitive abilities to store all individual prior actions of all agents in the population.
>
> ***Rule 2 adjusted.*** To explore the role of individual incomplete information instead of solely relying on the information at the group level, we include the last 10 individual cooperation decisions (C_{10}) as input for the matching algorithm of all agents instead of G_{10}.

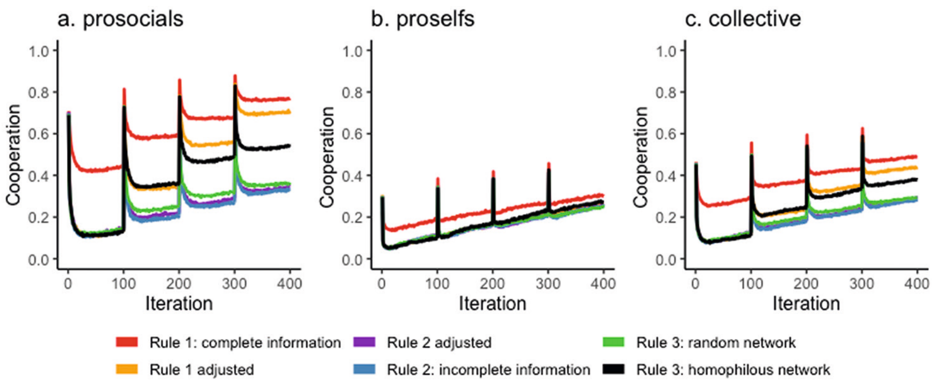


Figure A1: Average level of cooperation for prosocials (a), proselfs (b), and the collective (c), separated by the 'original' and additional matching rules. Parameter settings: $m = 5$; $l = 0.5$; $PA = 0.4$; $r = 0.05$.

The results are visualized in Figure A1, showing the average level of cooperation resulting from 100 simulation runs. *Rule 1 adjusted* shows that prosocials can harvest more of their innate cooperativeness compared to incomplete information rules. However, the possibility of initial mismatching, assumed under rule 1 adjusted, leads to lower cooperation among prosocials than under rule 1 (Figure A1a, yellow line). Figure A1b reports no radical impact of additional matching rules on cooperation levels among proself agents. The collective level of cooperation, akin to cooperation levels among prosocials, benefits from rule 1 adjusted. Moreover, findings regarding *rule 2 adjusted*, i.e., only the last 10 individual actions are known, are also detrimental to cooperation levels among prosocials (Figure A1a, purple line). An observation of individual behavior from the last 10 iterations in the context of that group is equally bad for good apples as when merit is solely group-based.

A2. Consequences of removing meritocratic matching: keeping groups fixed

In-between matching moments, we infer that cooperation increases linearly. This leads us to conduct exploratory simulation runs and inspect the role of keeping groups fixed. Do agents quickly learn to cooperate when we remove the dynamic part from the model? We first inspect how cooperation evolves over 50 rounds, specified per prosocial, proself, and overall (Figure A2a-c), followed by studying cooperation over more extended periods (Figure A2d-f). Lines with lighter shades of grey point to fewer prosocials in the group. It is important to note that agents always have a slight chance to cooperate even when defection prevails (and vice versa). Thus, due to the probabilistic nature of our model, we do not find smoothened curves in Figure A2 but mainly punctuated equilibria. We report the average level of cooperation of 100 simulation runs.

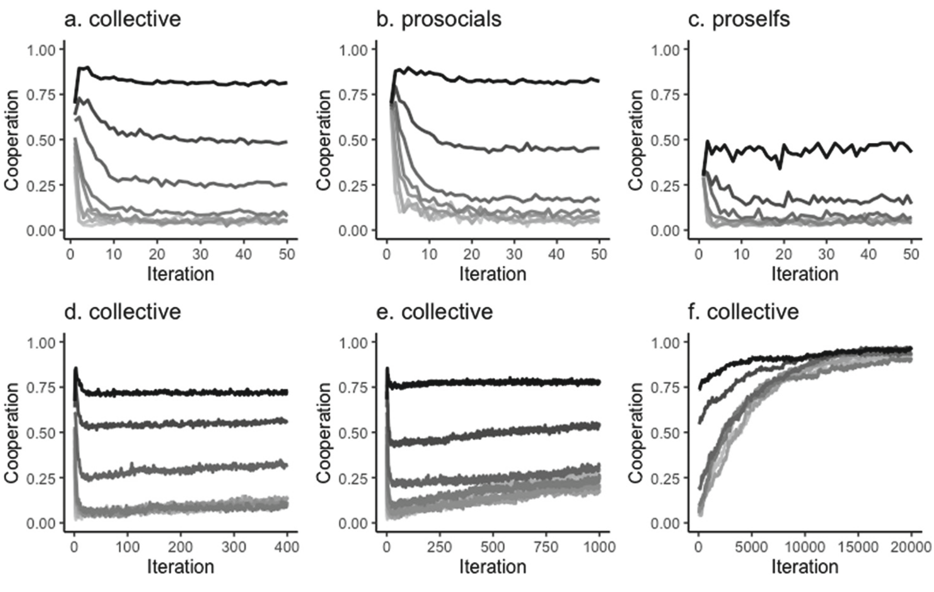


Figure A2: Average level of cooperation in fixed groups per 50, 400, 1000, and 20000 rounds (panels a, d, e, and f, respectively). Cooperation levels of prosocials (b) and proselfs (c) are reported until iteration 50. Darker shades mean more prosocials in the group (range = 0, 1, 2, …, 8 prosocials). Parameter settings: $m = 5$; $l = 0.5$; $PA = 0.4$; $r = 0.05$.

Figure A2f shows that cooperation needs time to arise. All group configurations tend to converge at all-out cooperation if we set the time horizon to 20000 rounds. However, groups with fewer prosocials tend to cooperate at lower levels when interacting over ≤ 5000 rounds. Prosocials are more able to cooperate when there are more prosocials in the group (darker lines), but defection is more likely when there are too many bad apples (lighter lines) in the group (Figure

A

A2d-e). Keeping groups fixed is thus a solution to promote collective success in the long run. Still, it may be an unrealistic scenario if we translate this finding to real life, where groups are usually dynamic while facing external influences.

Our explanation for the low levels of cooperation in groups with fewer prosocials in Figure A2a-c is as follows: Prosocials generally cooperate initially but quickly fall in line with their group mates' defection if the outcome of cooperation tends to be negative (Figure A2b-c). Figure A2b shows that the average level of cooperation at the start relates to the initial probability to cooperate for prosocial and proself agents, 0.7 and 0.3, respectively. But if cooperation does not generate positive outcomes – which is more likely in groups with more proself agents – we see that cooperation of prosocials stabilizes at substantially lower levels than the initial probability to cooperate, visualized by lighter shades of grey (Figure A2a-c). Prosocials are the drivers of cooperation in our model; defection prevails if they do not show up.

A3. Network cooperation

Please find the BehaviorSpace (100 simulation runs per condition) findings of network cooperation levels specified per network condition in Figure A3. The single run is reported in Chapter 3.

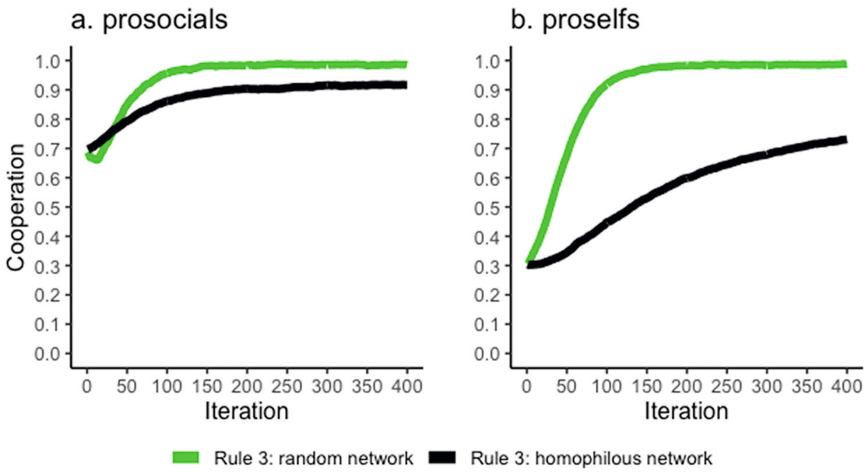


Figure A3: Average level of cooperation in the network, separated by random (green) and homophilous (black) network. Parameter settings: m = 5; l = 0.5; PA = 0.4; r = 0.05.

A4. Sensitivity analysis: Testing the impact of *m*, *l*, and *PA*

To check the impact of little or much noise in the decision algorithm, we incorporate values of m = {1, 5, 10}. Also, we inspect the implications for slow and rapid learning by assuming l = {0.1, 0.3, 0.5, 0.7, 0.9}. Next, we control how the initial distribution of prosocial agents affect model outcomes, PA = {0.2, 0.4, 0.6, 0.8}. We run 100 independent simulations per condition. There are overall 6 x 3 x 5 x 4 conditions – matching, slope, learning rate, and proportion prosocials, respectively – realizing a total of 36000 simulation runs, all things being equal.

We present the average level of cooperation of prosocials (Figure A5), proselfs (Figure A6), and the collective (Figure A7), as well as for the prosociality segregation MS index (Figure A4). We only included the last interaction of runs. Although the variance across matching rules decreases when comparing rules per *PA*, the order of rule 1 (complete information) > rule 3 (homophily) > role 2 (incomplete information) remains fairly the same when inspecting cooperation levels of prosocials, leading us to infer that our findings are relatively robust when $m = 5$ and $l < 0.9$. In what follows, we only discuss findings regarding prosocials given that the picture does not radically change when inspecting how robust the findings are for proselfs and the collective.

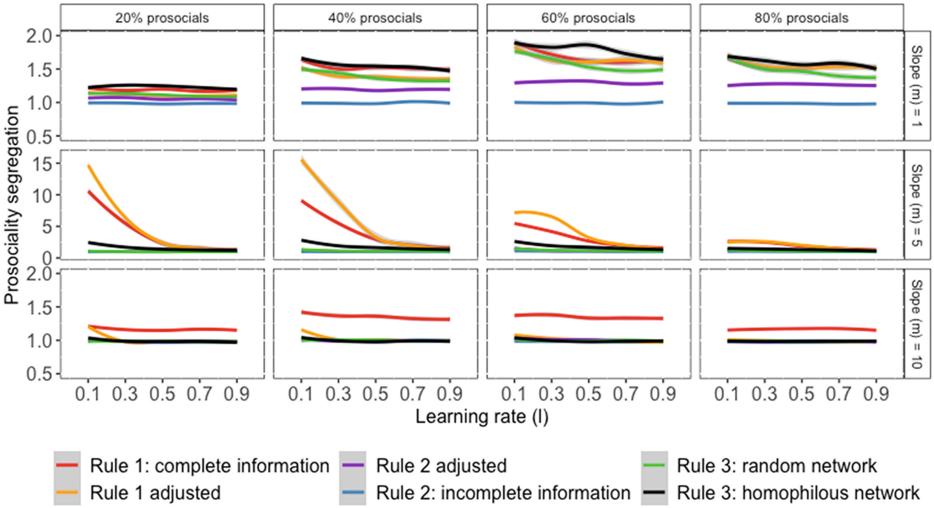


Figure A4: Sensitivity analysis of prosociality segregation per l, m, and PA. Mean and 95% confidence interval are depicted. Note that we report different values for the y-axis when $m = 5$.

Changes in cooperation outcomes regarding learning and noise may be expected. Macy (1991, p. 739) states that "rapid learning suggests pragmatic error-correction, while slow learning may indicate habitual or norm-guided behavior […] that take somewhat longer to change." We find the same in Figure A5, middle row. High learning rates allow agents to coordinate the best course of action quickly. Whereas slow learning rates, $l = 0.1$ or 0.3, show the need for more time to avoid negative outcomes.

More noise ($m = 1$) leads to a self-correcting equilibrium (Macy & Flache, 2002) where cooperation levels hover around 0.34 (Figure A5, top row). In such an equilibrium, the expected change in cooperation levels is minimal as the benefits of cooperation and defection counter each other. Contrarily, prosocials end up in an equilibrium of all-out cooperation (Figure A5, bottom row) if noise decreases to $m = 10$. Macy and Flache dubbed this a self-reinforcing equilibrium in which positive payoffs reinforce behavior even if alternative actions, i.e., defection, may lead to a higher payoff.

Figure A4 allows us further to assess the robustness of our prosociality segregation finding. The middle row in Figure A4 shows that prosociality segregation does not arise when prosocials

A

quickly learn to cooperate when *l* = 0.7 or 0.9, omitting the need to team up to spur cooperation. Next, more noise (*m* = 1) in the behavioral algorithm has a peculiar effect on homophily as it becomes the go-to matching rule to spur prosociality segregation. Still, cooperation levels hover around 0.34 (see figure A5). Furthermore, lower levels of noise (*m* = 10) point to the importance of initial grouping for prosociality segregation to arise. The possibility of pre-game grouping – in which prosocials are more likely to team up – leads to substantially higher prosocial segregation levels than initial random grouping, irrespective of *PA* and *l*.

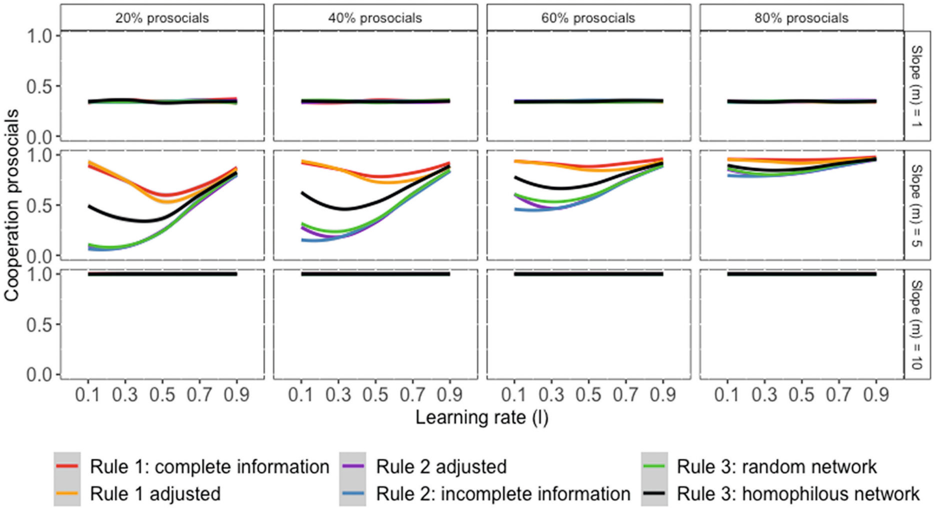


Figure A5: The effect of m, l, and PA on the mean level of cooperation of prosocials at iteration 400. Mean and 95% confidence interval are presented per rule.

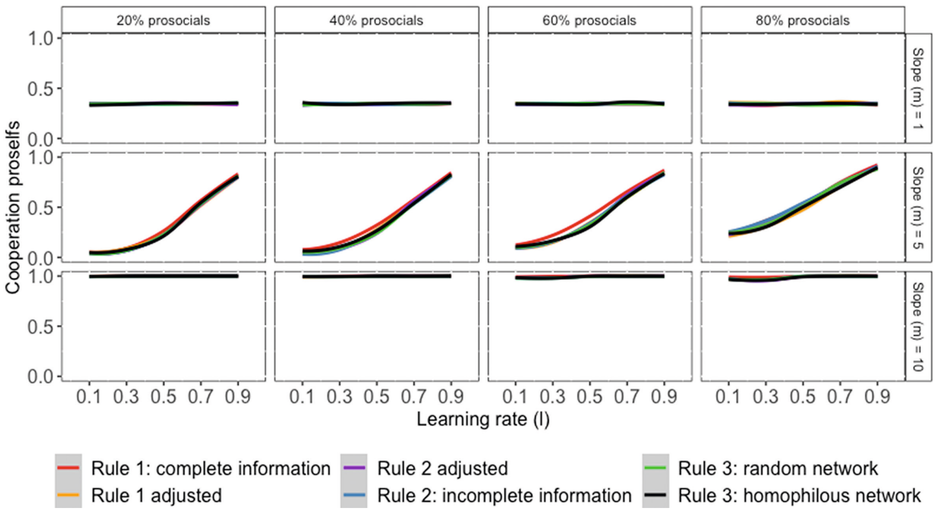


Figure A6: The effect of m, l, and PA on the mean level of cooperation of proselfs at iteration 400. Mean and 95% confidence interval are presented per rule.

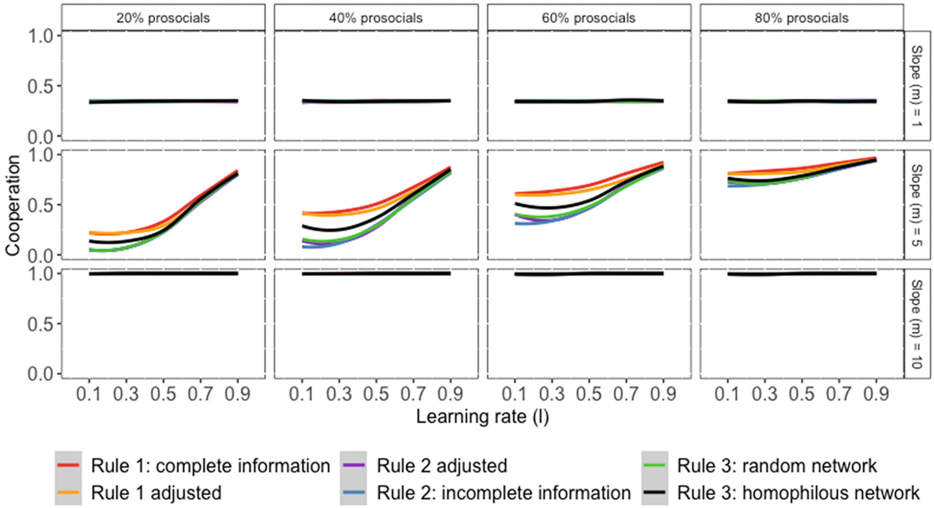


Figure A7: The effect of m, l, and PA on the mean level of collective cooperation at iteration 400. Mean and 95% confidence interval are presented per rule.

A5. Sensitivity analysis: Noise in leave-stay procedure

Noise is essential in threshold models (Macy & Evtushenko, 2020) but may also impact the leave-stay procedure. Some may wrongly want to leave the group, even if they are happy with the group performance. We aim to answer if a higher level of noise in the leave-stay procedure impedes or promotes the effectiveness of meritocratic matching in the long haul (noise leave-stay incorporates values: 0, 0.01, 0.05, 0.25), all other parameter settings being equal. The parameter set refers to the chance of activation, meaning that with 0.25, there is a 25 percent chance of noise implementation. Also, the parameter refers to the proportion of stayers put in the leavers pool. For example, when we set noise to 0.25, a random selection of 25 percent of stayers is put in the leavers pool. We run 100 independent simulations per condition. There are in total 6 x 4 conditions – matching and noise in leave-stay, respectively – realizing 2400 simulation runs. We report the average level of cooperation of prosocial agents in Figure A8.

Figure A8 elucidates that noise in the leave-stay procedure does not promote or impede cooperation among prosocials, proselfs, or the collective, as well as for prosociality segregation compared to our main findings reported in Chapter 3. Inserting more empty slots in groups and promoting more movement does not make our solution to the bad barrels problem stronger or weaker.

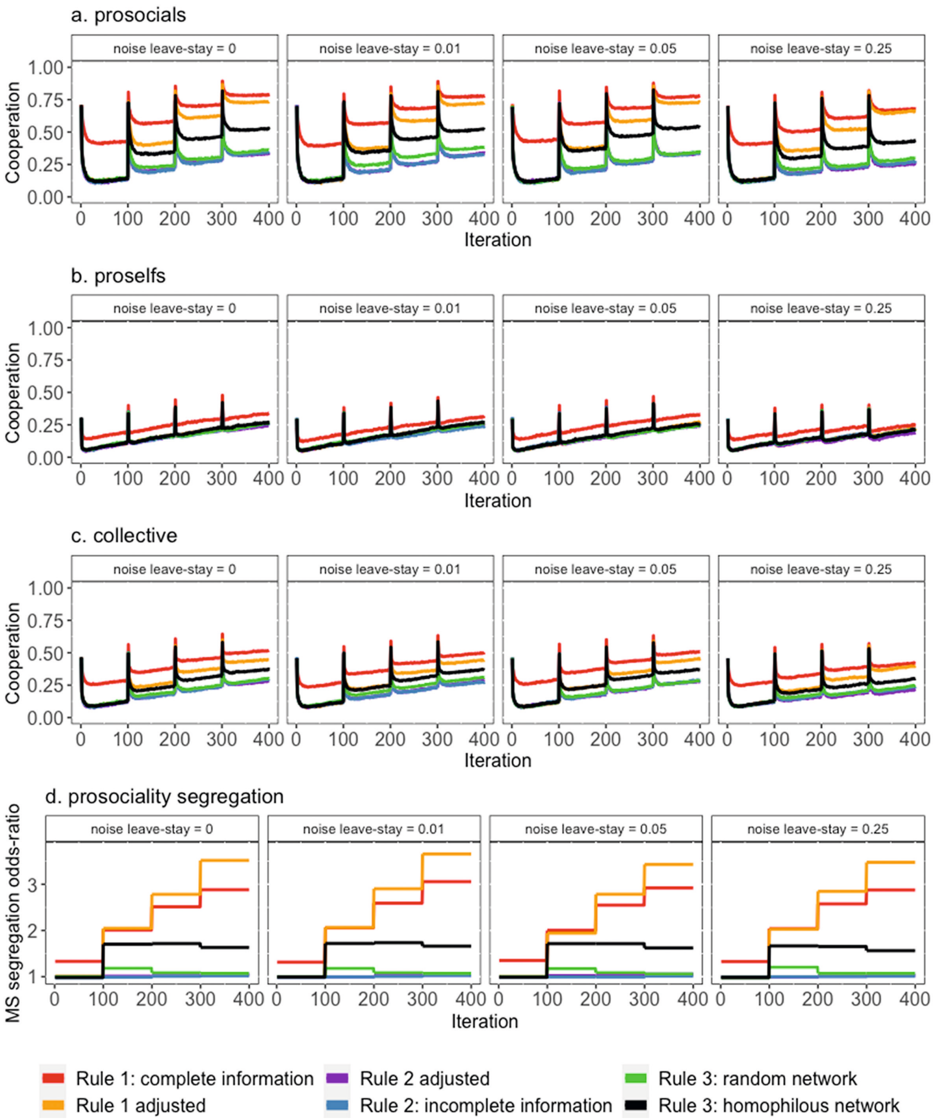


Figure A8: Robustness check of noise in the leave-stay procedure, separated per matching rule. We report the average level of cooperation among prosocial (a) and proself (b) agents as well as for the collective (c). The degree of prosociality segregation in the group context is reported in panel d. Parameter settings: $m = 5$; $l = 0.5$; $PA = 0.4$; $r = 0.05$.

A6. Sensitivity analysis: Altering input leave-stay procedure

Figure A9 shows that our findings are robust when altering agent-level input in the leave-stay procedure. 100 independent simulation runs per condition, a total of 1800 runs (6 x 3 x 100).

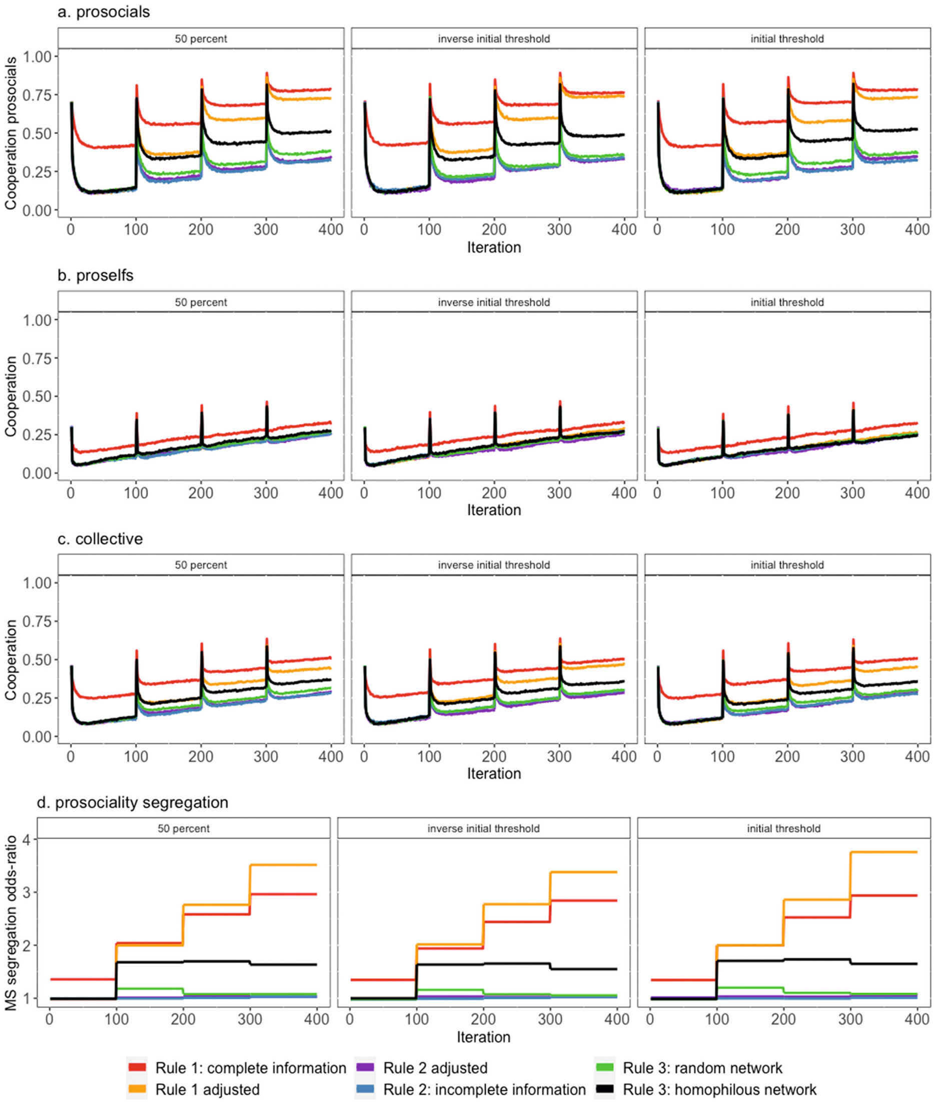


Figure A9: Robustness check of input in the leave-stay procedure, separated per matching rule. Agents leave either when $0.5 > G_{10}$, $1 - \tau_i > G_{10}$, or $\tau_i > G_{10}$. The baseline in Chapter 3 is $\tau_i > G_{10}$. We report the average level of cooperation among prosocial (a) and proself (b) agents as well as for the collective (c). The degree of prosociality segregation in the in the group context is reported in d. Parameter settings: $m = 5$; $l = 0.5$; $PA = 0.4$; $r = 0.05$.

A7. Sensitivity analysis: Chances of network dyad selection

In this section, we inspect if the homophily solution still works when we vary the chances of dyad selection. Parameter *r* was set to 0.05, reflecting a 5% chance of being selected as a dyad to play the 2-person PD in the social network.

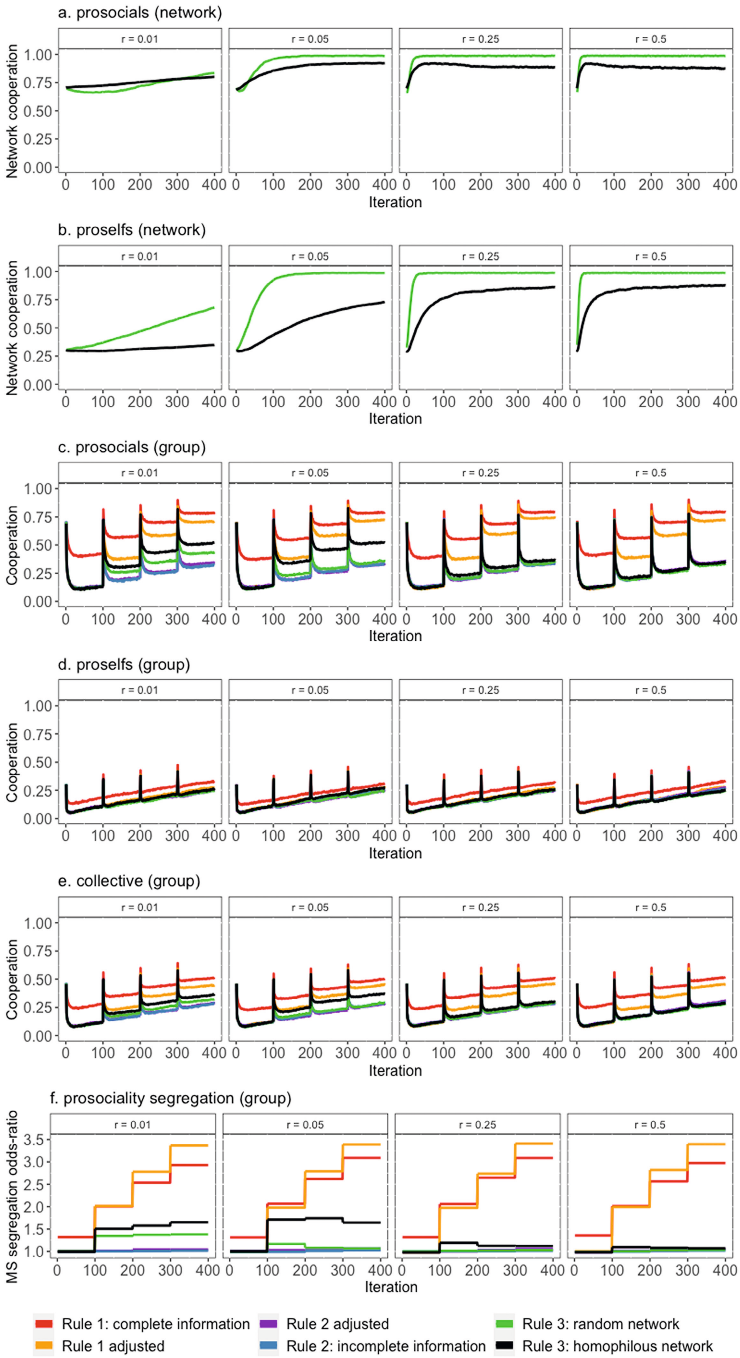


Figure A10: Impact of the chances of network dyad selection (r) on network cooperation (a-b), cooperation in the group (c-e), and prosociality segregation in the group context (f), separated by matching rules. The chances for dyad selection in the network differ, ranging from the lower bound (r = 0.01) to the upper bound (r = 0.50). Parameter settings: m = 5; l = 0.5; PA = 0.4.

We question if agents can still differentiate in the social network between prosocials and proselfs when the chances of interaction are slimmer or higher (***r*** = {0.01, 0.05, 0.25, 0.5}), all things being equal. The upper bound is ***r*** = 0.5 because a dyad consists of two agents. We run 100 independent simulations per condition. 6 x 4 conditions – matching and ***r***, respectively – realize 2400 simulation runs. We replicate all figures reported in Chapter 3 but separate per value of ***r***.

We deduce from Figure A10 that it becomes harder to differentiate between prosocials and proselfs in the social network when the chances for dyad selection increase. Prosocial agents cooperate more (Figure A10a), albeit somewhat lower in homophilous social networks. This makes sense because cooperation cannot cascade freely through the network but is restricted via behavioral homophily. A prosocial defector, surrounded by cooperators, may not be selected as an interaction partner when behavioral homophily is allowed. Proselfs learn more quickly to cooperate when chances for dyad selection increase, facilitating possibilities to learn how to cooperate (Figure A10b).

Our network and homophily solution for the bad barrels problem only works when agents can readily differentiate between types. Rule 3 with homophily is still more beneficial than other incomplete information rules when $r = 0.25$ (Figure A10c-e). However, the benefits of homophilous networks and information derived thereof disappear completely when $r = 0.5$. When r increases, the complete information matching rules are the best solution to increase cooperation. Prosociality segregation decreases from approximately 1.60 to 1.05, $r = 0.05$ and $r = 0.5$, respectively (Figure A10f).

A8. Sensitivity analysis: Impact of behavioral homophily

Here, we test how robust our findings are when behavioral homophily is implemented or not. We noted in Chapter 3 how structural homophily is implemented in the network. Prosocials are more likely to form ties to prosocials than to proselfs. The network solution is rather probabilistic, allowing for ties between proselfs and prosocials. We also implemented homophilous tendencies in 2-person PDs. Namely, behavioral homophily governs the interaction of agents. Cooperators only interact with other cooperators, leaving defectors to interact with other defectors. The only option for defectors to interact with cooperators is to change behavior from defection to cooperation. We inspect the consequences of removing behavioral homophily as a rule affecting who interacts with whom in the network. As such, we inspect the upper boundary of a favoring behavioral homophily rule protecting cooperators from exploitation by defectors and a lower boundary in which all linked agents can interact with one another. We ran 100 independent simulation runs per condition, matching rule 3 (homophily) times 2 (behavioral homophily), realizing 200 runs. The findings of this robustness check are visualized in Figure A11.

The model outcomes reported in Chapter 3 are fairly robust, qualitatively speaking. Yet, we find quantitively that our homophily solution fares slightly worse when behavioral homophily is not implemented. Prosocials still benefit when information from a homophilous network is paired with information from the group context, but not as well as when behavioral homophily is allowed (Figure A11c). An important reason prosocials cannot harness their cooperative potential is the inability to differentiate between proselfs and prosocials regarding network cooperation. Figures A11a and b indicate that both types quickly learn to cooperate when behavioral homophily is not implemented. The inability to identify similar others then affects the chances of grouping with similar others (Figure A11f). The homophily solution is thus less effective when behavioral homophily does not govern network interactions.

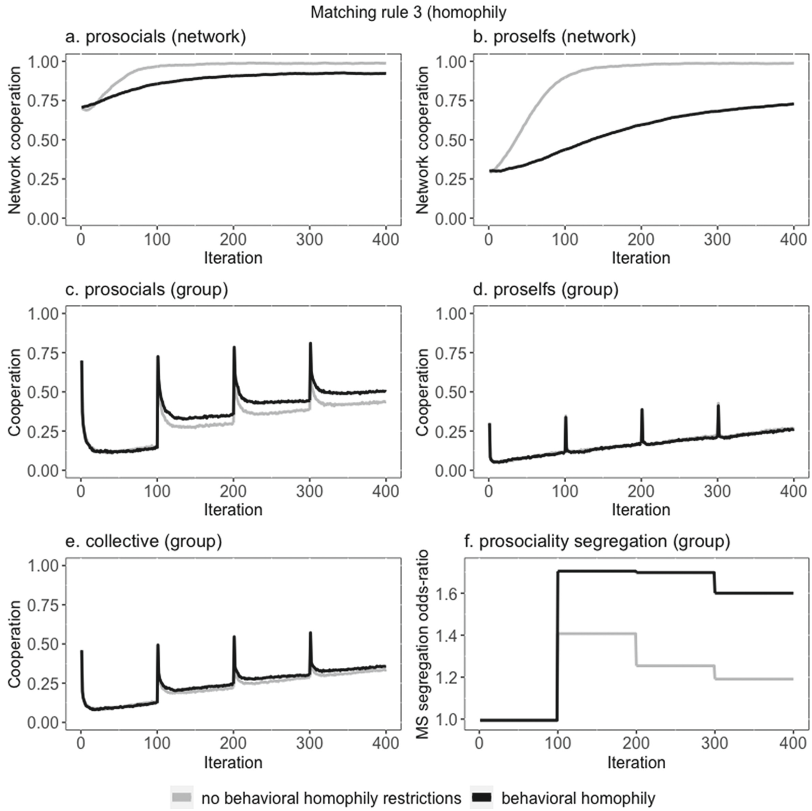


Figure A11: Impact of behavioral homophily on network cooperation (a-b), cooperation in the group (c-e), and prosociality segregation in the group context (f), separated by matching rules. Parameter settings: $m = 5$; $l = 0.5$; $PA = 0.4$; $r = 0.05$.

A9. Potential model extensions

Our model leaves room for several model extensions. We report seven potential avenues for extensions without specifying the in-depth model adjustment. First, we aimed to build on the current literature on meritocratic matching by studying prosocial and proself types and fixed next to adaptive thresholds to resemble individual and situational (i.e., group and network) influences. In the future, however, we may want to study more heterogeneity in the distribution of thresholds. For instance, agents could randomly draw an initial threshold from a normal, uniform, or polarized distribution (Macy & Evtushenko, 2020).

Second, initial thresholds can be affected by long-term learning effects in which one gradually adapts thresholds accordingly. Positive (negative) outcomes may gradually lower (increase) one's initial threshold. Agents may, for example, lower their initial threshold to a new initial threshold ($\tau_{i,new}$) when $\tau_{i,t}$ is lower than τ_i.[1] Contrarily, ($\tau_{i,new}$ increases when $\tau_{i,t}$ is larger than τ_i. For example, if a prosocial has a $\tau_{i,t} = 0.9$, then ($\tau_{i,new} \approx 0.5$. This extension would lead to an all-out adoption of the learning perspective.

Third, prosocials' fixed initial thresholds play an important role after joining a new group, but adaptive thresholds may incorporate some fixedness. For example, prosocials may be more likely to cooperate throughout the game than proselfs.[2] Typically, prosocials cooperate at a probability of 0.18 when $\tau_{i,t} = 0.3$, $k_t = 0$, and $m = 5$ compared to 0.26 with an added fixed prosociality effect of, for example, $a_i = 0.1$. Practically, including $a_i = 0.1$ would shift the logistic function structurally upwards for prosocials. Then we have a reversed ceiling effect.

Fourth, a step towards adding more matching dynamics can be to rely on a-synchronous matching in which agents can try to leave and join groups on their own accord. This resembles situations in real life when individuals quarrel with others in their group, resulting in some wanting to leave. Still, a-synchronous matching requires adjusting other fixed parts of the ABM, e.g., deviating from fixed group sizes and centralized leave-stay moments, while requiring a group acceptance and exclusion procedure. We leave this for future model considerations.

Fifth, including network information incorporates asymmetry of information. Some have local information about others' social network efforts, while others only have group information. But asymmetry can also occur within and between matching rules. Some agents may have complete information about a subset of agents' individual merits, while others have incomplete information about another subset. Even a tiny piece of individual information shows to ease the detrimental effect of group merit on cooperation in general (Nax et al., 2015a).

Sixth, homophily shows to buffer the bad barrels problem, but we can also envision finding the 'sweet spot' of homophily. We could investigate the upper and lower level of homophily required to aid spoiled prosocial agents. A consequence of this extension would be that we depart from the probabilistic nature of the random spatial graph algorithm, allowing us to vary the degree of homophily, e.g., from 0 (no homophily) to 1 (full homophily).

Seventh, agents may have different dyadic information per partner. Some may have a positive view of an agent, while others perceive the agent as a defector. Agents would then have different tabs about their network partners. Again, the consequence for the model may be severe as this extension may create peculiar dynamics of some not wanting an agent in the group while others do. We would then need to design an acceptance rule.

1 $\tau_{i,new} = \tau_i / \{\tau_i + exp[-(\tau_i + \tau_{i,t})]\}$

2 $p_{i,t}(c_{i,t} = 1) = 1 / \{1 + exp[m(\tau_{i,t} - k_t - a_i)]\}$

B. APPENDICES TO CHAPTER 4

Table B1: SAOM findings for friendship network selection.

	Model 1		Model 2	
	estimate *(SE)*	**RI**	**estimate *(SE)***	**RI**
Rate parameter				
1. Friendship rate $t = 1 \rightarrow t = 2$	9.46** (0.65)		12.20** (1.28)	
Structural network effects				
2. Out-degree (density)	−2.83** (0.31)	21.0%	−3.08** (0.22)	21.5%
3. Out-degree activity	0.05** (0.02)	4.7%	0.05** (0.01)	5.0%
4. In-degree popularity	−0.03 (0.03)	3.3%	−0.02 (0.03)	2.6%
5. In-degree activity	−0.31** (0.06)	12.2%	−0.24** (0.05)	10.9%
6. Reciprocity	3.03** (0.29)	11.5%	2.82** (0.23)	11.7%
7. Transitive triplets	0.60** (0.07)	12.7%	0.48** (0.07)	10.8%
8. Transitive reciprocated triplets	−0.34** (0.09)	5.1%	−0.33** (0.09)	4.8%
Cross-network effects				
9. Cooperation relation			0.62* (0.28)	2.9%
10. Reciprocated PFC tie			−0.45 (0.44)	0.1%
11. PFC popularity			0.16** (0.05)	3.3%
Individual attribute effects				
12. Female (ref = male) activity	−0.30* (0.15)	1.7%	−0.32** (0.12)	1.9%
13. Female (ref = male) popularity	−0.34* (0.14)	3.1%	−0.27* (0.13)	2.5%
14. Grades activity	0.31** (0.09)	6.2%	0.22** (0.06)	4.6%
15. Grade popularity	0.10 (0.06)	2.8%	0.06 (0.05)	1.9%
16. Agreeableness activity				
17. Agreeableness popularity				
18. Extraversion activity				
19. Extraversion popularity				
20. Conscientiousness activity				
21. Conscientiousness popularity				
22. Neuroticism activity				
23. Neuroticism popularity				

Model 3		Model 4	
estimate *(SE)*	RI	estimate *(SE)*	RI
11.88** (1.68)		10.70** (1.53)	
−3.09** (0.33)	20.8%	−2.99** (0.26)	17.4%
0.05** (0.02)	4.8%	0.05** (0.02)	4.0%
−0.02 (0.03)	2.2%	−0.05 (0.03)	4.5%
−0.26** (0.06)	10.9%	−0.26** (0.06)	9.5%
2.85** (0.30)	11.3%	2.77** (0.23)	9.5%
0.48** (0.08)	10.6%	0.49** (0.07)	9.4%
−0.32** (0.10)	4.6%	−0.31** (0.09)	3.8%
0.63 (0.33)	2.8%	0.64* (0.31)	2.3%
−0.45 (0.45)	0.9%	−0.40 (0.44)	0.7%
0.16** (0.05)	3.2%	0.15** (0.05)	2.4%
−0.35* (0.15)	2.0%	−0.48** (0.14)	2.1%
−0.28* (0.13)	2.4%	−0.34* (0.14)	2.3%
0.23** (0.07)	4.7%	0.27** (0.08)	4.3%
0.06 (0.06)	1.9%	0.12 (0.07)	2.6%
−0.20 (0.10)	1.0%	−0.26* (0.12)	1.1%
−0.10 (0.11)	0.9%	−0.19 (0.12)	1.2%
		0.05 (0.11)	0.3%
		−0.01 (0.10)	<0.0%
		0.06 (0.07)	0.4%
		−0.03 (0.07)	0.3%
		−0.05 (0.09)	0.3%
		−0.10 (0.09)	0.9%

[continued on next page]

A

Table B1: *[continued]*

	Model 1		Model 2	
	estimate *(SE)*	**RI**	**estimate *(SE)***	**RI**
24. Openness activity				
25. Openness popularity				
Homophily effects				
26. Gender	0.68** (0.14)	7.6%	0.64** (0.12)	7.5%
27. Grades	2.34** (0.58)	8.3%	1.95** (0.48)	7.3%
28. Agreeableness				
29. Extraversion				
30. Conscientiousness				
31. Neuroticism				
32. Openness				

Note. *SE* = standard error; RI = relative importance; Estimates, SEs, and RI percentages are rounded.
* $p < 0.05$; ** $p < 0.01$.

Table B2: SAOM findings for selection dynamics in the preference-for-collaboration (PFC) network.

	Model 1		Model 2	
	estimate *(SE)*	**RI**	**estimate *(SE)***	**RI**
Rate parameter				
1. Friendship rate $t = 1 \rightarrow t = 2$	10.19** (0.76)		14.83** (1.36)	
Structural network effects				
2. Out-degree (density)	−2.68** (0.28)	19.8%	−2.67** (0.22)	20.3%
3. Out-degree activity	0.03 (0.02)	3.6%	0.03* (0.01)	3.9%
4. In-degree popularity	−0.02 (0.03)	2.7%	−0.02 (0.03)	3.3%
5. In-degree activity	−0.21** (0.05)	9.8%	−0.18** (0.04)	9.8%
6. Reciprocity	2.79** (0.24)	12.2%	2.50** (0.19)	11.3%
7. Transitive triplets	0.52** (0.06)	14.0%	0.43** (0.05)	11.9%
8. Transitive reciprocated triplets	−0.35** (0.08)	6.4%	−0.33** (0.08)	6.0%
Cross-network effects				
9. Friends			0.84** (0.20)	4.2%
10. Reciprocated friends			−0.09 (0.32)	0.2%
11. Friendship popularity			0.09* (0.05)	2.1%

	Model 3		Model 4	
	estimate *(SE)*	RI	estimate *(SE)*	RI
			0.47* (0.11)	2.2%
			0.38** (0.11)	3.2%
	0.70** (0.14)	7.7%	0.72** (0.13)	6.2%
	1.97** (0.51)	7.0%	1.99** (0.65)	5.3%
	0.19 (0.40)	0.5%	0.57 (0.42)	1.2%
			0.60 (0.45)	1.2%
			0.32 (0.37)	0.8%
			0.02 (0.56)	<0.0%
			0.29 (0.31)	0.8%

	Model 3		Model 4	
	estimate *(SE)*	RI	estimate *(SE)*	RI
	14.61** (2.27)		14.86** (1.70)	
	−2.68** (0.26)	19.7%	−2.60** (0.25)	16.6%
	0.03 (0.01)	3.2%	0.03* (0.01)	3.2%
	−0.02 (0.02)	5.6%	−0.05* (0.02)	5.6%
	−0.19** (0.06)	8.3%	−0.18** (0.05)	8.3%
	2.50** (0.21)	11.0%	2.48** (0.21)	9.5%
	0.43** (0.05)	11.6%	0.45** (0.06)	10.3%
	−0.33** (0.07)	5.8%	−0.33** (0.07)	5.3%
	0.85** (0.25)	4.1%	0.88** (0.19)	3.5%
	−0.09 (0.38)	0.2%	−0.19 (0.33)	0.2%
	0.09 (0.04)	2.0%	0.08 (0.04)	1.4%

[continued on next page]

A

Table B2: *[continued]*

	Model 1		Model 2	
	estimate *(SE)*	**RI**	**estimate *(SE)***	**RI**
Individual attribute effects				
12. Female (ref = male) activity	−0.30* (0.13)	1.8%	−0.22* (0.11)	1.4%
13. Female (ref = male) popularity	−0.32** (0.11)	3.2%	−0.29** (0.11)	2.9%
14. Grades activity	0.32** (0.08)	6.6%	0.23** (0.06)	5.0%
15. Grades popularity	0.12* (0.06)	3.6%	0.11* (0.05)	3.6%
16. Agreeableness activity				
17. Agreeableness popularity				
18. Extraversion activity				
19. Extraversion popularity				
20. Conscientiousness activity				
21. Conscientiousness popularity				
22. Neuroticism activity				
23. Neuroticism popularity				
24. Openness activity				
25. Openness popularity				
Homophily effects				
26. Gender	0.66** (0.12)	7.7%	0.47** (0.10)	5.8%
27. Grades	2.25** (0.46)	8.5%	2.15** (0.42)	8.5%
28. Agreeableness				
29. Extraversion				
30. Conscientiousness				
31. Neuroticism				
32. Openness				

Note. *SE* = standard error; RI = relative importance; Estimates, SEs, and RI percentages are rounded.
* $p < 0.05$; ** $p < 0.01$.

Model 3		Model 4	
estimate *(SE)*	**RI**	**estimate *(SE)***	**RI**
−0.26* (0.13)	1.6%	−0.29** (0.13)	1.5%
−0.30* (0.12)	2.9%	−0.34* (0.12)	2.5%
0.24** (0.07)	5.0%	0.24** (0.06)	4.2%
0.11* (0.06)	3.4%	0.17** (0.06)	4.0%
−0.19 (0.10)	0.8%	−0.24* (0.12)	1.1%
−0.09 (0.10)	1.0%	−0.14 (0.11)	1.1%
		0.09 (0.09)	0.5%
		−0.11 (0.09)	1.0%
		0.01 (0.06)	0.1%
		0.05 (0.06)	0.7%
		−0.03 (0.07)	0.2%
		0.04 (0.07)	0.4%
		0.19* (0.08)	1.3%
		0.35** (0.09)	3.5%
0.52** (0.11)	6.0%	0.54** (0.12)	5.2%
2.18** (0.51)	8.1%	2.22** (0.45)	6.5%
−0.04 (0.35)	0.1%	0.08 (0.31)	0.2%
		0.39 (0.38)	0.9%
		0.12 (0.29)	0.3%
		0.43 (0.33)	1.0%
		−0.00 (0.26)	<0.0%

C. APPENDICES TO CHAPTER 5

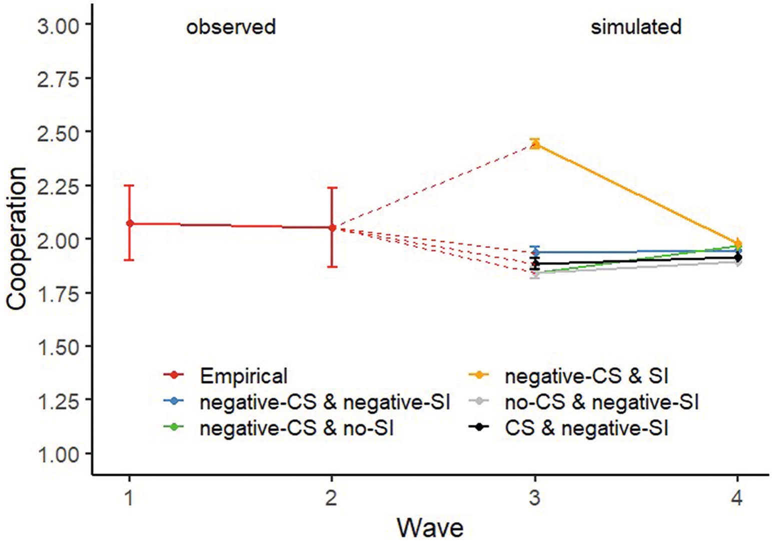


Figure C1: Visualizing the mean and 95% confidence intervals of cooperation levels generated via 50 simulation runs per condition in which cooperation selection (0.66) and social influence (3.54) take negative values. The empirical distribution of cooperation is included in wave 1 and 2 (red). Dashed red lines are solely used to link the empirical outcome at wave 2 to the starting point of the simulations in wave 3. *Note*. SI = social influence; CS = cooperation selection; negative = negative estimate value.

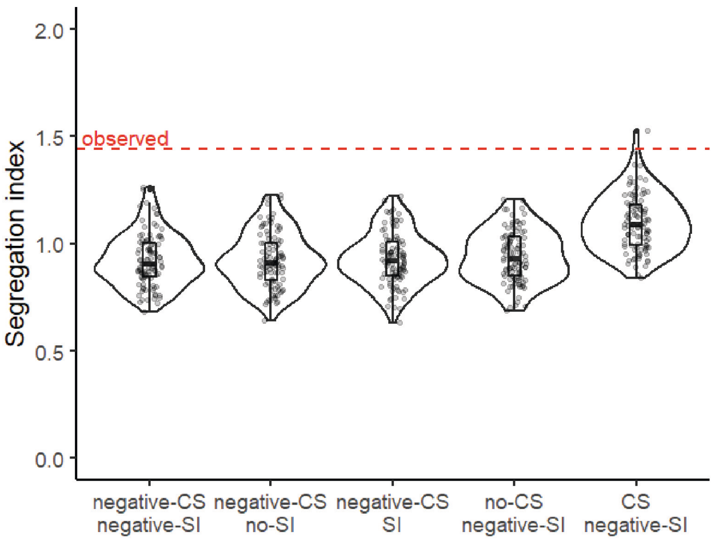


Figure C2: Visualizing network segregation by cooperation in waves = 3 and 4 combined per condition in which cooperation selection (0.66) and social influence (3.54) can take negative values. 50 simulation runs per condition are shown in the violin plots, boxplots, and via jittered data points. The dashed red line is the segregation level (1.44) found empirically at wave 2. *Note*. SI = social influence; CS = cooperation selection; negative = negative estimate value.

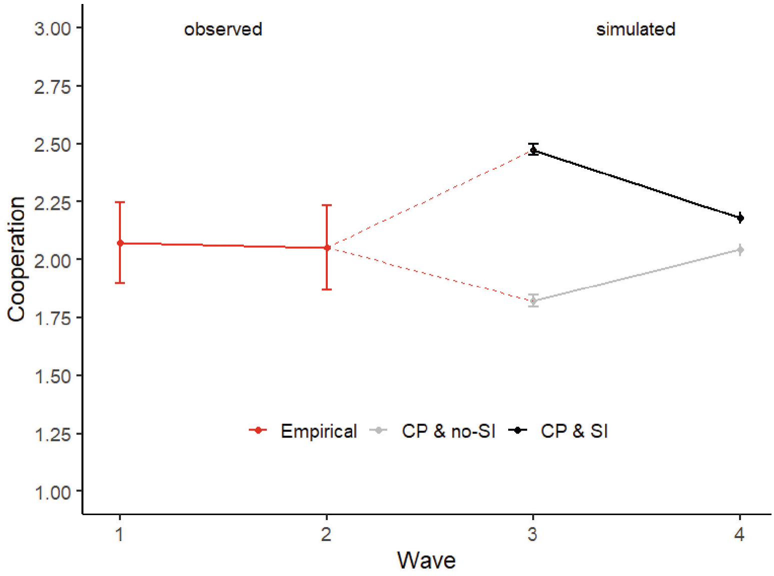


Figure C3: Visualizing the mean and 95% confidence intervals of cooperation levels generated via 50 simulation runs per condition in which cooperation popularity (0.06) is explored. Social influence (3.54) is present or not. The empirical distribution of cooperation is included in wave 1 and 2 (red). Dashed red lines are solely used to link the empirical outcome at wave 2 to the starting point of the simulations in wave 3. *Note*. SI = social influence; CP = cooperation popularity.

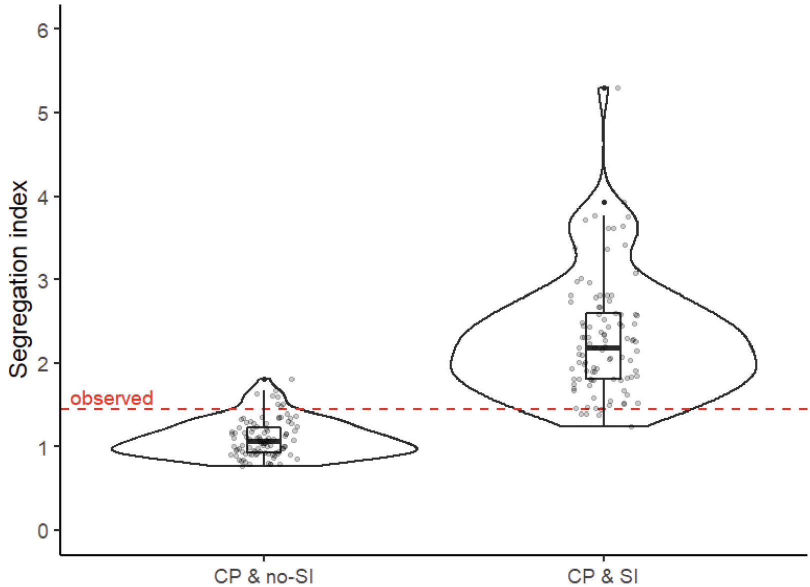


Figure C4: Visualizing network segregation by cooperation in waves = 3 and 4 combined per condition in which cooperation popularity (0.06) plays a central role. Social influence (3.54) is present or not. 50 simulation runs per condition are shown in the violin plots, boxplots, and via jittered data points. The dashed red line is the segregation level (1.44) found empirically at wave 2. *Note*. SI = social influence; CP = cooperation popularity.

A

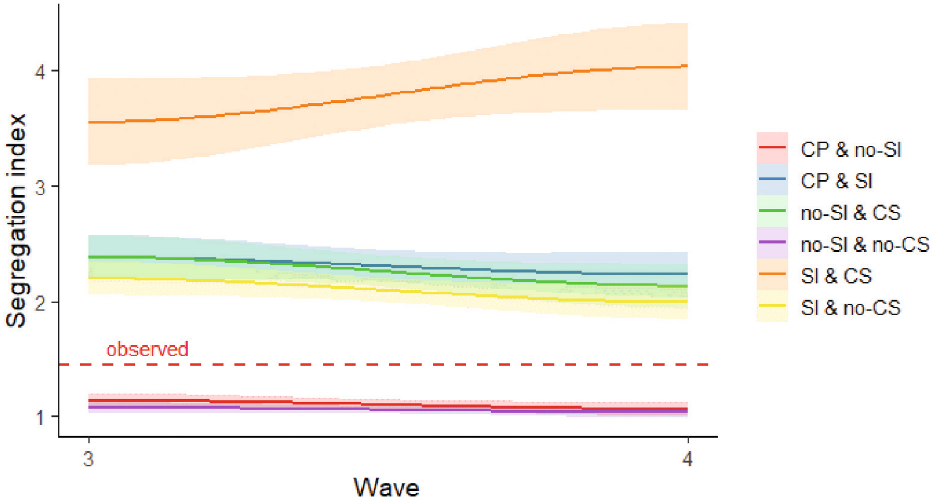


Figure C5: Visualizing network segregation by cooperation in wave 3 and 4 for conditions with CS, CP, and SI (excluding negative estimates). Mean and 95% confidence intervals are shown per condition. The dashed red line is the segregation level (1.44) found empirically at wave 2. *Note*. SI = social influence; CS = cooperation selection; CP = cooperation popularity.

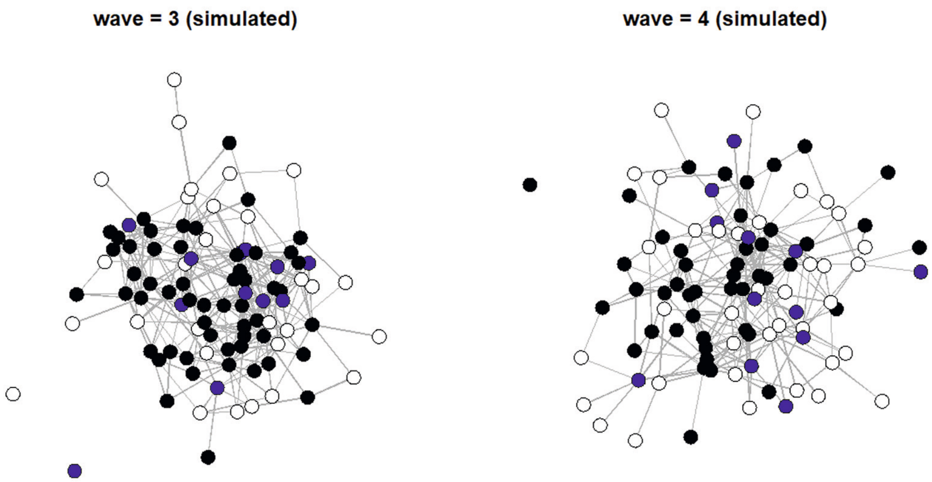


Figure C6: Visualizing two simulated networks of 95 agents. White nodes are defectors, blue are neutral ones, and black nodes are cooperators. The simulated networks are generated with CS set = 0 and SI = 3.54. Segregation slightly increased from 1.61 (wave 3) to 1.80 (wave 4). Average level of cooperation decreased from 2.26 (wave 3) to 2.09 (wave 4).

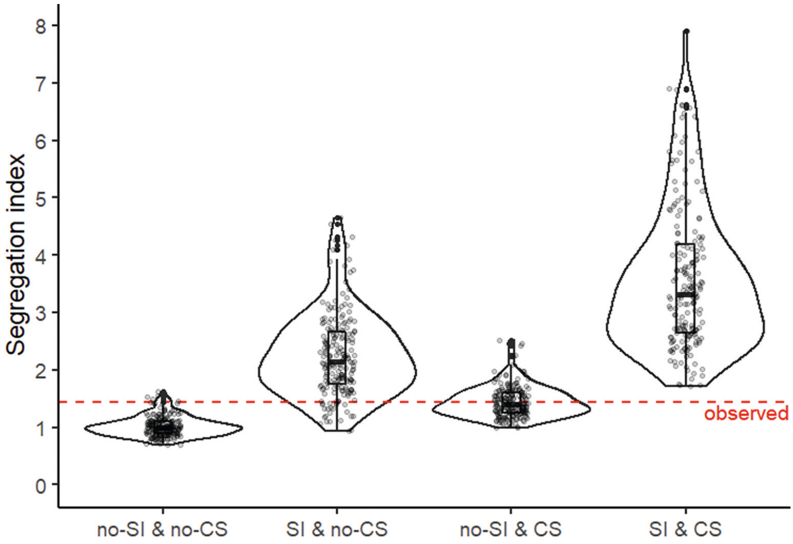


Figure C7: Visualizing network segregation by cooperation in waves = 3 and 4 combined. Input data from reversed cooperation levels are used. 100 simulation runs per condition are shown in the violin plots, box-plots, and via jittered data points. The dashed red line is the segregation level without reversed data (1.44) found empirically at wave 2. *Note*: SI = social influence; CS = cooperation selection.

A

D. APPENDICES TO CHAPTER 6

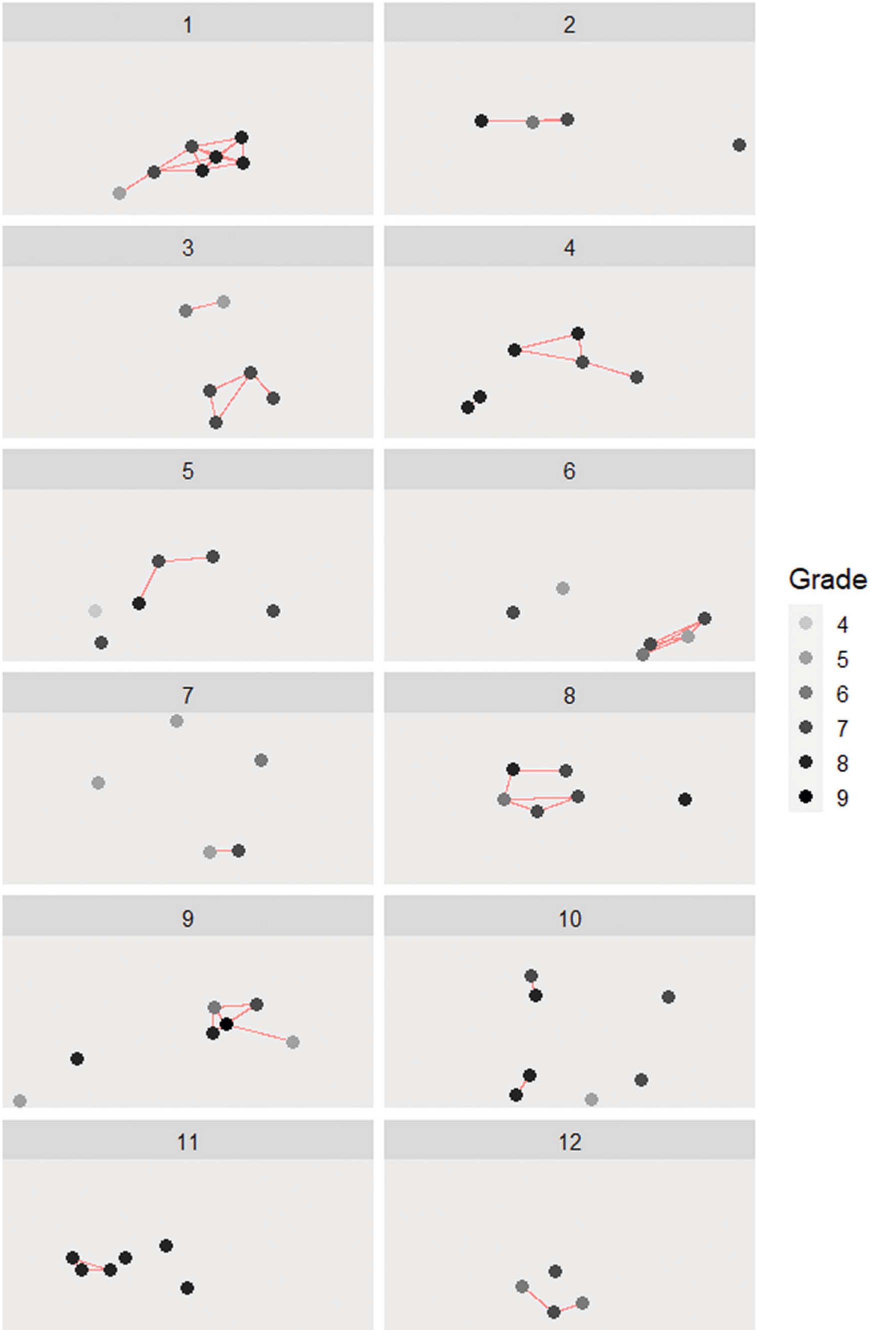


Figure D1: Visualizing familiarity relationships, gender, and grades per team. A light-red line between nodes refers to a present familiarity tie. A darker color refers to a higher grade of the student.

Table B1: Overview of the goodness of fit (GOF) figures and key results.

Figure number	Variable measured	A short explanation of the GOF
D2	Count of teams and team sizes	The data comprises 1 team with 5 students and 6 teams with 6 students. The observed number of teams comprising 4, 5, 6, or 7 students fits well with the simulated team sizes. The simulations show an over (size 5) and underestimation (size 6) of both empirical distributions, but the observed value remains in the simulation data interval.
D3	Distribution of grade similarity	The GOF plots show that including similarity –dyadically or on the team level – fits the opposite similarity index. The range (team similarity) is of interest for Model 1 in which dyadic similarity is included. The count of pairs (dyadic similarity) is estimated for Model 2, which includes team similarity in the ERPM.
D4	Intraclass correlation (ICC) of PFC popularity	Empirically, we have an ICC of 0.26 for PFC popularity. The simulated average in both models lies near the observed value.
D5	Number of grade pairs	We estimate the count of pairs with different grades and show the number of pairs of students with grades 4-and-5, 6, 7, or 8-and-9. Most have a grade of 7, and the number of pairs is empirically higher than the rest. The simulations capture this pattern to a satisfactory degree.
D6	Number of gender combinations	Students can be part of a male-male, female-female, or female-male pair. We compared the number of gender combinations empirically with the simulated data. We find that the fit is good. The ERPMs capture the tendency that there are more female-female pairs than male-male pairs. The results also show that cross-homophily pairs remain present empirically and in the simulated data.
D7	Number of friends	The average number of friends each student has in the partition is 3.67 (88 nominations divided by 12 and then divided by two because each tie consists of two students). The GOF plots show that both models capture the average number of friends in the partition well.
D8	Number of familiarity ties	The average number of familiar students each student has in the partition is 1.08 (26 nominations divided by 12, and then divided by two because each tie consists of two students). The GOF plots show that both models well capture the average number of familiar students in the partition.

Note. We present 7 goodness of fit (GOF) plots of Model 1 and Model 2. The plots show the distribution of simulated data (500 simulation runs per model) via ERPMs in violin plots, with the average reflected in a black dot (observed data = red diamond sign).

A

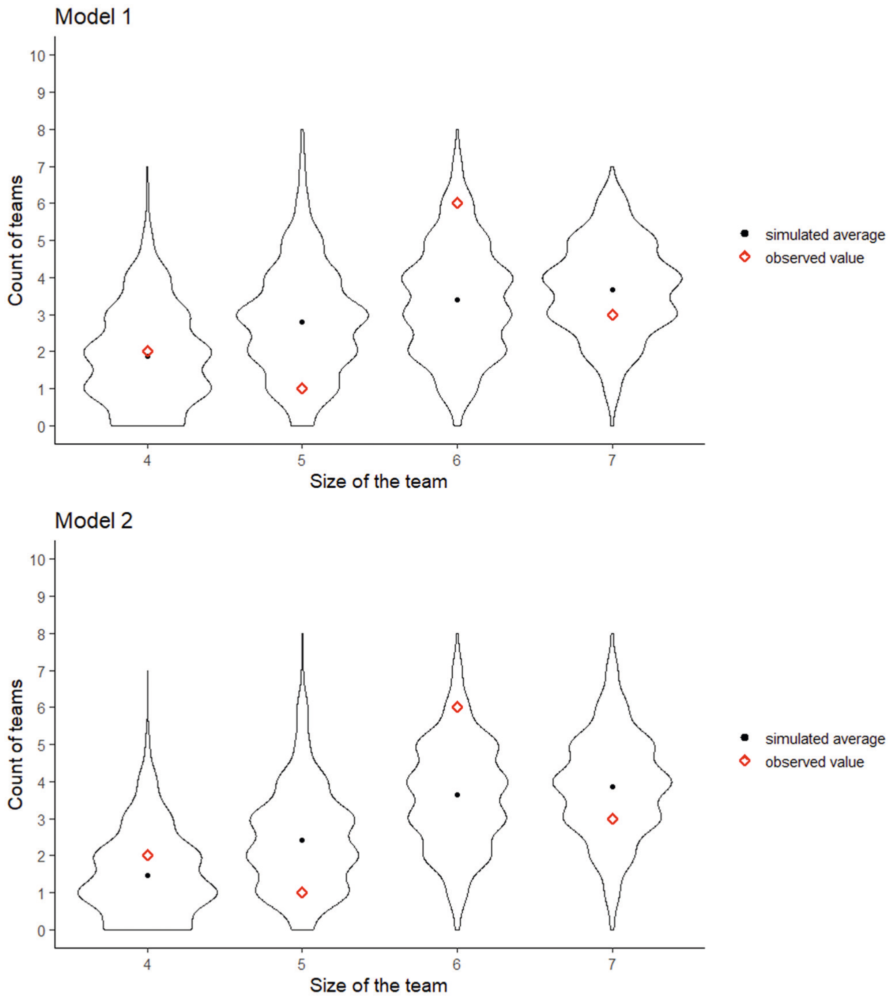


Figure D2: Count of teams per different size of the team. The violin plots represent the different distributions of the simulated data. The black dot shows the average value of the simulations. The red diamond shape shows the observed value that we have in our data.

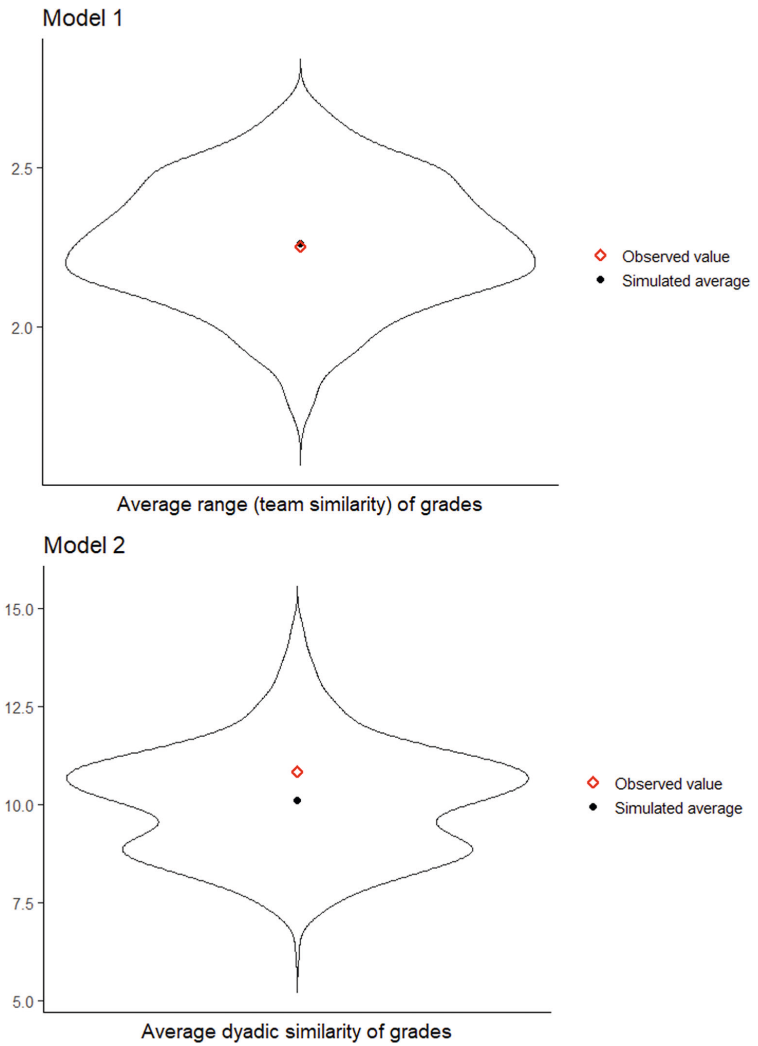


Figure D3: The average number of similarities on the team (Model 1) and dyadic (Model 2) level is shown here. Model 1 includes the range of grades whereas Model 2 measures the average number of similar pairs. The violin plots represent the different distributions of the simulated data. The black dot shows the average value of the simulations. The red diamond shape shows the observed value that we have in our data.

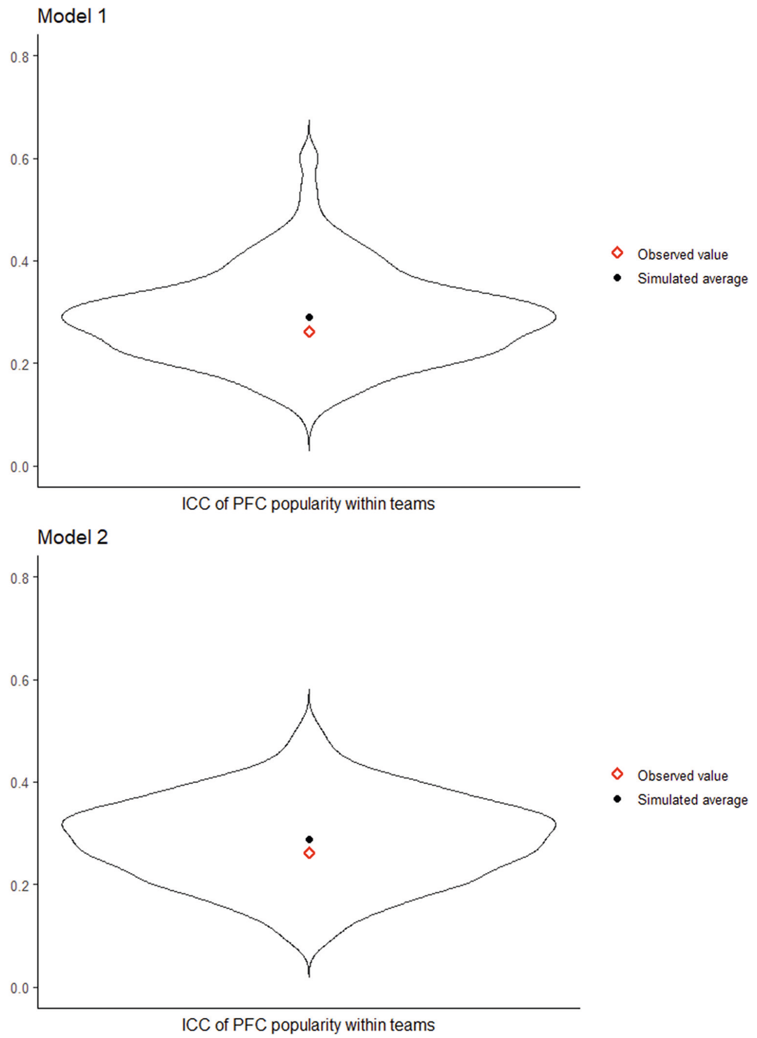


Figure D4: The distribution of the intraclass correlation of PFC popularity is shown here. The violin plots represent the different distributions of the simulated data. The black dot shows the average value of the simulations. The red diamond shape shows the observed value that we have in our data.

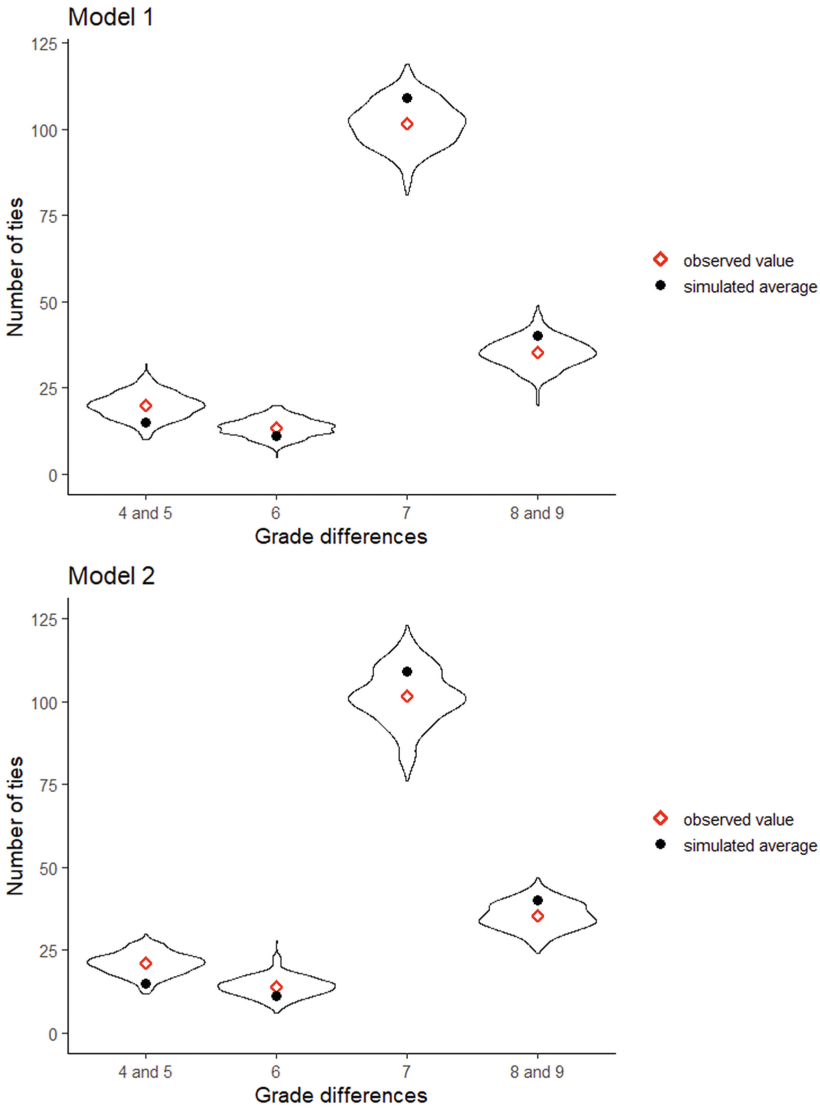


Figure D5: Showing differences and similarities in grades in the partition. The violin plots represent the different distributions of the simulated data. The black dot shows the average value of the simulations. The red diamond shape shows the observed value that we have in our data.

A

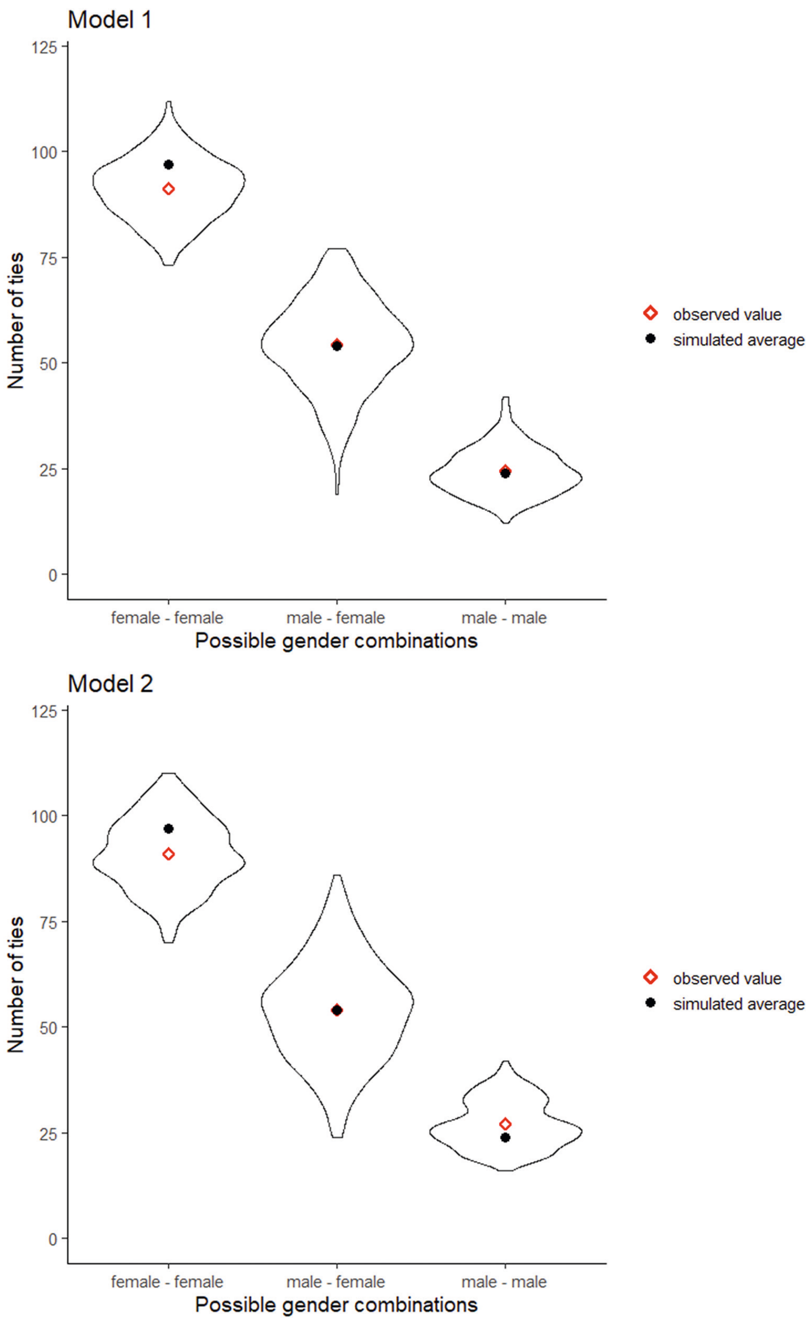


Figure D6: Exploring the count of gender combinations in the partition. The violin plots represent the different distributions of the simulated data. The black dot shows the average value of the simulations. The red diamond shape shows the observed value that we have in our data.

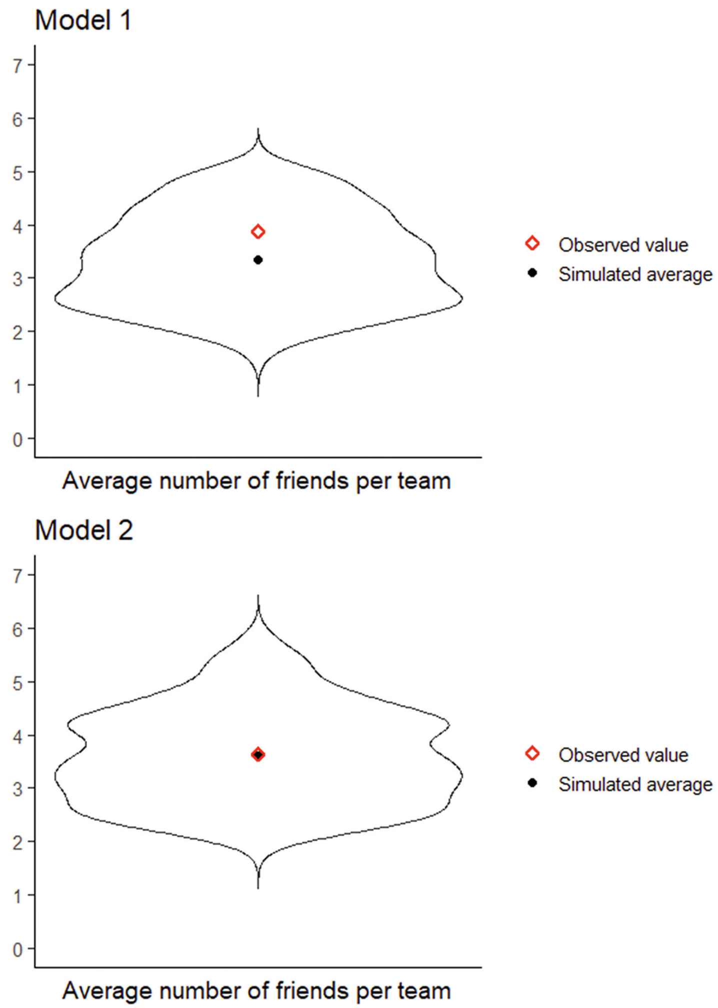


Figure D7: Average number of friends per team. The violin plots represent the different distributions of the simulated data. The black dot shows the average value of the simulations. The red diamond shape shows the observed value that we have in our data.

A

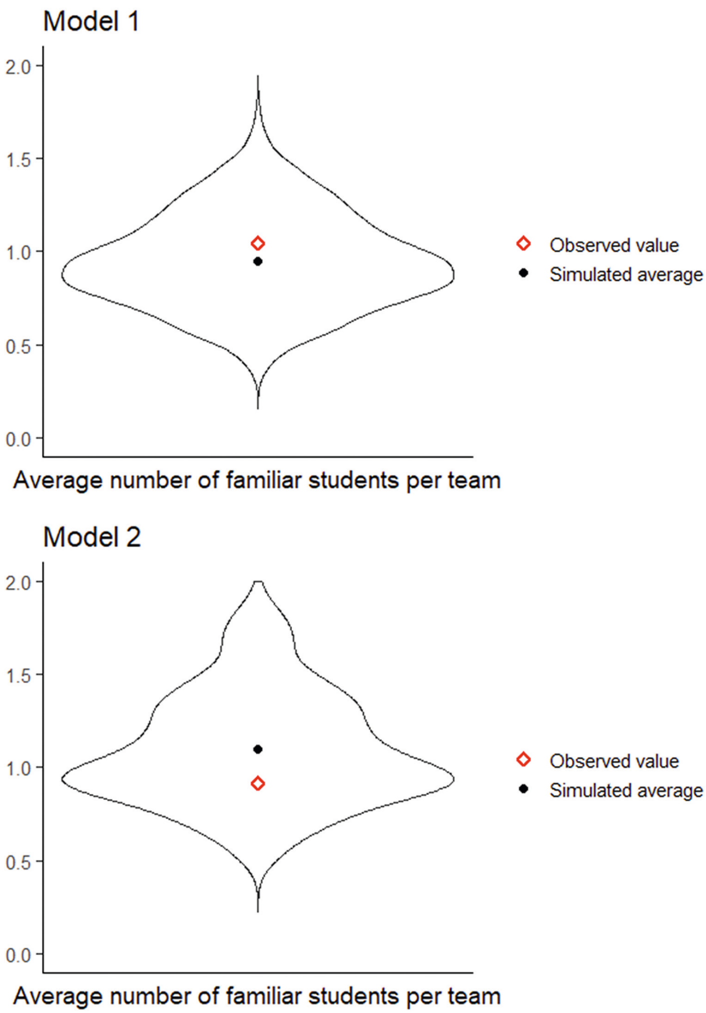


Figure D8: This plot shows the average number of familiarity ties per team. The violin plots represent the different distributions of the simulated data. The black dot shows the average value of the simulations. The red diamond shape shows the observed value that we have in our data.

A

REFERENCES

Ackermann, K. A., Fleiß, J., & Murphy, R. O. (2016). Reciprocity as an individual difference. *Journal of Conflict Resolution, 60*(2), 340–367. doi: 10.1177/0022002714541854

Ackermann, K. A., & Murphy, R. O. (2019). Explaining cooperative behavior in public goods games: How preferences and beliefs affect contribution levels. *Games,10*(1), 15. doi: 10.3390/g10010015

adams, j., & Schaefer, D. R. (2016). How initial prevalence moderates network-based smoking change: Estimating contextual effects with stochastic actor-based models. *Journal of Health and Social Behavior, 57*(1), 22–38. doi: 10.1177/0022146515 627848

Apicella, C. L., & Silk, J. B. (2019). The evolution of human cooperation. *Current Biology, 29* (11), R447–R450. doi: 10.1016/j.cub.2019.03.036

Apicella, C. L., Marlowe, F. W., Fowler, J. H., & Christakis, N. A. (2012). Social networks and cooperation in hunter-gatherers. *Nature, 481*(7382), 497–501. doi: 10.1038/nature10736

Arnett, J. J. (2007). Emerging adulthood: What is it, and what is it good for? *Child Development Perspectives, 1*(2), 68–73. doi: 10.1111/j.1750-8606.2007.00016.x

Asendorpf, J. B., & Wilpers, S. (1998). Personality effects on social relationships. *Journal of Personality and Social Psychology, 74*(6), 1531–1544. doi: 10.1037/0022-3514.74.6.1531

Attari, S., Krantz, D., & Weber, E. (2014). Reasons for cooperation and defection in real-world social dilemmas. *Judgment and Decision Making, 9*(4), 316–334. doi: 10.1017/S1930297500006197

Au, W., & Kwong, J. (2004). Measurements and effects of social-value orientation in social dilemmas: A review. In R. Suleiman, D. Budescu, I. Fischer, & D. Messick (Eds.), *Contemporary psychological research on social dilemmas* (pp. 71–98). New York, NY: Cambridge University Press.

Axelrod, R. (1997). The dissemination of culture: A model with local convergence and global polarization. *Journal of Conflict Resolution, 41*(2), 203–226. doi: 10.1177/0022002797041002001

Axelrod, R. (1984). *The evolution of cooperation*. New York, NY: Basic Books.

Bacharach, M., & Gambetta, D. (2001). Trust in signs. In K. Cook (Ed.), *Trust in society* (pp. 148–184). New York, NY: Russel Sage Foundation.

Barańczuk, U. (2019). The Five Factor Model of personality and social support: A meta-analysis. *Journal of Research in Personality, 81*, 38–46. doi: 10.1016/j.jrp.2019.05.002

Bailey, J. L., & Skvoretz, J. (2017). The social-psychological aspects of team formation: New avenues for research. *Sociology Compass, 11*(6), e12487. doi: 10.1111/soc4.12487

Bakker, D. M., & Dijkstra, J. (2021). Comparing the slider measure of social value orientation with its main alternatives. *Social Psychology Quarterly, 84*(3), 235–245,doi: 10.1177/01902725211008938

Bakker, D. M. (2019). *Cooperation and social control: effects of preferences, institutions, and social structure* (Doctoral dissertation, University of Groningen). doi: 10.33612/diss.98552819

Baldassarri, D. (2015). Cooperative networks: Altruism, group solidarity, reciprocity,and sanctioning in ugandan producer organizations. *American Journal of Sociology, 121*(2), 355–395. doi: 10.1086/682418

Balliet, D., Wu, J., & De Dreu, C. K. W. (2014). Ingroup favoritism in cooperation: A meta-analysis. *Psychological Bulletin, 140*(6), 1556–1581. doi: 10.1037/a0037737

Balliet, D. P., Parks, C. D., & Joireman, J. (2009). Social value orientation and cooperation in social dilemmas: A meta-analysis. *Group Processes and Intergroup Relations, 12*(4), 533–547. doi: 10.1177/1368430209105040

Bandura, A. (1977). *Social learning theory.* Englewood Cliffs, NJ: Prentice Hall.

Baumeister, R. F., Stillwell, A. M., & Heatherton, T. F. (1994). Guilt: An interpersonal approach. *Psychological Bulletin, 115*(2), 243–267. doi: 10.1037/0033-2909.115.2.243

Bell, S. T. (2007). Deep-level composition variables as predictors of team performance: a meta-analysis. *The Journal of Applied Psychology*, *92*(3), 595–615. doi: 10.1037/0021-9010.92.3.595

Bercovitz, J., & Feldman, M. (2011). The mechanisms of collaboration in inventive teams: Composition, social networks, and geography. *Research Policy*, *40*(1), 81–93. doi: 10.1016/j.respol.2010.09.008

Bianchi, F., Flache, A., & Squazzoni, F. (2020). Solidarity in collaboration networks when everyone competes for the strongest partner: A stochastic actor-based simulation model. *Journal of Mathematical Sociology, 44*(4), 249–266. doi: 10.1080/0022250x.2019.1704284

Bianchi, F., & Squazzoni, F. (2015). Agent-based models in sociology. *Wiley Interdisciplinary Reviews: Computational Statistics, 7*(4), 284–306. doi: 10.1002/wics.1356

Bicchieri, C. (1990). Norms of cooperation. *Ethics, 100*(4), 838–861. doi: 10.1086/293237

Blau, P. M. (1977). *Inequality and heterogeneity: A primitive theory of social structure*. New York, NY: Free Press.

Block, P., & Grund, T. (2014). Multidimensional homophily in friendship networks. *Network Science, 2*(2), 189–212. doi: 10.1017/nws.2014.17.

Blumenfeld, P. C., Marx, R. W., Soloway, E., & Krajcik, J. (1996). Learning with peers: From small group cooperation to collaborative communities. *Educational Researcher, 25*(8), 37–39. doi: 10.3102/0013189X025008037

Bogaert, S., Boone, C., & Declerck, C. (2008). Social value orientation and cooperation in social dilemmas: A review and conceptual model. *British Journal of Social Psychology, 47*(3), 453–480. doi: 10.1348/014466607X244970

Bowles, S., & Gintis, H. (2004). The evolution of strong reciprocity: Cooperation in heterogeneous populations. *Theoretical Population Biology, 65*(1), 17–28. doi: 10.1016/j.tpb.2003.07.001

Brouwer, J., & de Matos Fernandes, C.A. (2023). Using stochastic actor-oriented models to explain collaboration intentionality as a prerequisite for peer feedback and learning in networks. In O. Noroozi, & B. De Wever (Eds.), *The power of peer learning. Fostering students' learning processes and outcomes* (pp. 103–120). Cham, Switzerland: Springer.

Brouwer, J., de Matos Fernandes, C. A., Steglich, C. E. G., Jansen, E. P. W. A., Hofman, W. H. A., & Flache, A. (2022). The development of peer networks and academic performance in learning communities in higher education. *Learning and Instruction, 80*, 101603. doi: 10.1016/j.learninstruc.2022.101603

Brouwer, J., & Engels, M. C. (2021). The role of prosocial attitudes and academic achievement in peer networks in higher education. *European Journal of Psychology of Education*, 1–18. doi: 10.1007/s10212-020-00526-w

Brouwer, J., Flache, A., Jansen, E., Hofman, A., & Steglich, C. E. G. (2018). Emergent achievement segregation in freshmen learning community networks. *Higher Education, 76*(3), 483–500. doi: 10.1007/s10734-017-0221-2

Brouwer, J., Jansen, E. P. W. A., Flache, A., & Hofman, A. (2016). The impact of social capital on self-efficacy and study success among first-year university students. *Learning and Individual Differences, 52*, 109–118. doi: 10.1016/j.lindif.2016.09.016

Burt, R. S. (1992). *Structural holes: The social structure of competition*. Cambridge, MA: Harvard University Press.

Buskens, V., & Raub, W. (2002). Embedded trust: Control and learning. *Advances in Group Processes, 19*, 167–202. doi: 10.1016/s0882-6145(02)19007-2

Butts, C. T. (2007). Models for generalized location systems. *Sociological Methodology,37*(1), 283–348. doi: 10.1111/j.1467-9531.2006.00187.x

Byrne, D. E. (1971). *The attraction paradigm*. New York, NY: Academic Press.

Camerer, C. F. (2003). *Behavioral game theory: Experiments on strategic interaction*. Princeton, NJ: Princeton University Press.

Campbell, L. G., Mehtani, S., Dozier, M. E., & Rinehart, J. (2013). Gender-heterogeneous working groups produce higher quality science. *PLOS ONE*, *8*(10), e79147. doi: 10.1371/journal.pone.0079147

Chapman, K. J., Meuter, M., Toy, D., & Wright, L. (2006). Can't we pick our own groups? The influence of group selection method on group dynamics and outcomes. *Journal of Management Education*, *30*(4), 557–569. doi: 10.1177/1052562905284872

Chattoe-Brown, E. (1998). Just how (un)realistic are evolutionary algorithms as representations of social processes? *Journal of Artificial Societies and Social-Simulation, 1*(3), 2. doi: 10.18564/jasss.1819

Chaudhuri, A. (2011). Sustaining cooperation in laboratory public goods experiments: A selective survey of the literature. *Experimental Economics, 14*(1), 47–83. doi: 10.1007/s10683-010-9257-1

Christie, H., Munro, M., & Fisher, T. (2004). Leaving university early: exploring the differences between continuing and non-continuing students. *Studies in Higher Education, 29*(5), 617–636. doi: 10.1080/0307507042000261580

Chwe, M. S. (1999). Structure and strategy in collective action. *American Journal of Sociology, 105*(1), 128–156. doi: 10.1086/210269

Cialdini, R. B., Kallgren, C. A., & Reno, R. R. (1991). A focus theory of normative conduct: A theoretical refinement and reevaluation of the role of norms inhuman behavior. *Advances in Experimental Social Psychology, 24*, 201–234. doi: 10.1016/S0065-2601(08)60330-5

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20*(1), 37–46. doi: 10.1177/001316446002000104

Coleman, J. S. (1990). *Foundations of social theory*. Cambridge, MA: Harvard University Press.

Dawes, R. M. (1980). Social dilemmas. *Annual Review of Psychology, 31*, 169–193. doi: 10.1146/annurev.ps.31.020180.001125

de Jong, W. A., de Kleijn, R. A. M., Lockhorst, D., Brouwer, J., Noordegraaf, M., & van Tartwijk, J. W. F. (2023). Collaborative spirit: Understanding distributed leadership practices in and around teacher teams. *Teaching and Teacher Education*, *123*, 103977. doi: 10.1016/j.tate.2022.103977 de Matos Fernandes, Carlos A., Bakker, D. M., & Dijkstra, J. (2022a). Assessing the test-retest reliability of the social value orientation slider measure. *Judgment and Decision Making*, 17(1), 31–49. doi: 10.1017/S1930297500009013

de Matos Fernandes, C. A., Flache, A., Bakker, D. M., & Dijkstra, J. (2022b). A bad barrel spoils a good apple: How uncertainty and networks affect whether matching rules can foster cooperation. *Journal of Artificial Societies and Social Simulation, 25*(1), 6, doi: 10.18564/jasss.4754

de Matos Fernandes, C. A., & Keijzer, M. A. (2020) No one can predict the future: More than a semantic dispute. *Review of Artificial Societies and Social Simulation*.https://rofasss.org/2020/04/15/no-one-can-predict-the-future/

Deventer, J., Wagner, J., Lüdtke, O., & Trautwein, U. (2019). Are personality traits and relationship characteristics reciprocally related? Longitudinal analyses of codevelopment in the transition out of high school and beyond. *Journal of Personality and Social Psychology, 116*, 331–347. doi: 10.1037/pspp0000191

DeYoung, C. G. (2015). Cybernetic Big Five theory. *Journal of Research in Personality, 56*,33–58, doi: 10.1016/j.jrp.2014.07.004

Duca, S., & Nax, H. H. (2018). Groups and scores: The decline of cooperation. *Journal of the Royal Society Interface, 15*(144), 20180158. doi: 10.1098/rsif.2018.0158

Duca, S., Helbing, D., & Nax, H. H. (2018). Assortative matching with inequality involuntary contribution games. *Computational Economics, 52*(3), 1029–1043. doi: 10.1007/s10614-017-9774-5

Dunbar, R. I. M. (2018). The anatomy of friendship. *Trends in Cognitive Sciences, 22*(1),32–51. doi: 10.1016/j.tics.2017.10.004

Dunlop, P. D., Bharadwaj, A. A., & Parker, S. K. (2021). Two-year stability and change among the honesty-humility, agreeableness, and conscientiousness scales of theHEXACO100 in an Australian cohort, aged 24–29 years. *Personality and Individual Differences, 172*, 110601. doi: 10.1016/j.paid.2020.110601

Edelmann, A., Wolff, T., Montagne, D., & Bail, C. A. (2020). Computational social science and sociology. *Annual Review of Sociology, 46*(1), 61–81. doi: 10.1146/annurev-soc-121919-054621

Ehlert, A., Kindschi, M., Algesheimer, R., & Rauhut, H. (2020). Human social preferences cluster and spread in the field. *Proceedings of the National Academy of Sciences of the United States of America, 117*(37), 22787–22792. doi: 10.1073/pnas.2000824117

Eisinga, R., te Grotenhuis, M., & Pelzer, B. (2013). The reliability of a two-item scale: Pearson, Cronbach, or Spearman-Brown? doi: 10.1007/s00038-012-0416-3

Elster, J. (1986). Introduction. In J. Elster (Ed.), *Rational choice* (pp. 1–33). New York,NY: New York University Press.

Epstein, J. M. (1998). Zones of cooperation in demographic prisoner's dilemma. *Complexity, 4*(2), 36–48. doi: 10.1002/(SICI)1099-0526(199811/12)4:2(36::AID-CPLX9)3.0.CO;2-Z

Fang, R., Landis, B., Zhang, Z., Anderson, M. H., Shaw, J. D., & Kilduff, M. (2015).Integrating personality and social networks: A meta-analysis of personality,network position, and work outcomes in organizations. *Organization Science,26*(4), 1243–1260. doi: 10.1287/orsc.2015.0972

Fang, H., & Moro, A. (2011). Theories of statistical discrimination and affirmative action: A survey. In J. Benhabib, A. Bisin, & M. O. Jackson (Eds.), *Handbook of social economics*, (pp. 133–200). Amsterdam: North-Holland.

Fehl, K., van der Post, D. J., & Semmann, D. (2011). Co-evolution of behaviour and social network structure promotes human cooperation. *Ecology Letters, 14*(6),546–551. doi: 10.1111/j.1461-0248.2011.01615.x

Fehr, E., & Gintis, H. (2007). Human motivation and social cooperation: experimental and analytical foundations. *Annual Review of Sociology, 33*(1), 43–64. doi: 10.1146/annurev.soc.33.040406.131812

Fehr, E., & Gächter, S. (2002). Altruistic punishment in humans. *Nature, 415*, 137–140.doi: 10.1038/415137a

Fischbacher, U., Gächter, S., & Fehr, E. (2001). Are people conditionally cooperative? Evidence from a public goods experiment. *Economics Letters, 71*(3), 43–64. doi: 10.1016/S0165-1765(01)00394-9

Flache, A., Mäs, M., & Keijzer, M. A. (2022). Computational approaches in rigorous sociology: agent-based computational modeling and computational social science. In K. Gërxhani, N. de Graaf, & W. Raub (Eds.), *Handbook of sociological science: Contributions to rigorous sociology* (pp. 57–72). Cheltenham, UK: Edward Elgar Publishing. doi: 10.4337/9781789909432.00011

Flache, A., & de Matos Fernandes, C. A. (2021). Agent-based computational models. InG. Manzo (Ed.), *Research handbook on analytical sociology* (pp. 453–473).Cheltenham, UK: Edward Elgar Publishing. doi: 10.4337/9781789906851.00033

Flache, A., Bakker, D., Mäs, M., & Dijkstra, J. (2017a). The double edge of counter–sanctions: Is peer sanctioning robust to counter–punishment but vulnerable to counter–reward? In B. Jann, & W. Przepiorka (Eds.), *Social dilemmas, institutions,and the evolution of cooperation* (pp. 279-302). Oldenbourg: De Gruyter. doi: 10.1515/9783110472974-014

Flache, A., Mäs, M., Feliciani, T., Chattoe-Brown, E., Deffuant, G., Huet, S., & Lorenz,J. (2017b). Models of social influence: Towards the next frontiers. *Journal of Artificial Societies and Social Simulation, 20*(4), 2. doi: 10.18564/jasss.3521

Flache, A. (2002). The rational weakness of strong ties: Failure of group solidarity in a highly cohesive group of rational agents. *The Journal of Mathematical Sociology,26*(3), 189–216. doi: 10.1080/00222500212988

Flache, A., & Hegselmann, R. (1999). Rationality vs. learning in the evolution of solidarity networks: A theoretical comparison. *Computational and Mathematical Organization Theory, 5,* 97–127. doi: 10.1023/A:1009662602975

Flache, A., & Macy, M. W. (1996). The weakness of strong ties: Collective action failure in a highly cohesive group. *The Journal of Mathematical Sociology, 21*(1-2), 3–28.doi: 10.1080/0022250X.1996.9990172

Fowler, J. H., & Christakis, N. A. (2010). Cooperative behavior cascades in human social networks. *Proceedings of the National Academy of Sciences of the United States of America, 107*(12), 5334–5338. doi: 10.1073/pnas.0913149107

Fleeson, W. (2001). Toward a structure-and process-integrated view of personality: Traits as density distributions of states. *Journal of personality and social psychology,80*(6), 1011–1027. doi: 10.1037/0022-3514.80.6.1011

Fleiss, J. L. (1971). Measuring nominal scale agreement among many raters. *Psychological Bulletin, 76*(5), 378–382. doi: 10.1037/h0031619

Friedkin, N. E. (1998). *A structural theory of social influence.* Cambridge, England: Cambridge University Press.

Fu, F., Tarnita, C. E., Christakis, N. A., Wang, L., Rand, D. G., & Nowak, M. A. (2012). Evolution of in-group favoritism. *Scientific Reports, 2,* 460. doi: 10.1038/srep00460

Gambetta, D. (2009). Signalling. In P. Hedström & P. Bearman (Eds.), *The Oxford handbook of analytical sociology* (pp. 168–194). Oxford, UK: Oxford University Press. doi: 10.1093/oxfordhb/9780199215362.013.8

Giardini, F., & Wittek, R. P. M. (2019). Gossip, reputation, and sustainable cooperation: Sociological foundations. In F. Giardini & R. P. M. Wittek (Eds.), *The Oxford handbook of gossip and reputation* (pp. 23–46). Oxford, UK: Oxford University Press.

Girard, Y., Hett, F., & Schunk, D. (2015). How individual characteristics shape the structure of social networks. *Journal of Economic Behavior and Organization, 115*,197–216. doi: 10.1016/j.jebo.2014.12.005

Goedkoop, F., Sloot, D., Jans, L., Dijkstra, J., Flache, A., & Steg, L. (2022). The role of community in understanding involvement in community energy initiatives. *Frontiers in Psychology, 12*, 775752. doi: 10.3389/fpsyg.2021.775752

Gosling, S. D., Rentfrow, P. J., & Swann, W. B. (2003). A very brief measure of the Big-Five personality domains. *Journal of Research in Personality, 37*(6), 504–528. doi: 10.1016/S0092-6566(03)00046-1

Gould, R. V. (1993). Collective action and network structure. *American Sociological Review, 58*(2), 182–196. doi: 10.2307/2095965

Granovetter, M. (1985). Economic action and social structure: The problem of embeddedness. *American Journal of Sociology, 91*(3), 481–510. doi: 10.1086/228311

Granovetter, M. (1978). Threshold models of collective behavior. *American Journal of Sociology, 83*(6), 1420–1443. doi: 10.1086/226707

Greijdanus, H., de Matos Fernandes, C. A., Turner-Zwinkels, F., Honari, A., Roos, C.A., Rosenbusch, H., & Postmes, T. (2020). The psychology of online activism and social movements: relations between online and offline collective action. *Current Opinion in Psychology, 35*, 49–54. doi: 10.1016/j.copsyc.2020.03.003

Gross, J., & Dreu, C. K. W. D. (2019). The rise and fall of cooperation through reputation and group polarization. *Nature Communications, 10*(1), 1–10. doi: 10.1038/s41467-019-08727-8

Grow, A., Flache, A., & Wittek, R. P. M. (2017a). A model of global diversity and local consensus in status beliefs. *CoMSES Computational Model Library* (Version 1.2.0).Retrieved from: https://www.comses.net/codebases/5493/releases/1.2.0/

Grow, A., Flache, A., & Wittek, R. P. M. (2017b). Global diversity and local consensus in status beliefs: The role of network clustering and resistance to belief change. *Sociological Science, 4*(25), 611–640. doi: 10.15195/v4.a25

Guido, A., Robbett, A., & Romaniuc, R. (2019). Group formation and cooperation in social dilemmas: A survey and meta-analytic evidence. *Journal of Economic Behavior and Organization, 159,* 192–209. doi: 10.1016/j.jebo.2019.02.009

Gulati, R. (1995). Does familiarity breed trust? The implications of repeated ties for contractual choice in alliances. *Academy of Management Journal*, *38*(1), 85–112.doi: 10.5465/256729

Gunnthorsdottir, A., Vragov, R., Seifert, S., & McCabe, K. (2010). Near-efficiente quilibria in contribution-based competitive grouping. *Journal of PublicEconomics, 94*(11–12), 987–994. doi: 10.1016/j.jpubeco.2010.07.004

A

Gunnthorsdottir, A., Houser, D., & McCabe, K. (2007). Disposition, history, and contributions in public goods experiments. *Journal of Economic Behavior and Organization, 62*(2), 304–315. doi: 10.1016/j.jebo.2005.03.008

Hardin, G. (1968). The tragedy of the commons. *Science, 162*(3859), 1243–1248. doi: 10.1126/science.162.3859.1243

Harris, K., & Vazire, S. (2016). On friendship development and the Big Five personality traits. *Social and Personality Psychology Compass, 10*(11), 647–667. doi: 10.1111/spc3.12287

Heckathorn, D. D. (1996). Dynamics and dilemmas of collective action. *American Sociological Review, 61*(2), 250–277. doi: 10.2307/2096334

Hedström, P., & Ylikoski, P. (2010). Causal mechanisms in the social sciences. *Annual Review of Sociology, 36*, 49–67. doi: 10.1146/annurev.soc.012809.102632

Hedström, P. (2005). *Dissecting the social: On the principles of analytical sociology*. Cambridge, UK: Cambridge University Press.

Helbing, D., Yu, W. J., & Rauhut, H. (2011). Self-organization and emergence in social systems: Modeling the coevolution of social environments and cooperative behavior. The *Journal of Mathematical Sociology, 35*(1–3), 177–208. doi: 10.1080/0022250X.2010.532258

Helbing, D., & Yu, W. J. (2009). The outbreak of cooperation among success-driven individuals under noisy conditions. *Proceedings of the National Academy of Sciences of the United States of America, 106*(10), 3680–3685. doi: 10.1073/pnas.0811503106

Henry, S., & Mõttus, R. (2020). Traits and adaptations: A theoretical examination and new empirical evidence. *European Journal of Personality, 34*(3), 265–284, doi: 10.1002/per.2248

Hoffman, M., Block, P., & Snijders, T. A. B. (2023). Modeling partitions of individuals. *Sociological Methodology, 53*(1), 1–41. doi: 10.1177/00811750221145166

Hofman, J. M., Sharma, A., & Watts, D. J. (2017). *Prediction and explanation in social systems. Science, 355*(6324), 486–488. doi: 10.1126/science.aal3856

Höglinger, M., & Wehrli, S. (2017). Measuring social preferences on Amazon Mechanical Turk. In B. Jann & W. Przepriorka (Eds.), *Social dilemmas,institutions, and the evolution of cooperation* (pp. 527–546). Berlin, Boston: DeGruyter Oldenbourg.

Homans, G. C. (1974). *Social behavior. Its elementary forms*. New York, NY: Harcourt Brace Jovanovich.

Hooijsma, M., Huitsing, G., Kisfalusi, D., Dijkstra, J. K., Flache, A., & Veenstra, D. R.(2020). Multidimensional similarity in multiplex networks: Friendships between same- and cross-gender bullies and same- and cross-gender victims. *Network Science, 8*(1), 79–96. doi: 10.1017/nws.2020.1

Huckman, R. S., Staats, B. R., & Upton, D. M. (2009). Team familiarity, role experience, and performance: Evidence from Indian software services. *Management Science*, *55*(1), 85–100. doi: 10.1287/mnsc.1080.0921

Ilgen, D. R., Hollenbeck, J. R., Johnson, M., & Jundt, D. (2005). Teams in organizations: From input-process-output models to IMOI models. *Annual Review of Psychology*, *56*(1), 517–543. doi: 10.1146/annurev.psych.56.091103.070250

Indlekofer, N., & Brandes, U. (2013). Relative importance of effects in stochastic actor-oriented models. *Network Science, 1*(3), 278–304. doi: 10.1017/nws.2013.21

Iyengar, S., Lelkes, Y., Levendusky, M., Malhotra, N., & Westwood, S.J. (2019). The origins and consequences of affective polarization in the United States. *Annual Review of Political Science, 22*(1), 129–146. doi: 10.1146/annurev-polisci-051117-073034

Jackson, S. E., Joshi, A., & Erhardt, N. L. (2003). Recent research on team and organizational diversity: SWOT analysis and implications. *Journal of Management*, *29*(6), 801–830. doi: 10.1016/S0149-2063_03_00080-1

Jensen-Campbell, L. A., & Malcolm, K. T. (2007). The importance of conscientiousness in adolescent interpersonal relationships. *Personality and Social Psychology Bulletin, 33*(3), 368–383. doi: 10.1177/0146167206296104

Kadushin C. (2012). *Understanding social networks*. New York, NY: Oxford University Press.

Kaven, E., Kaven, I., Gómez-Zará, D., DeChurch, L., & Contractor, N. (2021). Assessing how team task influences team assembly through network analysis. In R. M. Benito, C. Cherifi, H. Cherifi, E. Moro, L. M. Rocha, & M. Sales-Pardo (Eds.),*Complex networks & their applications IX* (pp. 322–334). Springer International Publishing.

Keijzer, M. A., Mäs, M., & Flache, A. (2018). Communication in online social networks fosters cultural isolation. *Complexity*, 9502872. doi: 10.1155/2018/9502872

Kim, H., & Bearman, P. S. (1997). The structure and dynamics of movement participation. *American Sociological Review, 62*(1), 70–93. doi: 10.2307/2657453

King, E. B., Hebl, M. R., & Beal, D. J. (2009). Conflict and cooperation in diverse workgroups. *Journal of Social Issues*, *65*(2), 261–285. doi: 10.1111/j.1540-4560.2009.01600.x

Kivelä, M., Arenas, A., Barthelemy, M., Gleeson, J. P., Moreno, Y., & Porter, M. A.(2014). Multilayer networks. *Journal of Complex Networks, 2*(3), 203–271. doi: 10.1093/comnet/cnu016

Komorita, S. S., & Parks, C. D. (1996). *Social dilemmas*. Boulder, CO: Westview.

Korn, L., Böhm, R., Meier, N. W., & Betsch, C. (2020). Vaccination as a social contract. *Proceedings of the National Academy of Sciences of the United States of America,117*(26), 14890-14899. doi: 10.1073/pnas.1919666117

Kossinets, G., & Watts, D. J. (2009). Origins of homophily in an evolving social network. *American Journal of Sociology, 115*(2), 405–450. doi: 10.1086/599247

Krackhardt, D. (1988). Predicting with networks: Nonparametric multiple regressionanalysis of dyadic data. *Social Networks*, *10*(4), 359–381. doi: 10.1016/0378-8733(88)90004-4

Kraft-Todd, G., Yoeli, E., Bhanot, S., & Rand, D. G. (2015). Promoting cooperation in the field. *Current Opinion in Behavioral Sciences, 3,* 96–101. doi: 10.1016/j.cobeha.2015.02.006

Kreienkamp, J., Agostini, M., Kunz, M. C., Meyerhuber, M. I., & de Matos Fernandes, C. A. (2020). Normative influences in science and their impact on (objective)empirical research. In A. M. Bauer, & M. I. Meyerhuber (Eds.), *Empirical research and normative theory: Transdisciplinary perspectives on two methodical traditions between separation and interdependence* (pp. 75-104). Berlin, Boston: De Gruyter.

Kretschmer, D., & Leszczensky, L. (2022). In-group bias or out-group reluctance? The interplay of gender and religion in creating religious friendship segregation among Muslim youth. *Social Forces*. doi: 10.1093/sf/soab029

LaHuis, D. M., Hartman, M. J., Hakoyama, S., & Clark, P. C. (2014). Explained variance measures for multilevel models. *Organizational Research Methods, 17*(4), 433–451.doi: 10.1177/1094428114541701

Lazarsfeld, P. F., & Merton, R. K. (1954). Friendship as social process: Substantive and methodological analysis. In M. Berger, T. Abel, & C. H. Page (Eds.), *Freedom and control in modern society* (pp. 18–66). New York, NY: Van Nostrand.

Ledyard, J. O. (1995). Public goods: A survey of experimental research. In J. H. Kagel & A. E. Roth (Eds.), *Handbook of experimental economics* (pp. 111–194). Princeton,NJ: Princeton University Press.

Leggat, S. G. (2007). Effective healthcare teams require effective team members: defining teamwork competencies. *BMC Health Services Research*, *7*(1), 17. doi: 10.1186/1472-6963-7-17

Lin, N. (1999). Social networks and status attainment. *Annual Review of Sociology, 25*(1),467–487. doi: 10.1146/annurev.soc.25.1.467

Lindenberg, S. (1990). Homo Socio-oeconomicus: The emergence of a general model of man in the social sciences. *Journal of Institutional and Theoretical Economics,146*(4), 727–748.

Lomi, A., Snijders, T. A. B., Steglich, C. E. G., & Torló, V. J. (2011). Why are some more peer than others? Evidence from a longitudinal study of social networks and individual academic performance. *Social Science Research, 40*(6), 1506–1520. doi: 10.1016/j.ssresearch.2011.06.010

Lungeanu, A., Huang, Y., & Contractor, N. S. (2014). Understanding the assembly of interdisciplinary teams and its impact on performance. *Journal of Informetrics*,*8*(1), 59–70. doi: 10.1016/j.joi.2013.10.006

Lusher, D., Koskinen, J., & Robins, G. (2013). *Exponential random graph models for social networks: Theory, methods, and applications*. Cambridge, UK: Cambridge University Press.

Macy, M. W., & Evtushenko, A. (2020). Threshold models of collective behavior II: The predictability paradox and spontaneous instigation. *Sociological Science, 7*(26),628–648. doi: 10.15195/v7.a26

Macy, M. W., & Flache, A. (2002). Learning dynamics in social dilemmas. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(suppl 3), 7229–7236. doi: 10.1073/pnas.092080099

Macy, M. W., & Flache, A. (1995). Beyond rationality in models of choice. *Annual Review of Sociology, 21*, 73–91. doi: 10.1146/annurev.so.21.080195.000445

Macy, M. W. (1991a). Chains of cooperation: Threshold effects in collective action. *American Sociological Review, 56*(6), 730–747. doi: 10.2307/2096252

Macy, M. W. (1991b). Learning to cooperate: Stochastic and tacit collusion in social exchange. *American Journal of Sociology, 97*(3), 808–843. doi: 10.1086/229821

Manesi, Z., Van Lange, P. A. M., Van Doesum, N. J., & Pollet, T. V. (2019). What are the most powerful predictors of charitable giving to victims of typhoon Haiyan: Prosocial traits, socio-demographic variables, or eye cues? *Personality and Individual Differences, 146,* 217–225. doi: 10.1016/j.paid.2018.03.024

Manzo, G., Gabbriellini, S., Roux, V., & M'Mbogori, F. N. (2018). Complex contagions and the diffusion of innovations: Evidence from a small-N study. *Journal of Archaeological Method and Theory, 25*(4), 1109–1154. doi: 10.1007/s10816-018-9393-z

Marwell, G., & Ames, R. E. (1981). Economists free ride, does anyone else? Experiments on the provision of public goods, IV. *Journal of Public Economics, 15*(3), 295–310.doi: 10.1016/0047-2727(81)90013-x

A

Mäs, M. (2021). Analytical sociology and complexity research. In G. Manzo (Ed.),*Research handbook on analytical sociology* (pp. 100–118). Cheltenham, UK: Edward Elgar Publishing.
doi: 10.4337/9781789906851.00011

Mäs, M., & Nax, H. H. (2016). A behavioral study of "noise" in coordination games. *Journal of Economic Theory, 162*, 195–208. doi: 10.1016/j.jet.2015.12.010

Mäs, M., & Opp, K.-D. (2016). When is ignorance bliss? Disclosing true information and cascades of norm violation in networks. *Social Networks, 47*, 116–129. doi: 10.1016/j.socnet.2016.05.004

Mäs, M., Flache, A., Takács, K., & Jehn, K. A. (2013). In the short term we divide, in the long term we unite: Demographic crisscrossing and the effects of faultlines on subgroup polarization. *Organization Science, 24*(3), 716–736. doi: 10.1287/orsc.1120.0767

Mathieu, J. E., Maynard, M. T., Rapp, T., & Gilson, L. L. (2008). Team effectiveness1997-2007: A review of recent advancements and a glimpse into the future. *Journal of Management*, *34*(3), 410–476.
doi: 10.1177/0149206308316061

Maynard, M. T., Mathieu, J. E., Gilson, L. L., Sanchez, D. R., & Dean, M. D. (2019). Do I really know you and does it matter? Unpacking the relationship between familiarity and information elaboration in global virtual teams. *Group &Organization Management*, *44*(1), 3–37. doi: 10.1177/1059601118785842

McCrae, R. R., Kurtz, J. E., Yamagata, S., & Terracciano, A. (2011). Internal consistency,retest reliability, and their implications for personality scale validity. *Personality and Social Psychology Review, 15*(1), 28–50. doi: 10.1177/1088868310366253

McCrae, R. R., & John, O. P. (1992). An introduction to the Five-Factor Model and its applications. *Journal of Personality, 60*(2), 175–215.
doi: 10.1111/j.1467-6494.1992.tb00970.x

McPherson, J. M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a feather: Homophilyin social networks. *Annual Review of Sociology, 27*(1), 415–444.
doi: 10.1146/annurev.soc.27.1.415

McPherson, M., & Smith-Lovin, L. (1987). Homophily in voluntary organizations: Status distance and the composition of face-to-face groups. *American Sociological Review, 52*(3), 370–379. doi: 10.2307/2095356

Melamed, D., Munn, C. W., Simpson, B., Abernathy, J. Z., Harrell, A., & Sweitzer, M.(2020). Homophily and segregation in cooperative networks. *American Journal of Sociology, 125*(4), 1084–1127. doi: 10.1086/708142

Melamed, D., Harrell, A., & Simpson, B. (2018). Cooperation, clustering, and assortative mixing in dynamic networks. *Proceedings of the National Academy of Sciences of the United States of America, 115*(5), 951–956. doi: 10.1073/pnas.1715357115

Merton, R. K. (1968). *Social theory and social structure*. New York, NY: Free Press

Messick, D. M., & McClintock, C. G. (1968). Motivational bases of choice in experimental games. *Journal of Experimental Social Psychology, 4*(1), 1–25. doi: 10.1016/ 0022-1031(68)90046-2

Mijs, J. J. B. (2016). The unfulfillable promise of meritocracy: Three lessons and their implications for justice in education. *Social Justice Research, 29*(1), 14–34. doi: 10.1007/s11211-014-0228-0

Moody, J. (2001). Race, school integration, and friendship segregation in America. *American Journal of Sociology, 10*(3), 679–716. doi: 10.1086/338954

Murphy, R. O., & Ackermann, K. A. (2015). Social preferences, positive expectations,and trust based cooperation. *Journal of Mathematical Psychology, 67*, 45–50. doi: 10.1016/j.jmp.2015.06.001

Murphy, R. O., & Ackermann, K. A. (2014). Social value orientation: Theoretical and measurement issues in the study of social preferences. *Personality and Social Psychology Review, 18*(1), 13–41.
doi: 10.1177/1088868313501745

Murphy, R. O., Ackermann, K. A., & Handgraaf, M. J. J. (2011). Measuring social value orientation. *Judgment and Decision Making, 6*(8), 771–781.

Muskat, B., Anand, A., Contessotto, C., Tan, A. H. T., & Park, G. (2022). Team familiarity—Boon for routines, bane for innovation? A review and future research agenda. *Human Resource Management Review*, 100892. doi: 10.1016/j.hrmr.2021.100892

Nax, H. H., Murphy, R. O., Duca, S., & Helbing, D. (2017a). Contribution-based grouping under noise. *Games, 8*(4), 50. doi: 10.3390/g8040050

Nax, H. H., Murphy, R. O., & Helbing, D. (2017b). Nash dynamics, meritocratic matching, and cooperation. In B. Jann & W. Przepriorka (Eds.), *Social dilemmas,institutions, and the evolution of cooperation* (pp. 447–469). Berlin, Boston: DeGruyter Oldenbourg.

Nax, H. H., & Rigos, A. (2016). Assortativity evolving from social dilemmas. *Journal of Theoretical Biology, 395*, 194–203. doi: 10.1016/j.jtbi.2016.01.032

Nax, H. H., Perc, M., Szolnoki, A., & Helbing, D. (2015a). Stability of cooperation under image scoring in group interactions. *Scientific Reports, 5*(1), 12145. doi: 10.1038/srep12145

Nax, H. H., Balietti, S., Murphy, R. O., & Helbing, D. (2015b), Meritocratic matching can dissolve the efficiency-equality tradeoff: the case of voluntary contributions. *LSE Research Online Documents on Economics, London School of Economics and Political Science, LSE Library.*

Nikiforakis, N., Noussair, C. N., & Wilkening, T. (2012). Normative conflict and feuds: The limits of self-enforcement. *Journal of Public Economics, 96*(9–10), 797–807.doi: 10.1016/J.JPUBECO.2012.05.014

Nowak, M. A. (2006). Five rules for the evolution of cooperation. *Science, 314*(5805),1560–1563. doi: 10.1126/science.1133755

Nowak, M. A., & Sigmund, K. (1998). Evolution of indirect reciprocity by image scoring. *Nature, 393*(6685), 573–577. doi: 10.1038/31225

Nowak, M. A., & May, R. M. (1992). Evolutionary games and spatial chaos. *Nature, 359*,826–829. doi: 10.1038/359826a0

Olson, M. (1965). *The logic of collective action: Public goods and the theory of groups.* Cambridge, MA: Harvard University Press.

Opp, K.-D., & Gern, C. (1993). Dissident groups, personal networks, and spontaneous cooperation: The East German revolution of 1989. *American Sociological Review,58*(5), 659–680. doi: 10.2307/2096280

Orbell, J. M., & Dawes, R. M. (1993). Social welfare, cooperators' advantage, and the option of not playing the game. *American Sociological Review, 58*(6), 787–800. doi: 10.2307/2095951

Orbell, J., Schwartz-Shea, P., & Simmons, R. (1984). Do cooperators exit more readily than defectors? *American Political Science Review, 78*(1), 147–162. doi: 10.2307/1961254

Ostrom, E. (2010). Polycentric systems for coping with collective action and global environmental change. *Global Environmental Change, 20*, 550–557. doi: 10.1016/j.gloenvcha.2010.07.004

Ostrom, E. (2009). Building trust to solve commons dilemmas: Taking small steps to test an evolving theory of collective action. In S. A. Levin (Ed.), *Games, groups, and the global good* (pp 207–228). Berlin, Heidelberg: Springer.

Parker, P. D., Lüdtke, O., Trautwein, U., & Roberts, B. W. (2012). Personality and relationship quality during the transition from high school to early adulthood. *Journal of Personality, 80*(4), 1061–1089. doi: 10.1111/j.1467-6494.2012.00766.x

Perc, M., Jordan, J. J., Rand, D. G., Wang, Z., Boccaletti, S., & Szolnoki, A. (2017).Statistical physics of human cooperation. *Physics Reports, 687*, 1–51. doi: 10.1016/j.physrep.2017.05.004

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

R Core Team. (2021). *R: A language and environment for statistical computing* [Computer software manual]. Vienna, Austria.

Rand, D. G., Nowak, M. A., Fowler, J. H., & Christakis, N. A. (2014). Static network structure can stabilize human cooperation. *Proceedings of the National Academy of Sciences of the United States of America, 111*(48), 17093–17098. doi: 10.1073/pnas.1400406111

Rand, D. G., & Nowak, M. A. (2013). Human cooperation. *Trends in Cognitive Sciences,17*(8), 413–425. doi: 10.1016/J.TICS.2013.06.003

Raub, W., & Weesie, J. (1990). Reputation and efficiency in social interactions: An example of network effects. *American Journal of Sociology, 96*(3), 626–654. doi: 10.1086/229574

Reyna, C., Belaus, A., Mola, D., Ortiz, M. V., & Acosta, C. (2018). Social values orientation slider measure: Evidences of validity and reliability among Argentine undergraduate students. *Testing, Psychometrics, Methodology in Applied Psychology, 25*(3), 395–408. doi: 10.4473/TPM25.3.5

Ripley, R. M., Snijders, T. A. B., Boda, Z., Vörös, A., & Preciado, P. (2021). *Manual for siena version 4.0* [Computer software manual]. Oxford, UK: University of Oxford, Department of Statistics, Nuffield College.

Robbins, H., & Monro, S. (1951). A stochastic approximation method. *The Annals of Mathematical Statistics*, *22*(3), 400–407. http://www.jstor.org/stable/2236626

Robins, R. W., Caspi, A., & Moffitt, T. E. (2002). It's not just who you're with, it's who you are: Personality and relationship experiences across multiple relationships. *Journal of Personality, 70*(6), 925–964. doi: 10.1111/1467-6494.05028

Ross, L. (1977). The intuitive psychologist and his shortcomings: Distortions in the attribution process. *Advances in Experimental Social Psychology, 10*, 173–220. doi: 10.1016/s0065-2601(08)60357-3

Rossetti, C. S. L., Hilbe, C., & Hauser, O. P. (2022). (Mis)perceiving cooperativeness. *Current Opinion in Psychology, 43,* 151–155. doi: 10.1016/j.copsyc.2021.06.020

Ruggeri, K., Stock, F., Haslam, S. A., Capraro, V., Boggio, P., Ellemers, N., … Willer, R.(2022). Evaluating expectations from social and behavioral science aboutCOVID-19 and lessons for the next pandemic. *PsyArXiv*. doi: 10.31234/osf.io/58udn

Sandel, M. (2020). *The tyranny of merit: What's become of the common good?* New York,NY: Farrar, Strauss, and Giroux.

Schaefer, D. R., adams, j., & Haas, S. A. (2013). Social networks and smoking: Exploring the effects of peer influence and smoker popularity through simulations. *Health Education & Behavior, 40*(1 suppl), 24S-32S. doi: 10.1177/1090198113493091

Scheffer, M., Van Bavel, B., Van De Leemput, I. A., & Van Nes, E. H. (2017). Inequality in nature and society. *Proceedings of the National Academy of Sciences of the United States of America, 114*(50), 13154–13157. doi: 10.1073/pnas.1706412114

Schelling, T. C. (1978). *Micromotives and macrobehavior.* New York, NY: W.W. Norton &Company.

Schelling, T. C. (1971). Dynamic models of segregation. *The Journal of Mathematical Sociology, 1*(2), 143–186. doi: 10.1080/0022250X.1971.9989794

Schuessler, R. (1989). Exit threats and cooperation under anonymity. *Journal of Conflict Resolution, 33*(4), 728–749. doi: 10.1177/0022002789033004007

Selfhout, M., Burk, W., Branje, S., Denissen, J., Van Aken, M., & Meeus, W. (2010).Emerging late adolescent friendship networks and Big Five personality traits: A social network approach. *Journal of Personality, 78*(2), 509–538. doi: 10.1111/j.1467-6494.2010.00625.x

Shchebetenko, S. (2019). Do personality characteristics explain the associations between self-esteem and online social networking behaviour? *Computers in Human Behavior, 91*, 17–23. doi: 10.1016/j.chb.2018.09.017

Simpson, B., & Willer, R. (2015). Beyond altruism: Sociological foundations of cooperation and prosocial behavior. *Annual Review of Sociology, 41*, 43–63. doi: 10.1146/annurev-soc-073014-112242

Simpson, B., Brashears, M., Gladstone, E., & Harrell, A. (2014). Birds of different feathers cooperate together: No evidence for altruism homophily in networks. *Sociological Science, 1,* 542–564. doi: 10.15195/v1.a30

Smith, B. L., MacGregor, J., Matthews, R. S., & Gabelnick, F. (2004). *Learning communities: Reforming undergraduate education*. San Francisco, CA: Jossey-Bass Publishers.

Snijders, T. A. B., & Lomi, A. (2019). Beyond homophily: Incorporating actor variables in statistical network models. *Network Science, 7*(1), 1–19. doi: 10.1017/nws.2018.30

Snijders, T. A. B. (2017). Stochastic actor-oriented models for network dynamics. *Annual Review of Statistics and Its Application*, *4*(1), 343–363. doi: 10.1146/annurev-statistics-060116-054035

Snijders, T. A. B., & Steglich, C. E. G. (2015). Representing micro–macro linkages byactor-based dynamic network models. *Sociological Methods and Research, 44*(2),222–271. doi: 10.1177/0049124113494573

Snijders, T. A. B., Lomi, A., & Torló, V. J. (2013). A model for the multiplex dynamics of two-mode and one-mode networks, with an application to employment preference, friendship, and advice. *Social Networks, 35*(2), 265–276. doi: 10.1016/j.socnet.2012.05.005

Snijders, T. A. B., van de Bunt, G. G., & Steglich, C. E. G. (2010). Introduction to stochastic actor-based models for network dynamics. *Social Networks, 32*(1), 44–60. doi: 10.1016/j.socnet.2009.02.004.

Snijders, T. A. B. (2002). Markov chain Monte Carlo estimation of exponential random graph models. *Journal of Social Structure*, *3*(2), 1–40.

Snijders, T. A. B. & Bosker, R. J. (1994). Modeled variance in two-level models. *Sociological Methods & Research, 22*(3), 342–363, doi: 10.1177/0049124194022003004

Spence, M. (1973). Job market signaling. *The Quarterly Journal of Economics, 87*(3), 355–374. doi: 10.2307/1882010

Stadtfeld, C., Takács, K., & Vörös, A. (2020). The emergence and stability of groups in social networks. *Social Networks, 60*, 129–145. doi: 10.1016/j.socnet.2019.10.008

Stadtfeld, C., Vörös, A., Elmer, T., Boda, Z., & Raabe, I. J. (2019). Integration in emerging social networks explains academic failure and success. *Proceedings of the National Academy of Sciences of the United States of America, 116*(3), 792–797. doi: 10.1073/PNAS.1811388115

Stadtfeld, C. (2018). The micro-macro link in social networks. *Emerging Trends in the Social and Behavioral Sciences*, 1–15. doi: 10.1002/ 9781118900772.etrds0463

Stanley, W., & Faust, K. (1994). *Social network analysis: Methods and applications*. New York, NY: Cambridge University Press.

Steglich, C. E. G., & Snijders, T. A. B. (2022). Stochastic network modelling as generative social science. In K. Gërxhani, N. D. de Graaf, & W. Raub (Eds.),*Handbook of sociological science: Contributions to rigorous sociology* (pp. 73–99).Cheltenham, UK: Edward Elgar Publishing.

Steglich, C. E. G. (2018). Why echo chambers form and network interventions fail: Selection outpaces influence in dynamic networks. *arXiv*. doi: 10.48550/ARXIV.1810.00211

Steglich, C. E. G., Snijders, T. A. B., & Pearson, M. (2010). Dynamic networks and behavior: Separating selection from influence. *Sociological Methodology, 40*(1),329–393. doi: 10.1111/j.1467-9531.2010.01225.x

Stiglitz, J. E. (2012). *The price of inequality: How today's divided society endangers our future*. New York, NY: W.W. Norton & Company.

Tackman, A. M., Srivastava, S., Pfeifer, J. H., & Dapretto, M. (2017). Development of conscientiousness in childhood and adolescence: Typical trajectories and associations with academic, health, and relationship changes. *Journal of Research in Personality, 67*, 85–96. doi: 10.1016/j.jrp.2016.05.002

Takács K., Gross, J., Testori, M., Letina, S., Kenny, A. R., Power, E. A. & Wittek, R. P.M. (2021). Networks of reliable reputations and cooperation: A review. *Philosophical Transactions of the Royal Society B: Biological Sciences, 376*, 20200297. doi: 10.1098/rstb.2020.0297

Takács, K., Janky, B., & Flache, A. (2008). Collective action and network change. *Social Networks, 30*(3), 177–189. doi: 10.1016/j.socnet.2008.02.003

Tannenbaum, S. I., Traylor, A. M., Thomas, E. J., & Salas, E. (2021). Managing teamwork in the face of pandemic: evidence-based tips. *BMJ Quality & Safety*,*30*(1), 59–63. doi: 10.1136/bmjqs-2020-011447

Thaler, R. H, & Sunstein, C. R. (2008). *Nudge: Improving decisions about health, wealth,and happiness*. New Haven, CT: Yale University Press.

Thielmann, I., Spadaro, G., & Balliet, D. (2020). Personality and prosocial behavior: A theoretical framework and meta-analysis. *Psychological Bulletin, 146*(1), 30–90. doi: 10.1037/bul0000217

Trautwein, C., & Bosse, E. (2017). The first year in higher education—critical requirements from the student perspective. *Higher Education, 73*(3), 371–387.doi: 10.1007/s10734-016-0098-5

Truijens, D. (2022). Coherence between theory and policy in Nudge and Boost: Is it relevant for evidence-based policy-making? *Rationality and Society, 34*(3), 368–394. doi: 10.1177/10434631221093741

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185*(4157), 1124–1131. doi: 10.1126/science.185.4157.1124

Udehn, L. (2001). *Methodological individualism: Background, history and meaning*. London: Taylor & Francis.

Uzzi, B. (1997). Social Structure and competition in interfirm networks: The paradox of embeddedness. *Administrative Science Quarterly*, *42*(1), 35–67. doi: 10.2307/2393808

Van Bavel, J. J., Baicker, K., Boggio, P. S., Capraro, V., Cichocka, A., Cikara, M.,Crockett, M. J., Crum, A. J., Douglas, K. M., Druckman, J. N., Drury, J., Dube, O., Ellemers, N., Finkel, E. J., Fowler, J. H., Gelfand, M., Han, S., Haslam, S. A.,Jetten, J., … Willer, R. (2020). Using social and behavioural science to support-COVID-19 pandemic response. *Nature Human Behaviour, 4*(5), 460–471. doi: 10.1038/s41562-020-0884-z

van de Bunt, G. G., van Duijn, M. A. J., & Snijders, T. A. B. (1999). Friendship networks through time: An actor-oriented dynamic statistical network model. *Computational & Mathematical Organization Theory*, *5*(2), 167–192. doi: 10.1023/A:1009683123448

van der Zanden, P. J. A. C., Denessen, E., Cillessen, A. H. N., & Meijer, P. C. (2018).Domains and predictors of first-year student success: A systematic review. *Educational Research Review, 23*, 57–77. doi: 10.1016/j.edurev.2018.01.001

van Duijn, M. A. J., Zeggelink, E. P. H., Huisman, M., Stokman, F. N., & Wasseur, F.W. (2003). Evolution of sociology freshmen into a friendship network. *Journal of Mathematical Sociology*, *27*, 153–191. doi: 10.1080/00222500390213137

van Knippenberg, D., & Schippers, M. C. (2007). Work group diversity. *Annual Review of Psychology*, *58*(1), 515–541. doi: 10.1146/annurev.psych.58.110405.085546

Van Lange, P. A. M., & Rand, D. G. (2022). Human cooperation and the crises of climate change, COVID-19, and misinformation. *Annual Review of Psychology, 73*(1),379–402. doi: 10.1146/annurev-psych-020821-110044

Van Lange, P. A. M., Balliet, D. P., Parks, C. D., & Van Vugt, M. (2014). *Social dilemmas: Understanding human cooperation.* Oxford, UK: Oxford University Press.

Van Lange, P. A. M., Joireman, J., Parks, C. D., & van Dijk, E. (2013). The psychology of social dilemmas: A review. *Organizational Behavior and Human Decision Processes, 120*(2), 125–141. doi: 10.1016/j.obhdp.2012.11.003

Van Lange, P. A. M. (1999). The pursuit of joint outcomes and equality in outcomes: An integrative model of social value orientation. *Journal of Personality and Social Psychology, 77*(2), 337–349. doi: 10.1037/0022-3514.77.2.337

van Zalk, M. H. W., Nestler, S., Geukes, K., Hutteman, R., & Back, M. D. (2020). The codevelopment of extraversion and friendships: Bonding and behavioral interaction mechanisms in friendship networks. *Journal of Personality and Social Psychology, 118*, 1269–1290. doi: 10.1037/pspp0000253

Verbrugge, L. M. (1977). The structure of adult friendship choices. *Social Forces*, *56*(2),576–597. doi: 10.1093/sf/56.2.576

von Neumann, J., & Morgenstern, O. (1944) *Theory of games and economic behavior.* Princeton, NJ: Princeton University Press.

Waldeck, R. (2013). Segregated cooperation. *Journal of Artificial Societies and Social Simulation, 16*(4), 14. doi: 10.18564/jasss.2363

Wang, P., Sharpe, K., Robins, G. L., & Pattison, P. E. (2009). Exponential random graph(p*) models for affiliation networks. *Social Networks*, *31*(1), 12–25. doi: 10.1016/j.socnet.2008.08.002

Watts, D. J. (2017). Should social science be more solution-oriented? *Nature Human Behaviour, 1*(1), 15. doi: 10.1038/s41562-016-0015

Weber, H., Schwenzer, M., & Hillmert, S. (2020). Homophily in the formation and development of learning networks among university students. *Network Science,8*(4), 469–491. doi: 10.1017/nws.2020.10

Wilcox, P., Winn, S., & Fyvie-Gauld, M. (2005). 'It was nothing to do with the university, it was just the people': the role of social support in the first-year experience of higher education. *Studies in Higher Education, 30*(6), 707–722. doi: 10.1080/03075070500340036

Wilensky, U. (1999). NetLogo. *Center for connected learning and computer-based modeling,Northwestern University*.

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

Wilmot, M. P., & Ones, D. S. (2019). A century of research on conscientiousness at work. *Proceedings of the National Academy of Sciences of the United States of America, 116*(46), 23004–23010. doi: 10.1073/pnas.1908430116

Winter, F., & Kataria, M. (2020). You are who your friends are? An experiment onhomophily in trustworthiness among friends. *Rationality and Society, 32*(2), 223–251. doi: 10.1177/1043463120919380

Wittek, R. P. M., & Bekkers, R. (2015). The sociology of altruism and prosocial behavior. In J. D. Wright (Ed.), *International encyclopedia of the social and behavioral sciences*(pp. 579–583). Elsevier.

Wittek, R. P. M., Snijders, T. A. B., & Nee, V. (2013). Introduction: Rational choice social research. In R. P. M. Wittek, T. A. B. Snijders, & V. Nee (Eds.), *The handbook of rational choice social research* (pp. 1–30). Stanford, CA: Stanford University Press.

Wong, L. H., Pattison, P., & Robins, G. (2006). A spatial model for social networks. *Physica A: Statistical Mechanics and Its Applications, 360*(1), 99–120. doi: 10.1016/j.physa.2005.04.029

Wu J., Balliet, D., & Van Lange, P. A. M. (2016). Reputation, gossip, and human cooperation. *Social and Personality Psychology Compass, 10*, 350–364. doi: 10.1111/spc3.12255

Yang, Y., Tian, T. Y., Woodruff, T. K., Jones, B. F., & Uzzi, B. (2022). Gender-diverse teams produce more novel and higher-impact scientific ideas. *Proceedings of the National Academy of Sciences of the United States of America*, *119*(36), e2200841119. doi: 10.1073/pnas.2200841119

Yinger, J. M. (1982). *Countercultures: The promise and peril of a world turned upside down.* New York, NY: Free Press.

Young, M. (1958). *The rise of the meritocracy, 1870–2033: An essay on education and society.* London, UK: Thames and Hudson.

Zhao, C.-M., & Kuh, G. D. (2004). Adding value: Learning communities and student engagement. *Research in Higher Education*, *45*(2), 115–138. doi: 10.1023/B:RIHE.0000015692.88534.de

Zhu, X., Woo, S. E., Porter, C., & Brzezinski, M. (2013). Pathways to happiness: From personality to social networks and perceived support. *Social Networks, 35*(3),382–393. doi: 10.1016/j.socnet.2013.04.005

A

ACKNOWLEDGMENTS (DANKWOORD)

Mijn dank gaat uit naar mijn promotoren Andreas, Jacob en Dieko. Zij gaven mij de ruimte om te groeien in mijn promotietraject. Voor mijn gevoel zat ik 5 jaar geleden eerst op de achterbank wat meer de kat uit de boom aan het kijken (en te twijfelen of ik wel op de goede route zat), terwijl ik gaandeweg steeds meer opschoof naar de chauffeursstoel en het heft in handen nam. En die ruimte kreeg ik ook van jullie.

Allereerst wil ik graag Andreas hartelijk bedanken. Tien jaar geleden begon ik bij Sociologie met een college van jou bij het vak *Inleiding in de Sociologie* en toen – nu nog steeds trouwens – vond ik de Sociologie al bijster interessant. De passie die jij voor het vak hebt, is aanstekelijk en inspirerend. Je hebt mijn manier van analyseren en denken over sociale mechanismen op een positieve manier vormgegeven. Ik vond het daarnaast fantastisch om betrokken te raken bij het vak *Inleiding in de Sociologie*. Het feit dat je mij de ruimte en het vertrouwen gaf om dat ook daadwerkelijk te doen, zegt veel over jou als begeleider en als mens. Ik kan uren doorgaan over hoeveel ik van je heb geleerd, hoe scherp je mijn stukken las, en hoe goed en gedetailleerd je feedback altijd was (en hoe verrast ik altijd was dat ik daar zelf niet aan had gedacht). Maar ik ben vooral dankbaar dat je mijn mentor wilde zijn tijdens de Research Master en mijn PhD-project en mij de kans gaf om van je te leren over van alles en nog wat binnen maar ook buiten de wetenschap. Ik wil nog niet zo ver gaan dat ik mij ga abonneren op *The Economist*, maar ik denk er wel over na.

Terwijl Andreas onderwijs gaf in semester 1A in mijn eerste jaar in de Bachelor, gaf Jacob het eerste vak dat ik volgde in semester 1B (Markt). Waar ik destijds, maar vooral ook in mijn PhD-project veel van heb geleerd, is hoe je complexe vraagstukken terug kan brengen naar begrijpelijke voorbeelden. Ik moet toegeven: Mijn voorbeelden zijn niet zo kleurrijk als die van jou. Het is een kracht om humor te koppelen aan inhoudelijk sterke feedback, en ik durf wel te stellen dat jij dat kan als de beste. De gesprekken die we hadden als team verliepen niet altijd even efficiënt, maar waren wel enorm leerzaam met bij tijd en wijlen wat zijpaadjes. Waar ik veel aan heb gehad, speelde zich ook af in de beginfase van mijn PhD-project. Ik twijfelde veel. Jij (en Andreas zeer zeker ook) gaven mij alle ruimte om tot rust te komen. Dat tekent ook jullie vriendelijkheid en compassie. In een mail schreef je het volgende: *Be true to yourself, and you can't go wrong*. Dat heb ik ter harte genomen (en doe ik nog steeds).

De komst van Dieko als begeleider fungeerde in mijn PhD-project als een vliegwieleffect. Jouw kennis over samenwerkingsproblemen paste perfect met waar ik al een tijdje mee bezig was, maar gaf het essentiële duwtje in de rug. Daarnaast gaf je mij de ruimte om tijdens de bilaterale gesprekken het eens over wat anders te hebben. Veel dank dat jij je tijd en energie – terwijl je al elders werkzaam was – ook wilde stoppen in mijn project. En het is mooi dat we elkaar weer tegenkomen in onze andere werkzaamheden.

As I embarked on this journey towards becoming "sufficiently able to conduct independent research," I was happy to be part of room B.0111. Our Ph.D. projects were diverse, mirroring the different research lines we have in our department. If you put all the Department of Sociology research lines in a meat grinder, office B.0111 would have been the results. We agreed on many things, such as the importance of the Coleman boat, that it was a shame that we didn't have a workplace watching over the pond, and that it was too hot in the summer and too cold in the winter in our office. We founded the "Journal of Institutional Heating Policies" to address heating issues. For example, in one paper, it was noted that "*to better understand the dynamics of heater views, future research should adopt a longitudinal approach to study the effect of monthly variations in the average temperature on heating views. This research, however, provides a strong base for future research and heating policy in office B.0111*." On a serious note, I was fascinated by the topics my officemates studied (and still study) and learned a lot. Thank you, Xingna, Eleonora, Wouter, Simon, Rowan, Louisa, and Aliona.

Additionally, I want to acknowledge Jasperina and Marion, whose intellectual stimulation and support have contributed to my academic growth. The fruitful discussions, collaborative projects, and shared experiences have broadened my perspective and encouraged my research. Nog een zin toevoegen: We cooperated well with each other (to put it in terms closely related to this dissertation).

I want to express my sincere gratitude to the Norms and Networks Cluster members for their invaluable contributions and the enriching experience I had as the secretary of the group. The discussions within the cluster were fruitful, sparking intellectual curiosity, and fostering deep insights into the dynamics of norms and networks. Furthermore, I want to emphasize the low-key atmosphere that allowed for open dialogue and the free exchange of ideas. Thank you to all current and prior NNC members – Michael, Dieko, Verena, Coby, Loes, Shuo, Melis, Lars, Fleur, Echo, Dennis, Daniel, Alla, Jonas, Hendrik, Thomas, Tibor, Tanzhe, Vincenz, Jacob, Jochem, and Andreas – for making the cluster an inspiring and enriching community.

En ja, ook ik heb een leven buiten de wetenschap. Ik wil graag mijn dankbaarheid uiten aan mijn vrienden en mede-TKB'ers, die gedurende mijn hele promotietraject – en daarvoor – een onmisbare bron van steun, motivatie en gezelligheid zijn geweest. We zijn verenigd geweest in diverse collectieven: Heren 18, maten van het judo, Heren 17, Heren 6, Velo 4, KB, IT en Rosé. De gemene deler is sport. Ik ga niet iedereen bij langs, maar de mannen in de bovenstaande gremia hebben in alles een gemeenschapsgevoel dat ik als zeer fijn heb ervaren en nog steeds ervaar. Of dat ligt aan het bier weet ik niet zeker, maar het was (en is) altijd gezellig en feestelijk (met altijd ook ruimte voor een serieus gesprek over het leven). Ik wil toch stilstaan bij één clubje, de *Fantastic Four*, bestaande uit Daniël, Jurgen, Eyal en ik. Op een dag hadden we bedacht dat we maar eens lid moesten worden van een voetbalclub. En zo geschiedde. Het is een unaniem belachelijk groot succes op en naast het veld geworden: Pullen, Willy Bomba, Hodor, pasta pesto, Huize Zwaan, niet een bal afgeven als je de kans hebt om tijdens je debuut te scoren, rugbyen in plaats van voetballen, Subway, Haka's; ik kijk altijd terug met een glimlach op mijn gezicht. Jullie hebben mijn promotietraject verrijkt. Het is alleen maar passend dat mijn paranimfen twee heren zijn geworden – Daniël end Jurgen (ook bekend als Hans) – die ik al sinds de middelbare school ken, die ook de faculteit (faccu voor intimi) kennen en er altijd zijn voor eenieder die hen nodig hebben (zo ook voor mij).

Ik ben ook bevoorrecht om schoonouders te hebben in Abe en Marleen die altijd voor mij klaar stonden (en nog steeds staan). Wijze raad, hulp in alle vormen, een klap op mijn schouders, een goede of slechte grap, schop onder mijn kont, heerlijk eten, een lekker biertje of jullie tijd; jullie gaven het altijd onvoorwaardelijk waar nodig. Vanaf het eerste moment dat ik binnen kwam, voelde ik mij thuis. En dat komt door jullie.

Mijn diepste waardering wil ik uitspreken richting mijn ouders voor hun steun en opoffering gedurende mijn hele sportieve en academische reis. Ik ben jullie enorm dankbaar voor de offers die jullie hebben gebracht om Fabio en mij van een kleine zolderkamer naar waar we nu staan als gezin te brengen. Ons succes is jullie succes. Het is ook heel speciaal dat Fabio en ik als broers binnen 5 maanden van elkaar vader zijn geworden. Het is mooi om zo'n lieve familie te hebben. Bedankt papa, mama, Fabio, Nurlaila en Matteo.

Aan het begin van dit proefschrift staat "voor mijn gezin" en daar wil ik ook graag mee afsluiten. Marije, je bent al 12 jaar mijn steun en toeverlaat in alles en jij snapt mij. Samen met onze Seppe en ons hummeltje dat onderweg is, begint onze reis maar net. Dit boek is voor jullie.

ICS DISSERTATION SERIES

The ICS series presents dissertations of the Interuniversity Center for Social Science Theory and Methodology. Each of these studies aims at integrating explicit theory formation with state-of-the-art empirical research or at the development of advanced methods for empirical research. The ICS was founded in 1986 as a cooperative effort of the universities of Groningen and Utrecht. Since 1992, the ICS has expanded to the University of Nijmegen and, since 2017, to the University of Amsterdam. Most of the projects are financed by the participating universities or by the Dutch Research Council. The international composition of the ICS graduate students is mirrored in the increasing international orientation of the projects and thus of the ICS series itself.

1. Kees van Liere (1990), *"Lastige leerlingen: Een empirisch onderzoek naar sociale oorzaken van probleemgedrag op basisscholen."* Amsterdam: Thesis Publishers
2. Marco van Leeuwen (1990), *"Bijstand in Amsterdam, ca. 1800-1850: Armenzorg als beheersings en overlevingsstrategie."* ICS dissertation, Utrecht
3. Ineke Maas (1990), *"Deelname aan podiumkunsten via de podia, de media en actieve beoefening: Substitutie of leereffecten?"* Amsterdam: Thesis Publishers
4. Marjolein Broese van Groenou (1991), *"Gescheiden netwerken: De relaties met vrienden en verwanten na echtscheiding."* Amsterdam: Thesis Publishers
5. Jan van den Bos (1991), *"Dutch EC policy making: A model guided approach to coordination and negotiation."* Amsterdam: Thesis Publishers
6. Karin Sanders (1991), *"Vrouwelijke pioniers: Vrouwen en mannen met een 'mannelijke' hogere beroepsopleiding aan het begin van hun loopbaan."* Amsterdam: Thesis Publishers
7. Sjerp de Vries (1991), *"Egoism, altruism, and social justice: Theory and experiments on cooperation in social dilemmas."* Amsterdam: Thesis Publishers
8. Ronald Batenburg (1991), *"Automatisering in bedrijf."* Amsterdam: Thesis Publishers
9. Rudi Wielers (1991), *"Selectie en allocatie op de arbeidsmarkt. Een uitwerking voor de informele en geïnstitutionaliseerde kinderopvang."* Amsterdam: Thesis Publishers
10. Gert Westert (1991), *"Verschillen in ziekenhuisgebruik."* ICS dissertation, Groningen
11. Hanneke Hermsen (1992), *"Votes and policy preferences: Equilibria in party systems."* Amsterdam: Thesis Publishers
12. Cora Maas (1992), *"Probleemleerlingen in het basisonderwijs."* Amsterdam: Thesis Publishers
13. Ed Boxman (1992), *"Contacten en carrière : Een empirisch theoretisch onderzoek naar de relatie tussen sociale netwerken en arbeidsmarktposities"* Amsterdam: Thesis Publishers
14. Conny Taes (1992), *"Kijken naar banen: Een onderzoek naar de inschatting van arbeidsmarktkansen bij schoolverlaters uit het middelbaar beroepsonderwijs."* Amsterdam: Thesis Publishers
15. Peter van Roozendaal (1992), *"Cabinets in multi party democracies: The effect of dominant and central parties on cabinet composition and durability."* Amsterdam: Thesis Publishers
16. Marcel van Dam (1992), *"Regio zonder regie: Verschillen in en effectiviteit van gemeentelijk arbeidsmarktbeleid."* Amsterdam: Thesis Publishers

17. Tanja van der Lippe (1993), *"Arbeidsverdeling tussen mannen en vrouwen."* Amsterdam: Thesis Publishers
18. Marc Jacobs (1993), *"Software: Kopen of kopiëren? Een sociaal wetenschappelijk onderzoek onder PC gebruikers."* Amsterdam: Thesis Publishers
19. Peter van der Meer (1993), *"Verdringing op de Nederlandse arbeidsmarkt: Sector- en seksev-erschillen."* Amsterdam: Thesis Publishers
20. Gerbert Kraaykamp (1993), *"Over lezen gesproken: Een studie naar sociale differentiatie in leesgedrag."* Amsterdam: Thesis Publishers
21. Evelien Zeggelink (1993), *"Strangers into friends: The evolution of friendship networks using an individual oriented modeling approach."* Amsterdam: Thesis Publishers
22. Jaco Berveling (1994), *"Het stempel op de besluitvorming: Macht, invloed en besluitvorming op twee Amsterdamse beleidsterreinen."* Amsterdam: Thesis Publishers
23. Wim Bernasco (1994), *"Coupled careers: The effects of spouse's resources on success at work."* Amsterdam: Thesis Publishers
24. Liset van Dijk (1994), *"Choices in child care: The distribution of child care among mothers, fathers and non parental care providers."* Amsterdam: Thesis Publishers
25. Jos de Haan (1994), *"Research groups in Dutch sociology."* Amsterdam: Thesis Publishers
26. Kwasi Boahene (1995), *"Innovation adoption as a socioeconomic process: The case of the Ghanaian cocoa industry."* Amsterdam: Thesis Publishers
27. Paul Ligthart (1995), *"solidarity in economic transactions: An experimental study of framing effects in bargaining and contracting."* Amsterdam: Thesis Publishers
28. Roger Leenders (1995), *"Structure and influence: Statistical models for the dynamics of actor attributes, network structure, and their interdependence."* Amsterdam: Thesis Publishers
29. Beate Volker (1995), *"Should auld acquaintance be forgot...? Institutions of communism, the transition to capitalism and personal networks: The case of East Germany."* Amsterdam: Thesis Publishers
30. Anneke Cancrinus-Matthijsse (1995), *"Tussen hulpverlening en ondernemerschap: Beroep-suitoefening en taakopvattingen van openbare apothekers in een aantal West-Europese landen."* Amsterdam: Thesis Publishers
31. Nardi Steverink (1996), *"Zo lang mogelijk zelfstandig: Naar een verklaring van verschillen in oriëntatie ten aanzien van opname in een verzorgingstehuis onder fysiek kwetsbare ouderen."* Amsterdam: Thesis Publishers
32. Ellen Lindeman (1996), *"Participatie in vrijwilligerswerk."* Amsterdam: Thesis Publishers
33. Chris Snijders (1996), "*Trust and commitments.*" Amsterdam: Thesis Publishers
34. Koos Postma (1996), *"Changing prejudice in Hungary. A study on the collapse of state socialism and its impact on prejudice against Gypsies and Jews."* Amsterdam: Thesis Publishers
35. Jooske van Busschbach (1996), *"Uit het oog, uit het hart? Stabiliteit en verandering in persoonlijke relaties."* Amsterdam: Thesis Publishers
36. René Torenvlied (1996), *"Besluiten in uitvoering: Theorieën over beleidsuitvoering model-matig getoetst op sociale vernieuwing in drie gemeenten."* Amsterdam: Thesis Publishers

37. Andreas Flache (1996), *"The double edge of networks: An analysis of the effect of informal networks on cooperation in social dilemmas."* Amsterdam: Thesis Publishers
38. Kees van Veen (1997), *"Inside an internal labor market: Formal rules, flexibility and career lines in a Dutch manufacturing company."* Amsterdam: Thesis Publishers
39. Lucienne van Eijk (1997), *"Activity and wellbeing in the elderly."* Amsterdam: Thesis Publishers
40. Róbert Gál (1997), *"Unreliability: Contract discipline and contract governance under economic transition."* Amsterdam: Thesis Publishers
41. Anne-Geerte van de Goor (1997), *"Effects of regulation on disability duration."* ICS dissertation, Utrecht
42. Boris Blumberg (1997), *"Das Management von Technologiekooperationen: Partnersuche und Verhandlungen mit dem Partner aus Empirisch Theoretischer Perspektive."* ICS dissertation, Utrecht
43. Marijke von Bergh (1997), *"Loopbanen van oudere werknemers."* Amsterdam: Thesis Publishers
44. Anna Petra Nieboer (1997), *"Life events and well-being: A prospective study on changes in well-being of elderly people due to a serious illness event or death of the spouse."* Amsterdam: Thesis Publishers
45. Jacques Niehof (1997), *"Resources and social reproduction: The effects of cultural and material resources on educational and occupational careers in industrial nations at the end of the twentieth century."* ICS dissertation, Nijmegen
46. Ariana Need (1997), *"The kindred vote: Individual and family effects of social class and religion on electoral change in the Netherlands, 1956 1994."* ICS dissertation, Nijmegen
47. Jim Allen (1997), *"Sector composition and the effect of education on Wages: An international comparison."* Amsterdam: Thesis Publishers
48. Jack Hutten (1998), *"Workload and provision of care in general practice: An empirical study of the relation between workload of Dutch general practitioners and the content and quality of their care."* ICS dissertation, Utrecht
49. Per Kropp (1998), *"Berufserfolg im Transformationsprozeß: Eine theoretisch empirische Studie über die Gewinner und Verlierer der Wende in Ostdeutschland."* ICS dissertation, Utrecht
50. Maarten Wolbers (1998), *"Diploma-inflatie en verdringing op de arbeidsmarkt: Een studie naar ontwikkelingen in de opbrengsten van diploma's in Nederland."* ICS dissertation, Nijmegen
51. Wilma Smeenk (1998), *"Opportunity and marriage: The impact of individual resources and marriage market structure on first marriage timing and partner choice in the Netherlands."* ICS dissertation, Nijmegen
52. Marinus Spreen (1999), *"Sampling personal network structures: Statistical inference in ego graphs."* ICS dissertation, Groningen
53. Vincent Buskens (1999), *"Social networks and trust."* ICS dissertation, Utrecht
54. Susanne Rijken (1999), *"Educational expansion and status attainment: A cross-national and over-time comparison."* ICS dissertation, Utrecht

55. Mérove Gijsberts (1999), *"The legitimation of inequality in state-socialist and market societies, 1987-1996."* ICS dissertation, Utrecht
56. Gerhard van de Bunt (1999), *"Friends by choice: An actor-oriented statistical network model for friendship networks through time."* ICS dissertation, Groningen
57. Robert Thomson (1999), *"The party mandate: Election pledges and government actions in the Netherlands, 1986 1998."* Amsterdam: Thela Thesis
58. Corine Baarda (1999), *"Politieke besluiten en boeren beslissingen: Het draagvlak van het mestbeleid tot 2000."* ICS dissertation, Groningen
59. Rafael Wittek (1999), *"Interdependence and informal control in organizations."* ICS dissertation, Groningen
60. Diane Payne (1999), *"Policy Making in the European Union: An analysis of the impact of the reform of the structural funds in Ireland."* ICS dissertation, Groningen
61. René Veenstra (1999), *"Leerlingen, Klassen, Scholen: Prestaties en vorderingen van leerlingen in het voortgezet onderwijs."* Amsterdam, Thela Thesis
62. Marjolein Achterkamp (1999), *"Influence strategies in collective decision making: A comparison of two models."* ICS dissertation, Groningen
63. Peter Mühlau (2000), *"The governance of the employment relation: Arelational signaling perspective."* ICS dissertation, Groningen
64. Agnes Akkerman (2000), *"Verdeelde vakbeweging en stakingen: Concurrentie om leden."* ICS dissertation, Groningen
65. Sandra van Thiel (2000), *"Quangocratization: Trends, causes and consequences."* ICS dissertation, Utrecht
66. Sylvia Peacock-Korupp (2000), *"Mothers and the process of social stratification."* ICS dissertation, Utrecht
67. Rudi Turksema (2000), *"Supply of day care."* ICS dissertation, Utrecht
68. Bernard Nijstad (2000), *"How the group affects the mind: Effects of communication in idea generating groups."* ICS dissertation, Utrecht
69. Inge de Wolf (2000), *"Opleidingsspecialisatie en arbeidsmarktsucces van sociale wetenschappers."* ICS dissertation, Utrecht
70. Jan Kratzer (2001), *"Communication and performance: An empirical study in innovation teams."* ICS dissertation, Groningen
71. Madelon Kroneman (2001), *"Healthcare systems and hospital bed use."* ICS/NIVEL-dissertation, Utrecht
72. Herman van de Werfhorst (2001), *"Field of study and social inequality: Four types of educational resources in the process of stratification in the Netherlands."* ICS dissertation, Nijmegen
73. Tamás Bartus (2001), *"Social capital and earnings inequalities: The role of informal job search in Hungary."* ICS dissertation Groningen
74. Hester Moerbeek (2001), *"Friends and foes in the occupational career: The influence of sweet and sour social capital on the labour market."* ICS dissertation, Nijmegen
75. Marcel van Assen (2001), *"Essays on actor perspectives in exchange networks and social dilemmas."* ICS dissertation, Groningen

76. Inge Sieben (2001), *"Sibling similarities and social stratification: The impact of family background across countries and cohorts."* ICS dissertation, Nijmegen
77. Alinda van Bruggen (2001), *"Individual production of social well-being: An exploratory study."* ICS dissertation, Groningen
78. Marcel Coenders (2001), *"Nationalistic attitudes and ethnic exclusionism in a comparative perspective: An empirical study of attitudes toward the country and ethnic immigrants in 22 countries."* ICS dissertation, Nijmegen
79. Marcel Lubbers (2001), *"Exclusionistic electorates: Extreme right-wing voting in Western Europe."* ICS dissertation, Nijmegen
80. Uwe Matzat (2001), *"Social networks and cooperation in electronic communities: A theoretical-empirical analysis of academic communication and internet discussion groups."* ICS dissertation, Groningen
81. Jacques Janssen (2002), *"Do opposites attract divorce? Dimensions of mixed marriage and the risk of divorce in the Netherlands."* ICS dissertation, Nijmegen
82. Miranda Jansen (2002), *"Waardenoriëntaties en partnerrelaties: Een panelstudie naar wederzijdse invloeden."* ICS dissertation, Utrecht
83. Anne-Rigt Poortman (2002), *"Socioeconomic causes and consequences of divorce."* ICS dissertation, Utrecht
84. Alexander Gattig (2002), *"Intertemporal decision making."* ICS dissertation, Groningen
85. Gerrit Rooks (2000), *"Contract en conflict: Strategisch management van inkooptransacties."* ICS dissertation, Utrecht
86. Károly Takács (2002), *"Social networks and intergroup conflict."* ICS dissertation, Groningen
87. Thomas Gautschi (2002), *"Trust and exchange, effects of temporal embeddedness and network embeddedness on providing and dividing a surplus."* ICS dissertation, Utrecht
88. Hilde Bras (2002), *"Zeeuwse meiden: Dienen in de levensloop van vrouwen, ca. 1850 – 1950."* Aksant Academic Publishers, Amsterdam
89. Merijn Rengers (2002), *"Economic lives of artists: Studies into careers and the labour market in the cultural sector."* ICS dissertation, Utrecht
90. Annelies Kassenberg (2002), *"Wat scholieren bindt: Sociale gemeenschap in scholen."* ICS dissertation, Groningen
91. Marc Verboord (2003), *"Moet de meester dalen of de leerling klimmen? De invloed van literatuuronderwijs en ouders op het lezen van boeken tussen 1975 en 2000."* ICS dissertation, Utrecht
92. Marcel van Egmond (2003), *"Rain falls on all of us (but some manage to get more wet than others): Political context and electoral participation."* ICS dissertation, Nijmegen
93. Justine Horgan (2003), *"High-performance human resource management in Ireland and the Netherlands: Adoption and effectiveness."* ICS dissertation, Groningen
94. Corine Hoeben (2003), *"LETS' be a community: Community in local exchange trading systems."* ICS dissertation, Groningen
95. Christian Steglich (2003), *"The framing of decision situations: Automatic goal selection and rational goal pursuit."* ICS dissertation, Groningen

96. Johan van Wilsem (2003), *"Crime and context: The impact of individual, neighborhood, city and country characteristics on victimization."* ICS dissertation, Nijmegen
97. Christiaan Monden (2003), *"Education, inequality and health: The impact of partners and life course."* ICS dissertation, Nijmegen
98. Evelyn Hello (2003), *"Educational attainment and ethnic attitudes: How to explain their relationship."* ICS dissertation, Nijmegen
99. Marnix Croes en Peter Tammes (2004). *Gif laten wij niet voortbestaan: Een onderzoek naar de overlevingskansen van Joden in de Nederlandse gemeenten, 1940-1945.* Aksant Academic Publishers, Amsterdam.
100. Ineke Nagel (2004), *"Cultuurdeelname in de levensloop."* ICS dissertation, Utrecht
101. Marieke van der Wal (2004), *"Competencies to participate in life: Measurement and the impact of school."* ICS dissertation, Groningen
102. Vivian Meertens (2004), *"Depressive symptoms in the general population: A multifactorial social approach."* ICS dissertation, Nijmegen
103. Hanneke Schuurmans (2004), *"Promoting well-being in frail elderly people: Theory and intervention."* ICS dissertation, Groningen
104. Javier Arregui Moreno (2004), *"Negotiation in legislative decision-making in the European Union."* ICS dissertation, Groningen
105. Tamar Fischer (2004), *"Parental divorce, conflict and resources: The effects on children's behaviour problems, socioeconomic attainment, and transitions in the demographic career."* ICS dissertation, Nijmegen
106. René Bekkers (2004), *"Giving and volunteering in the Netherlands: Sociological and psychological perspectives."* ICS dissertation, Utrecht
107. Renée van der Hulst (2004), *"Gender differences in workplace authority: An empirical study on social networks."* ICS dissertation, Groningen
108. Rita Smaniotto (2004), *"'You scratch my back and I scratch yours' versus 'love thy neighbour': Two proximate mechanisms of reciprocal altruism."* ICS dissertation, Groningen
109. Maurice Gesthuizen (2004), *"The life course of the low-educated in the Netherlands: Social and economic risks."* ICS dissertation, Nijmegen
110. Carlijne Philips (2005), *"Vakantiegemeenschappen: Kwalitatief en kwantitatief onderzoek naar gelegenheid en refreshergemeenschap tijdens de vakantie."* ICS dissertation, Groningen
111. Esther de Ruijter (2005), *"Household outsourcing."* ICS dissertation, Utrecht
112. Frank van Tubergen (2005), *"The integration of immigrants in cross-national perspective: Origin, destination, and community effects."* ICS dissertation, Utrecht
113. Ferry Koster (2005), *"For the time being: Accounting for inconclusive findings concerning the effects of temporary employment relationships on solidary behavior of employees."* ICS dissertation, Groningen
114. Carolien Klein Haarhuis (2005), *"Promoting anti-corruption reforms: Evaluating the implementation of a World Bank anti-corruption program in seven African countries (1999-2001)."* ICS dissertation, Utrecht

115. Martin van der Gaag (2005), *"Measurement of individual social capital."* ICS dissertation, Groningen
116. Johan Hansen (2005), *"Shaping careers of men and women in organizational contexts."* ICS dissertation, Utrecht
117. Davide Barrera (2005), *"Trust in embedded settings."* ICS dissertation, Utrecht
118. Mattijs Lambooij (2005), *"Promoting cooperation: Studies into the effects of long-term and short-term rewards on cooperation of employees."* ICS dissertation, Utrecht
119. Lotte Vermeij (2006), *"What's cooking? Cultural boundaries among Dutch teenagers of different ethnic origins in the context of school."* ICS dissertation, Utrecht
120. Mathilde Strating (2006), *"Facing the challenge of rheumatoid arthritis: A 13-year prospective study among patients and cross-sectional study among their partners."* ICS dissertation, Groningen
121. Jannes de Vries (2006), *"Measurement error in family background variables: The bias in the intergenerational transmission of status, cultural consumption, party preference, and religiosity."* ICS dissertation, Nijmegen
122. Stefan Thau (2006), *"Workplace deviance: Four studies on employee motives and self-regulation."* ICS dissertation, Groningen
123. Mirjam Plantinga (2006), *"Employee motivation and employee performance in child care: The effects of the introduction of market forces on employees in the Dutch child-care sector."* ICS dissertation, Groningen
124. Helga de Valk (2006), *"Pathways into adulthood: A comparative study on family life transitions among migrant and Dutch youth."* ICS dissertation, Utrecht
125. Henrike Elzen (2006), *"Self-Management for chronically ill older people."* ICS dissertation, Groningen
126. Ayse Güveli (2007), *"New social classes within the service class in the Netherlands and Britain: Adjusting the EGP class schema for the technocrats and the social and cultural specialists."* ICS dissertation, Nijmegen
127. Willem-Jan Verhoeven (2007), *"Income attainment in post-communist societies."* ICS dissertation, Utrecht
128. Marieke Voorpostel (2007), *"Sibling support: The exchange of help among brothers and sisters in the Netherlands."* ICS dissertation, Utrecht
129. Jacob Dijkstra (2007), *"The effects of externalities on partner choice and payoffs in exchange networks."* ICS dissertation, Groningen
130. Patricia van Echtelt (2007), *"Time-greedy employment relationships: Four studies on the time claims of post-Fordist work."* ICS dissertation, Groningen
131. Sonja Vogt (2007), *"Heterogeneity in social dilemmas: The case of social support."* ICS dissertation, Utrecht
132. Michael Schweinberger (2007), *"Statistical methods for studying the evolution of networks and behavior."* ICS dissertation, Groningen
133. István Henrik Back (2007), *"Commitment and evolution: Connecting emotion and reason in long-term relationships."* ICS dissertation, Groningen

134. Ruben van Gaalen (2007), *"Solidarity and ambivalence in parent-child relationships."* ICS dissertation, Utrecht
135. Jan Reitsma (2007), *"Religiosity and solidarity: Dsimensions and relationships disentangled and tested."* ICS dissertation, Nijmegen
136. Jan Kornelis Dijkstra (2007), *"Status and affection among (pre)adolescents and their relation with antisocial and prosocial behavior."* ICS dissertation, Groningen
137. Wouter van Gils (2007), *"Full-time working couples in the Netherlands: Causes and consequences."* ICS dissertation, Nijmegen
138. Djamila Schans (2007), *"Ethnic diversity in intergenerational solidarity."* ICS dissertation, Utrecht
139. Ruud van der Meulen (2007), *"Brug over woelig water: Lidmaatschap van sportverenigingen, vriendschappen, kennissenkringen en veralgemeend vertrouwen."* ICS dissertation, Nijmegen
140. Andrea Knecht (2008), *"Friendship selection and friends' influence: Dynamics of networks and actor attributes in early adolescence."* ICS dissertation, Utrecht
141. Ingrid Doorten (2008), *"The division of unpaid work in the household: A stubborn pattern?"* ICS dissertation, Utrecht
142. Stijn Ruiter (2008), *"Association in context and association as context: Causes and consequences of voluntary association involvement."* ICS dissertation, Nijmegen
143. Janneke Joly (2008), *"People on our minds: When humanized contexts activate social norms."* ICS dissertation, Groningen
144. Margreet Frieling (2008), *"'Joint production' als motor voor actief burgerschap in de buurt."* ICS dissertation, Groningen
145. Ellen Verbakel (2008), *"The partner as resource or restriction? Labour market careers of husbands and wives and the consequences for inequality between couples."* ICS dissertation, Nijmegen
146. Gijs van Houten (2008), *"Beleidsuitvoering in gelaagde stelsels: De doorwerking van aanbevelingen van de Stichting van de Arbeid in het CAO-overleg."* ICS dissertation, Utrecht
147. Eva Jaspers (2008), *"Intolerance over time: Macro- and micro-level questions on attitudes towards euthanasia, homosexuality and ethnic minorities."* ICS dissertation, Nijmegen
148. Gijs Weijters (2008), *"Youth delinquency in Dutch cities and schools: A multilevel approach."* ICS dissertation, Nijmegen
149. Jessica Nooij (2009), *"The self in social rejection."* ICS dissertation, Groningen
150. Gerald Mollenhorst (2009), *"Networks in contexts: How meeting opportunities affect personal relationships."* ICS dissertation, Utrecht
151. Tom van der Meer (2009), *"States of freely associating citizens: Comparative studies into the impact of state institutions on social, civic and political participation."* ICS dissertation, Nijmegen
152. Manuela Vieth (2009), *"Commitments and reciprocity in trust situations: Experimental studies on obligation, indignation, and self-consistency."* ICS dissertation, Utrecht
153. Rense Corten (2009), *"Co-evolution of social networks and behavior in social dilemmas: Theoretical and empirical perspectives."* ICS dissertation, Utrecht

A

154. Arieke Rijken (2009), *"Happy families, high fertility? Childbearing choices in the context of family and partner relationships."* ICS dissertation, Utrecht
155. Jochem Tolsma (2009), *"Ethnic hostility among ethnic majority and minority groups in the Netherlands: An investigation into the impact of social mobility experiences, the local living environment and educational attainment on ethnic hostility."* ICS dissertation, Nijmegen
156. Freek Bucx (2009), *"Linked lives: Young adults' life course and relations with parents."* ICS dissertation, Utrecht
157. Philip Wotschack (2009), *"Household governance and time allocation: Four studies on the combination of work and care."* ICS dissertation, Groningen
158. Nienke Moor (2009), *"Explaining worldwide religious diversity: The relationship between subsistence technologies and ideas about the unknown in pre-industrial and (post-)industrial societies."* ICS dissertation, Nijmegen
159. Lieke ten Brummelhuis (2009), *"Family matters at work: Depleting and enriching effects of employees' family lives on work outcomes."* ICS dissertation, Utrecht
160. Renske Keizer (2010), *"Remaining childless: Causes and consequences from a life-course perspective."* ICS dissertation, Utrecht
161. Miranda Sentse (2010), *"Bridging contexts: The interplay between family, child, and peers in explaining problem behavior in early adolescence."* ICS dissertation, Groningen
162. Nicole Tieben (2010), *"Transitions, tracks and transformations: Social inequality in transitions into, through and out of secondary education in the Netherlands for cohorts born between 1914 and 1985."* ICS dissertation, Nijmegen
163. Birgit Pauksztat (2010), *"Speaking up in organizations: Four studies on employee voice."* ICS dissertation, Groningen
164. Richard Zijdeman (2010), *"Status attainment in the Netherlands, 1811-1941: Spatial and temporal variation before and during industrialization."* ICS dissertation, Utrecht
165. Rianne Kloosterman (2010), *"Social background and children's educational careers: The primary and secondary effects of social background over transitions and over time in the Netherlands."* ICS dissertation, Nijmegen
166. Olav Aarts (2010), *"Religious diversity and religious involvement: A study of religious markets in Western societies at the end of the twentieth century."* ICS dissertation, Nijmegen
167. Stephanie Wiesmann (2010), *"24/7 Negotiation in couples' transition to parenthood."* ICS dissertation, Utrecht
168. Borja Martinovic (2010), *"Interethnic contacts: A dynamic analysis of interaction between immigrants and natives in Western countries."* ICS dissertation, Utrecht
169. Anne Roeters (2010), *"Family life under pressure? Parents' paid work and the quantity and quality of parent-child and family time."* ICS dissertation, Utrecht
170. Jelle Sijtsema (2010), *"Adolescent aggressive behavior: Status and stimulation goals in relation to the peer context."* ICS dissertation, Groningen
171. Kees Keizer (2010), *"The spreading of disorder."* ICS dissertation, Groningen
172. Michael Mäs (2010), *"The diversity puzzle. explaining clustering and polarization of opinions."* ICS dissertation, Groningen

173. Marie-Louise Damen (2010), *"Cultuurdeelname en CKV: Studies naar effecten van kunsteducatie op de cultuurdeelname van leerlingen tijdens en na het voortgezet onderwijs."* ICS dissertation, Utrecht
174. Marieke van de Rakt (2011), *"Two generations of crime: The intergenerational transmission of convictions over the life course."* ICS dissertation, Nijmegen
175. Willem Huijnk (2011), *"Family life and ethnic attitudes: The role of the family for attitudes towards intermarriage and acculturation among minority and majority groups."* ICS dissertation, Utrecht
176. Tim Huijts (2011), *"Social ties and health in Europe: Individual associations, cross-national variations, and contextual explanations."* ICS dissertation, Nijmegen
177. Wouter Steenbeek (2011), *"Social and physical disorder: How community, business presence and entrepreneurs influence disorder in Dutch neighborhoods."* ICS dissertation, Utrecht
178. Miranda Vervoort (2011), *"Living together apart? Ethnic concentration in the neighborhood and ethnic minorities' social contacts and language practices."* ICS dissertation, Utrecht
179. Agnieszka Kanas (2011), *"The economic performance of immigrants: The role of human and social capital."* ICS dissertation, Utrecht
180. Lea Ellwardt (2011), *"Gossip in organizations: A social network study."* ICS dissertation, Groningen
181. Annemarije Oosterwaal (2011), *"The gap between decision and implementation: Decision making, delegation and compliance in governmental and organizational settings."* ICS dissertation, Utrecht
182. Natascha Notten (2011), *"Parents and the media: Causes and consequences of parental media socialization."* ICS dissertation, Nijmegen
183. Tobias Stark (2011), *"Integration in schools: A process perspective on students' interethnic attitudes and interpersonal relationships."* ICS dissertation, Groningen
184. Giedo Jansen (2011), *"Social cleavages and political choices: Large-scale comparisons of social class, religion and voting behavior in Western democracies."* ICS dissertation, Nijmegen
185. Ruud van der Horst (2011), *"Network effects on treatment results in a closed forensic psychiatric setting."* ICS dissertation, Groningen
186. Mark Levels (2011), *"Abortion laws in European Countries between 1960 and 2010: Legislative developments and their consequences for women's reproductive decision making."* ICS dissertation, Nijmegen
187. Marieke van Londen (2012), *"Exclusion of ethnic minorities in the Netherlands: The effects of individual and situational characteristics on opposition to ethnic policy and ethnically mixed neighbourhoods."* ICS dissertation, Nijmegen
188. Sigrid Mohnen (2012), *"Neighborhood context and health: How neighborhood social capital affects individual health."* ICS dissertation, Utrecht
189. Asya Zhelyazkova (2012), *"Compliance under controversy: Analysis of the transposition of European directives and their provisions."* ICS dissertation, Utrecht
190. Valeska Korff (2012), *"Between cause and control: Management in a humanitarian organization."* ICS dissertation, Groningen

191. Maike Gieling (2012), *"Dealing with diversity: Adolescents' support for civil liberties and immigrant rights."* ICS dissertation, Utrecht
192. Katya Ivanova (2012), *"From parents to partners: The impact of family on romantic relationships in adolescence and emerging adulthood."* ICS dissertation, Groningen
193. Jelmer Schalk (2012), *"The performance of public corporate actors: Essays on effects of institutional and network embeddedness in supranational, national, and local collaborative contexts."* ICS dissertation, Utrecht
194. Alona Labun (2012), *"Social networks and informal power in organizations."* ICS dissertation, Groningen
195. Michal Bojanowski (2012), *"Essays on social network formation in heterogeneous populations: Models, methods, and empirical analyses."* ICS dissertation, Utrecht
196. Anca Minescu (2012), *"Relative group position and intergroup attitudes in Russia."* ICS dissertation, Utrecht
197. Marieke van Schellen (2012), *"Marriage and crime over the life course: The criminal careers of convicts and their spouses."* ICS dissertation, Utrecht
198. Mieke Maliepaard (2012), *"Religious trends and social integration: Muslim minorities in the Netherlands."* ICS dissertation, Utrecht
199. Fransje Smits (2012), *"Turks and Moroccans in the Low Countries around the year 2000: Determinants of religiosity, trend in religiosity and determinants of the trend."* ICS dissertation, Nijmegen
200. Roderick Sluiter (2012), *"The diffusion of morality policies among Western European countries between 1960 and 2010: A comparison of temporal and spatial diffusion patterns of six morality and eleven non-morality policies."* ICS dissertation, Nijmegen
201. Nicoletta Balbo (2012), *"Family, friends and fertility."* ICS dissertation, Groningen
202. Anke Munniksma (2013), *"Crossing ethnic boundaries: Parental resistance to and consequences of adolescents' cross-ethnic peer relations"* ICS dissertation, Groningen
203. Anja-Kristin Abendroth (2013), *"Working women in Europe: How the country, workplace, and family context matter."* ICS dissertation, Utrecht
204. Katia Begall (2013), *"Occupational hazard? The relationship between working conditions and fertility."* ICS dissertation, Groningen
205. Hidde Bekhuis (2013), *"The popularity of domestic cultural products: Cross-national differences and the relation to globalization."* ICS dissertation, Utrecht
206. Lieselotte Blommaert (2013), *"Are Joris and Renske more employable than Rashid and Samira? A study on the prevalence and sources of ethnic discrimination in recruitment in the Netherlands using experimental and survey data."* ICS dissertation, Utrecht
207. Wiebke Schulz (2013), *"Careers of men and women in the 19th and 20th centuries."* ICS dissertation, Utrecht
208. Ozan Aksoy (2013), *"Essays on social preferences and beliefs in non-embedded social dilemmas."* ICS dissertation, Utrecht
209. Dominik Morbitzer (2013), *"Limited farsightedness in network formation."* ICS dissertation, Utrecht

210. Thomas de Vroome (2013), *"Earning your place: The relation between immigrants' economic and psychological integration in the Netherlands."* ICS dissertation, Utrecht
211. Marloes de Lange (2013), *"Causes and consequences of employment flexibility among young people: Recent developments in the Netherlands and Europe."* ICS dissertation, Nijmegen
212. Roza Meuleman (2014), *"Consuming the Nation: Domestic cultural consumption: Its stratification and relation with nationalist attitudes."* ICS dissertation, Utrecht
213. Esther Havekes (2014), *"Putting interethnic attitudes in context: The relationship between neighbourhood characteristics, interethnic attitudes and residential behaviour."* ICS dissertation, Utrecht
214. Zoltán Lippényi (2014), *"Transitions toward an open society? Intergenerational occupational mobility in Hungary in the 19th and 20th centuries."* ICS dissertation, Utrecht
215. Anouk Smeekes (2014), *"The presence of the past: Historical rooting of national identity and current group dynamics."* ICS dissertation, Utrecht
216. Michael Savelkoul (2014), *"Ethnic diversity and social capital: Testing underlying explanations derived from conflict and contact theories in Europe and the United States."* ICS dissertation, Nijmegen
217. Martijn Hogerbrugge (2014), *"Misfortune and family: How negative events, family ties, and lives are linked."* ICS dissertation, Utrecht
218. Gina-Felicia Potarca (2014), *"Modern love: Comparative insights in online dating preferences and assortative mating."* ICS dissertation, Groningen
219. Mariska van der Horst (2014), *"Gender, aspirations, and achievements: Relating work and family aspirations to occupational outcomes."* ICS dissertation, Utrecht
220. Gijs Huitsing (2014), *"A social network perspective on bullying"* ICS dissertation, Groningen
221. Thomas Kowalewski (2015), *"Personal growth in organizational contexts."* ICS dissertation, Groningen
222. Manuel Muñoz-Herrera (2015), *"The impact of individual differences on network relations: Social exclusion and inequality in productive exchange and coordination games."* ICS dissertation, Groningen
223. Tim Immerzeel (2015), *"Voting for a change: The democratic lure of populist radical right parties in voting behavior."* ICS dissertation, Utrecht
224. Fernando Nieto Morales (2015), *"The control imperative: Studies on reorganization in the public and private sectors."* ICS dissertation, Groningen
225. Jellie Sierksma (2015), *"Bounded helping: How morality and intergroup relations shape children's reasoning about helping."* ICS dissertation, Utrecht
226. Tinka Veldhuis (2015), *"Captivated by fear: An evaluation of terrorism detention policy."* ICS dissertation, Groningen
227. Miranda Visser (2015), *"Loyality in humanity: Turnover among expatriate humanitarian aid workers."* ICS dissertation, Groningen
228. Sarah Westphal (2015), *"Are the kids alright? Essays on postdivorce residence arrangements and children's well-being."* ICS dissertation, Utrecht
229. Britta Rüschoff (2015), *"Peers in careers: Peer relationships in the transition from school to work."* ICS dissertation, Groningen

230. Nynke van Miltenburg (2015), *"Cooperation under peer sanctioning institutions: Collective decisions, noise, and endogenous implementation."* ICS dissertation, Utrecht
231. Antonie Knigge (2015), *"Sources of sibling similarity: Status attainment in the Netherlands during modernization."* ICS dissertation, Utrecht
232. Sanne Smith (2015), *"Ethnic segregation in friendship networks: Studies of its determinants in English, German, Dutch, and Swedish school classes."* ICS dissertation, Utrecht
233. Patrick Präg (2015), *"Social stratification and health: Four essays on social determinants of health and wellbeing."* ICS dissertation, Groningen
234. Wike Been (2015), *"European top managers' support for work-life arrangements"* ICS dissertation, Utrecht
235. André Grow (2016), *"Status differentiation: New insights from agent-based modeling and social network analysis."* ICS dissertation, Groningen
236. Jesper Rözer (2016), *"Family and personal networks: How a partner and children affect social relationships."* ICS dissertation, Utrecht
237. Kim Pattiselanno (2016), *"At your own risk: The importance of group dynamics and peer processes in adolescent peer groups for adolescents' involvement in risk behaviors."* ICS-dissertation, Groningen
238. Vincenz Frey (2016), *"Network formation and trust."* ICS dissertation, Utrecht
239. Rozemarijn van der Ploeg (2016), *"Be a buddy, not a bully? Four studies on social and emotional processes related to bullying, defending, and victimization."* ICS dissertation, Groningen
240. Tali Spiegel (2016), *"Identity, career trajectories and wellbeing: A closer look at individuals with degenerative eye conditions."* ICS- dissertation, Groningen
241. Felix Christian Tropf (2016), *"Social science genetics and fertility."* ICS dissertation, Groningen
242. Sara Geven (2016), *"Adolescent problem behavior in school: The role of peer networks."* ICS dissertation, Utrecht
243. Josja Rokven (2016), *"The victimization-offending relationship from a longitudinal perspective."* ICS dissertation, Nijmegen
244. Maja Djundeva (2016), *"Healthy ageing in context: Family welfare state and the life course."* ICS dissertation, Groningen
245. Mark Visser (2017), *"Inequality between older workers and older couples in the Netherlands: A dynamic life course perspective on educational and social class differences in the late career."* ICS dissertation, Nijmegen
246. Beau Oldenburg (2017), *"Bullying in schools: The role of teachers and classmates."* ICS dissertation, Groningen
247. Tatang Muttaqin (2017), *"The education divide in Indonesia: Four essays on determinants of unequal access to and quality of education."* ICS dissertation, Groningen
248. Margriet van Hek (2017), *"Gender inequality in educational attainment and reading performance: A contextual approach."* ICS dissertation, Nijmegen

249. Melissa Verhoef-van Dorp (2017), *"Work schedules, childcare and well-being: Essays on the associations between modern-day job characteristics, childcare arrangements and the well-being of parents and children."* ICS dissertation, Utrecht
250. Timo Septer (2017), *"Goal priorities, cognition and conflict: Analyses of cognitive maps concerning organizational change."* ICS dissertation, Groningen
251. Bas Hofstra (2017), *"Online social networks: Essays on membership, privacy, and structure."* ICS dissertation, Utrecht
252. Yassine Khoudja (2018), *"Women's labor market participation across ethnic groups: The role of household conditions, gender role attitudes, and religiosity in different national contexts."* ICS dissertation, Utrecht
253. Joran Laméris (2018), *"Living together in diversity: Whether, why and where ethnic diversity affects social cohesion."* ICS dissertation, Nijmegen
254. Maaike van der Vleuten (2018), *"Gendered Choices: Fields of study of adolescents in the Netherlands."* ICS dissertation, Utrecht
255. Mala Sondang Silitonga (2018), *"Corruption in Indonesia: The impact of institutional change, norms, and networks."* ICS dissertation, Groningen
256. Manja Coopmans (2018), *"Rituals of the past in the context of the present: The role of Remembrance Day and Liberation Day in Dutch society."* ICS dissertation, Utrecht
257. Paul Hindriks (2018), *"The struggle for power: Attitudes towards the political participation of ethnic minorities."* ICS dissertation, Utrecht
258. Nynke Niezink (2018), *"Modeling the dynamics of networks and continuous behavior."* ICS dissertation, Groningen
259. Simon de Bruijn (2018), *"Reaching agreement after divorce and separation: Essays on the effectiveness of parenting plans and divorce mediation."* ICS dissertation, Utrecht
260. Susanne van 't Hoff-de Goede (2018), *"While you were locked up: An empirical study on the characteristics, social surroundings and wellbeing of partners of prisoners in The Netherlands."* ICS dissertation, Utrecht
261. Loes van Rijsewijk (2018), *"Antecedents and consequences of helping among adolescents."* ICS dissertation, Groningen
262. Mariola Gremmen (2018), *"Social network processes and academic functioning: The role of peers in students' school well-being, academic engagement, and academic achievement."* ICS dissertation, Groningen
263. Jeanette Renema (2018), *"Immigrants' support for welfare spending: The causes and consequences of welfare usage and welfare knowledgeability."* ICS dissertation, Nijmegen
264. Suwatin Miharti (2018), *"Community health centers in Indonesia in the era of decentralization: The impact of structure, staff composition and management on health outcomes."* ICS dissertation, Groningen
265. Chaïm la Roi (2019), *"Stigma and stress: Studies on attitudes towards sexual minority orientations and the association between sexual orientation and mental health."* ICS dissertation, Groningen
266. Jelle Lössbroek (2019), *"Turning grey into gold: Employer-employee interplay in an ageing workforce."* ICS dissertation, Utrecht

A

267. Nikki van Gerwen (2019), *"Employee cooperation through training: A multi-method approach."* ICS dissertation, Utrecht
268. Paula Thijs (2019), *"Trends in cultural conservatism: The role of educational expansion, secularisation, and changing national contexts."* ICS dissertation, Nijmegen
269. Renske Verweij (2019), *"Understanding childlessness: Unravelling the link with genes and the socio-environment."* ICS dissertation, Groningen
270. Niels Blom (2019), *"Partner relationship quality under pressing work conditions: Longitudinal and cross-national investigation."* ICS dissertation, Nijmegen
271. Müge Simsek (2019), *"The dynamics of religion among native and immigrant youth in Western Europe."* ICS dissertation, Utrecht
272. Leonie van Breeschoten (2019), *"Combining a career and childcare: The use and usefulness of work-family policies in European organizations."* ICS dissertation, Utrecht
273. Roos van der Zwan (2019), "*The political representation of ethnic minorities and their vote choice."* ICS dissertation, Nijmegen
274. Ashwin Rambaran (2019), "*The classroom as context for bullying: A social network approach*." ICS dissertation, Groningen
275. Dieko Bakker (2019), *"Cooperation and social control: Effects of preferences, institutions, and social structure."* ICS dissertation, Groningen
276. Femke van der Werf (2019), "*Shadow of a rainbow? National and ethnic belonging in Mauritius*." ICS dissertation, Utrecht
277. Robert Krause (2019), *"Multiple imputation for missing network data."* ICS dissertation, Groningen
278. Take Sipma (2020), *"Economic insecurity and populist radical right voting."* ICS dissertation, Nijmegen
279. Mathijs Kros (2020), *"The nature of negative contact: Studies on interethnic relations in Western societies."* ICS dissertation, Utrecht
280. Lonneke van den Berg (2020), *"Time to leave: Individual and contextual explanations for the timing of leaving home."* ICS dissertation, Amsterdam
281. Marianne Hooijsma (2020), "*Clashrooms: Interethnic peer relationships in schools."* ICS dissertation, Groningen
282. Marina Tulin (2020*), "Blind spots in social resource theory: Essays on the creation, maintenance and returns of social capital."* ICS dissertation, Amsterdam
283. Tessa Kaufman (2020), "*Toward tailored interventions: Explaining, assessing, and preventing persistent victimization of bullying."* ICS dissertation, Groningen
284. Lex Thijssen (2020), *"Racial and ethnic discrimination in western labor markets: Empirical evidence from field experiments."* ICS dissertation, Utrecht
285. Lukas Norbutas (2020), *"Trust on the dark web: An analysis of illegal online drug markets."* ICS dissertation, Utrecht
286. Tomáš Diviák (2020), *"Criminal networks: Actors, mechanisms, and structures."* ICS dissertation, Groningen
287. Tery Setiawan (2020), "*Support for interreligious conflict in Indonesia*." ICS dissertation, Nijmegen

288. Vera de Bel (2020), "*The ripple effect in family networks: Relational structures and well-being in divorced and non-divorced families*." ICS dissertation, Groningen
289. Diego Palacios (2020), *"How context and the perception of peers' behaviors shape relationships in adolescence: A multiplex social network perspective."* ICS dissertation, Groningen
290. Saskia Glas (2020), *"Where are the Muslim Feminists? Religiosity and support for gender equality in the Arab region*." ICS dissertation, Nijmegen
291. Tomas Turner-Zwinkels (2020), *"A new macro-micro approach to the study of political careers: Theoretical, Methodological and empirical challenges and solutions."* ICS dissertation, Groningen
292. Lotte Scheeren (2020), "*Not on the same track? Tracking age and gender inequality in education."* ICS dissertation, Amsterdam
293. Joris Broere (2020), "*Essays on how social network structure affects asymmetric coordination and trust*." ICS dissertation, Utrecht
294. Marcus Kristiansen (2021*), "Contact with benefits: How social networks affect benefit receipt dynamics in the Netherlands."* ICS dissertation, Utrecht
295. Judith Kas (2021), "*Trust and reputation in the peer-to-peer platform economy*." ICS dissertation, Utrecht
296. Andrea Forster (2021), "*Navigating educational institutions: Mechanisms of educational inequality and social mobility in different educational systems."* ICS dissertation, Amsterdam
297. Jannes ten Berge (2021), *"Technological change and work: The relation between technology implementation within organizations and changes in workers' employment."* ICS dissertation, Utrecht
298. Jolien Geerlings (2021), "*Teaching in culturally diverse classrooms: The importance of dyadic relations between teachers and children*." ICS dissertation, Utrecht
299. Kirsten van Houdt (2021), "*Stepfamilies in adulthood: Solidarity between parents and adult children*." ICS dissertation, Amsterdam
300. Suzanne de Leeuw (2021), "*The intergenerational transmission of educational attainment after divorce and remarriage."* ICS dissertation, Amsterdam
301. Fleur Goedkoop (2021), "*Involvement in bottom-up energy transitions: The role of local and contextual embeddedness*." ICS dissertation, Groningen
302. Eva Vriens (2021), "*Mutualism in the 21st century: The why, when, and how behind successful risk-sharing institutions*." ICS dissertation, Utrecht
303. Ardita Muja (2021), "*From school to work: The role of the vocational specificity of education in young people's labor market integration*." ICS dissertation, Nijmegen
304. Siyang Kong (2021), "*Women and work in contemporary China: The effect of market transition on women's employment, earnings, and status attainment*." ICS dissertation, Utrecht
305. Marijn Keijzer (2022), *"Opinion dynamics in online social media."* ICS dissertation, Groningen
306. Sander Kunst (2022), *"The educational divide in openness towards globalisation in Western Europe."* ICS dissertation, Amsterdam

A

307. Nella Geurts (2022), "*Puzzling pathways: The integration paradox among migrants in Western Europe."* ICS dissertation, Nijmegen
308. Dragana Stojmenovska (2022), *"Men's place: The incomplete integration of women in workplace authority."* ICS dissertation, Amsterdam
309. Bram Hogendoorn (2022), *"Divorce and inequality: Stratification in the risk and consequences of union dissolution."* ICS dissertation, Amsterdam
310. Tom Nijs (2022), *"This place is ours: Collective psychological ownership and its social consequences."* ICS dissertation, Utrecht
311. Nora Storz (2022), *"'This land is ours—but is it also theirs?' Collective ownership beliefs and reconciliation in territorial conflict regions."* ICS dissertation, Utrecht
312. Tara Koster (2022), *"Parenting and fairness in diverse families."* ICS dissertation, Utrecht
313. Danelien van Aalst (2022), "*Elements Contributing to Teachers' Role in Bullying."* ICS dissertation, Groningen
314. Wybren Nooitgedagt (2022), *"Who owns the country? Collective psychological ownership and intergroup relations in settler societies"* ICS dissertation, Utrecht
315. Marija Dangubić (2022), *"Rejecting Muslim minority practices: Principles and prejudices"* ICS dissertation, Utrecht
316. Sara Cvetkovska (2022), *"Lines in the shifting sand: The implications of being tolerated"* ICS dissertation, Utrecht
317. Maaike Hornstra (2022), "*Going beyond the dyad: Adult intergenerational closeness after divorce and remarriage"* ICS dissertation, Amsterdam
318. Wouter Kiekens (2022), *"Sexual and gender minority youth's mental health and substance use: Disparities, mechanisms, and protective factors"* ICS dissertation, Groningen
319. Carlijn Bussemakers (2022), *"Adversity and educational inequality: The interplay between adverse experiences and parental resources for children's educational attainment."* ICS dissertation, Nijmegen
320. Evi Velthuis (2022), " '*To tolerate or not to tolerate?' Reasons for tolerance of minority group practices among majority members in the Netherlands and Germany."* ICS dissertation, Utrecht
321. Hendrik Nunner (2023), *"Co-evolution of social networks and infectious diseases."* ICS dissertation, Utrecht
322. Carly van Mensvoort (2023), *"Inspiring leaders: An empirical study of female supervisors at work."* ICS dissertation, Nijmegen
323. Anne van der Put (2023), *"Healthy at work: The role of the work environment in worksite health promotion."* ICS dissertation, Utrecht
324. Rowan ten Kate (2023), "*Understanding loneliness among older migrants.*" ICS dissertation, Groningen
325. Sanne Kellij (2023), "*I see, I see what you don't see: Neural and behavioral social-cognitive processes underlying (persistent) victimization.*" ICS dissertation, Groningen
326. Inge Hendriks (2023), "*Changes and contrasts in attitudes towards ethnic minorities.*" ICS dissertation, Nijmegen

327. Eleonora Marucci (2023), "*Antecedents and consequences of teacher attunement in primary and secondary school classrooms.*" ICS dissertation, Groningen
328. Kasper Otten (2023), "*Cooperation in changing groups: How newcomers and norms shape public good provision in the lab, online games, and the field.*" ICS dissertation, Utrecht
329. Carlos de Matos Fernandes (2023), "*In or out? The paradox of exclusionary mechanisms in keeping cooperation going.*" ICS dissertation, Groningen

ABOUT THE AUTHOR

Carlos Antonio de Matos Fernandes was born in The Hague, the Netherlands, on the 7th of September, 1990. He obtained his Bachelor of Science in Sociology in 2016 at the University of Groningen. His bachelor thesis was on political distrust, perceived ethnic threat, and radical right-wing voting. After graduating, Carlos enrolled in the Research Master's program of the Faculty Behavioral and Social Sciences at the University of Groningen, focusing on Sociology. Carlos authored his Research Master's thesis on the interrelatedness of heterogeneity in valuing collective goods, social control, opposition to social control, and cooperation. He obtained his Master of Science (Research) in 2018. In 2018, he started his Ph.D. research at the Interuniversity Center for Social Science Theory and Methodology (ICS) at the Department of Sociology of the University of Groningen, the Netherlands. At the same time, Carlos was embedded in the Norms and Networks Cluster (NNC) and the research and training center dedicated to the interdisciplinary study of sustainable cooperation as a critical feature of resilient societies (SCOOP). Carlos published work in *Current Opinion in Psychology*, *Research Handbook of Analytical Sociology*, *Journal of Artificial Societies and Social Simulation*, and *Judgment and Decision Making*. As of 2022, Carlos works as a senior data scientist at DataFryslân.